
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 4376–4382, DOI 10.12694/scpe.v25i5.3136

THE APPLICATION OF INFORMATION TECHNOLOGY FOR ATHLETE DATA ANALYSIS AND AUTOMATIC GENERATION OF TRAINING PLANS

SHULI YUAN*

Abstract. In response to the demand for scientific training of sports athletes, the author combined data mining technology to study an improved sports training mode decision support evaluation system. In this regard, the author analyzed the characteristics of association rule algorithms and elaborated on their functions in data preprocessing, data mining, and pattern evaluation. Based on the software design of decision support systems, the characteristics of system operation were analyzed. At the same time, the author focused on explaining the data fusion processing of association rules in sports evaluation decision support systems, and proposed an improved Apriori algorithm output mode to improve the effectiveness of system evaluation. Compared with other algorithms such as Apriori, DC Apriori and Apriori, this algorithm has higher reliability. When the minimum confidence is increased, the advantage of prior information will gradually disappear, and the final result will be obtained. Experimental results show that this method can effectively provide support for sports training decision-making.

Key words: Data mining technology, Association rules, Sports training evaluation, Data fusion processing

1. Introduction. In modern sports competition, the analysis of athlete data and the formulation of training plans are crucial for improving competitive level and achieving excellent results. With the development of technology and continuous innovation in data collection technology, more and more athletes and coaches are using information technology to collect, analyze, and apply sports data. The training plan for athletes has gradually shifted from subjective experience in the past to objective decision-making based on data, making training more scientific and efficient. The application of data analysis in sports competitions has become a trend. By analyzing athlete data, coaches can understand their performance during training and competition, identify their strengths and weaknesses, and then develop targeted training plans [11]. For example, in basketball games, coaches can evaluate a player's performance in offense and defense by analyzing their shooting percentage, rebounds, and assist data, and provide targeted technical and tactical training. Through data analysis, coaches can discover the potential abilities of athletes and assist them in personalized training to improve their competitive level.

Data analysis can not only help coaches develop training plans, but also help athletes understand their performance and improve their space. By analyzing their own sports data, athletes can gain a deeper understanding of their strengths and weaknesses, identify their problems in the competition, and find ways to improve. For example, in track and field competitions, athletes can analyze their speed, endurance, and technical data to identify their weaknesses in training and conduct targeted training to improve their competitive level. Data analysis can enable athletes to have a more comprehensive understanding of their performance and potential, thereby formulating more scientific and effective training plans. In addition to data analysis, the application of information technology in sports training also includes data collection and application [4]. With the continuous advancement of technology, athletes can use various sensors and devices to collect exercise data, such as heart rate, step frequency, exercise trajectory, etc. These data can help athletes and coaches have a more comprehensive understanding of their sports status and performance, enabling more precise training and adjustments. For example, in football training, athletes can monitor their heart rate and movement track by wearing smart bracelets or chest bands, so as to adjust their intensity and rhythm in training. Through data collection and application, athletes and coaches can more scientifically manage training and competitive processes, and improve training effectiveness [1]. The application of data analysis and information technology has not only changed the training methods of athletes, but also put forward new requirements for the role of coaches. Traditionally,

*Zhengzhou University of Technology, Zhengzhou, 450044, China (mylifeyuan@163.com).

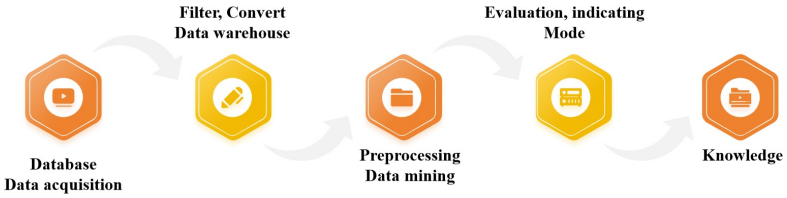

Fig. 2.1: Data mining process.

coaches relied mainly on their own experience and intuition to develop training plans, but now they need to have certain abilities in data analysis and information technology application. Coaches need to learn to collect, process, and analyze sports data, obtain valuable information from it, and apply it to training programs. This poses new requirements for the comprehensive quality of coaches, who need to constantly learn and update their knowledge to adapt to the needs of technological development and data analysis.

In summary, the application of data analysis and information technology in modern sports competitions has become a trend. Through data analysis, coaches can develop targeted training programs to help athletes improve their competitive skills. Athletes can also analyze their sports data to identify their problems and engage in targeted training. Data collection and application can help athletes and coaches have a more comprehensive understanding of sports status and performance, and improve the scientificity and effectiveness of training. The use of data analysis and information technology not only changes the teaching process, but also places new demands on the effectiveness of teachers. With the continuous development of techniques and technology, the use of statistical data and information technology in sports competitions is becoming more and more popular, challenging many methods and approaches to athletes and coaches.

The author aims to explore the application of information technology in athlete data analysis and automatic generation of training plans, and evaluate its impact on athlete training and competitive performance. By collecting and analyzing physiological, technical, and competitive data of athletes, combined with advanced data mining and machine learning algorithms, we hope to automatically generate personalized training plans to help athletes better tap into their potential, improve training effectiveness, and achieve better results in competitions. Through this study, we will be able to gain a deeper understanding of the current status and potential of the application of information technology in athlete training and competition, providing scientific training guidance and decision support for athletes and coaches [15].

2. Decision support system for sports training mode.

2.1. Overview of Association Rule Mining Algorithms. The most crucial aspect of data mining technology is the association rule algorithm, and the Apriori algorithm is a classic algorithm in association rule algorithms. At present, there are various ways to classify association rules, and the most common one is to classify them according to the dimensions of the data types in the association rules [7]. It can be classified into one-dimension and multiple-attribute.There are a lot of influential factors in practice for athletes' physical training, and the data types obtained are much bigger than that of 3D ones. Therefore, the multi-dimension association rules must be taken into account in the design of the DSS of sports training model. Multi-dimensional association rules are more complicated than single-dimension association rules. Usually, multidimensional association rules include data preprocessing, data mining, and model assessment.The concrete data mining process is illustrated in Figure 2.1.

The preprocessing stage in the data mining process mainly involves collecting, processing, and transforming data, which takes the longest time throughout the entire data mining process; The data mining stage mainly analyzes the data in the preprocessing stage through selected association rules, neural network techniques, etc [12]. The evaluation stage of the pattern mainly involves presenting the information obtained from data mining to users, or providing a visual program for real-time viewing and analysis.

2.2. Data Fusion Processing of Sports Training Evaluation Decision Support System. In the evaluation of sports training model based on large data mining, the related data should be integrated into the

DSS.This paper focuses on the classification of the extracted data by neural networks.

2.2.1. Fusion and Clustering of Sports Evaluation Decision Information. The data feature identification function of the sports evaluation decision support system is:

$$P_c = \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha(i,j) P(i,j) \tag{2.1}$$

Assuming that the starting symbol for each of the above attributes is: $C_0 = C_{N/2} = 0, C_{N-n} = C_n^*, n = 0, 1, 2 \ldots, N/2 - 1$, the model relationship between sports training evaluation decision data and cluster center distribution is:

$$P_r = \frac{P_t}{(4\pi)^2 \left(\frac{d}{\lambda}\right) r} \left[1 + \alpha^2 + 2\varepsilon \cos\left(\frac{4\pi h^2}{d\lambda}\right)\right] \tag{2.2}$$

Based on the association criteria of sports training modes, feature recognition of sports decision support systems is carried out based on the different types of data obtained. Among them, the attribute categories of association criteria in the system are:

$$R_\beta X = U\{E \in R \mid c(E,X) \leq \beta\} \tag{2.3}$$

$$R_\beta X = U\{E \in R \mid c(E,X) \leq 1 - \beta\} \tag{2.4}$$

For different data block types m_i and m_j combined with the association criteria of sports training modes, the iterative process of sports training decisions obtained by using the fuzzy mean method in the data types is as follows:

$$S_b = \sum_{i=1}^{e} p(\omega_i)(u_i - u)(u_i - u)^T \tag{2.5}$$

$$S_\omega = \sum_{i=1}^{e} p(\omega_i) E\left[\frac{(u_i - u)(u_i - u)^T}{\omega_i}\right] \tag{2.6}$$

$$S_i = S_b + S_\omega \tag{2.7}$$

In the formula, $p(\omega_i)$ is the set of association rule vectors for the sports training decision system. Based on the above calculation formula, the fusion of information in the sports training mode decision support system is achieved [14].

2.2.2. Improve Apriori algorithm output. It is necessary to create a policy organization that presents the information only during the request for information that will be used to make design decisions for sports training models. The author created a collection of sports training standards organization standards using a modified weight system, and the last important rule is the only product that accepts weight products.

$$\omega_{sij}(n_0 + 1) = \omega_{sij}(n_0) - \eta_{sij} \frac{\partial J}{\partial \omega_{sij}} \tag{2.8}$$

Through the similarity analysis of relevant data in the DSS, the adaptive learning process can be obtained as follows:

$$\alpha_{desira}^i = \alpha_1 \cdot \frac{Density_i}{\sum_i Density_i} + \alpha_2 \frac{AP_i}{AP_{init}} \tag{2.9}$$

Rearrange the relevant data stored in the system in quintuples to obtain the probability density function used for data mining:

$$P_s = P_{2D}^k (1 - P_{2D})^{N-1-k} \sum_{i=1}^{\infty} \lambda_s^i = \frac{\lambda_s}{1 - \lambda_s} \tag{2.10}$$

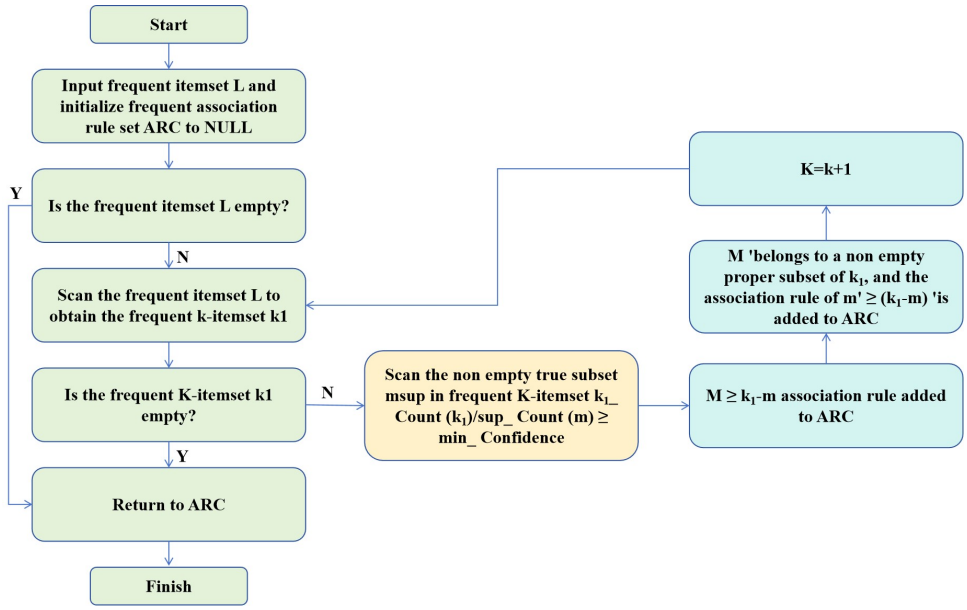

Fig. 2.2: Improved Apriori algorithm for frequent association rule process.

In the formula, λ_s and P_{2D} represent the correlation dimension of the data and the probability that the data can be effectively detected. The motion training model determines that the data clusters in the support system differ in the following ways:

$$\mathrm{DisSim}(A,B) = 1 - \left| \frac{\mathrm{SameDis(A)\text{-}\,SameDis(B)}}{\mathrm{Dis}(A) + \mathrm{Dis}(B)} \right| \tag{2.11}$$

In summary, the author can establish a system database model by mining the data of the sports training mode decision support system and extracting its association rules, and design the corresponding system in combination with software development [5].

2.3. Software Design. The decision support system for sports training mode is based on modern computers and uses computer and programming languages to simulate the sports training effects of athletes through human-computer interaction. It mainly addresses decision-making issues for managers during the implementation of plans. Decision systems can provide athletes with the convenience of obtaining real-time sports data, while also assisting in guiding and tracking the effectiveness of sports.

The author chooses the improved Apriori algorithm with frequent association rules as shown in Figure 2.2 to design the human-computer interaction system. Its main functions include:

1. The human-computer interaction system can provide a more convenient computer environment for decision-makers in sports training. It can check the physical fitness indicators of each athlete based on the front-end display settings, and use computers for tracking and processing.
2. Visually display the operational status of the decision support system, allowing users to fully understand the data changes during the system's operation and make timely adjustments.
3. Based on the output results of the system, targeted adjustments can be made to the training plan, and the simulation can be calculated to form the optimal training plan.
4. The human-computer interaction system can also correct erroneous information and perform preliminary verification and judgment on input data [6].

On the basis of analyzing the decision support system algorithm, the author integrated data mining technology with the system program through algorithm compilation to design a sports training mode decision support system [2]. The main functional modules of the system include communication module, program module,

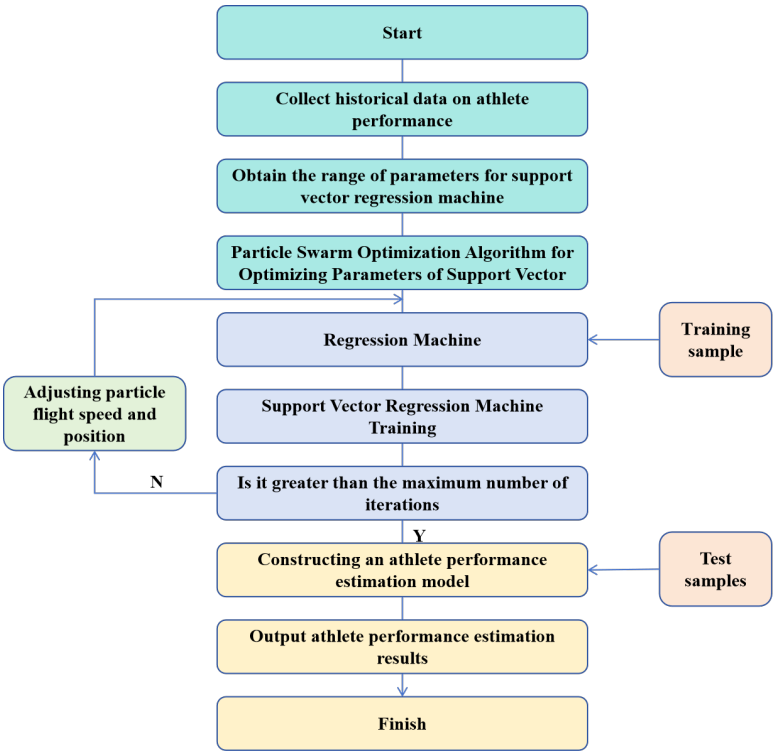

Fig. 2.3: Flow chart of athlete performance modeling and estimation.

database module, and data output module. Using Conti-ki bus technology to transmit and coordinate the data types of the sports training mode decision support system, meanwhile, combining VIX integrated control technology can achieve integrated control of the system. The data perception of decision support systems is built based on the 6LoWPAN protocol stack. The design of the wireless sensor network system adopts Atmel1284P as the main chip to control the overall IoT address allocation and mobilization of the sports training mode decision-making system. After the taskbar address is determined, the system's human-computer interaction is achieved through the TaskBasic interface program.

2.4. Athlete Performance Modeling and Estimation. The workflow of athlete performance modeling and estimation based on big data analysis technology is shown in Figure 2.3.

The steps for modeling and estimating athlete performance using big data analysis technology are as follows:

1. Collect historical data of athlete performance, process the historical data of athletes to obtain the range of athlete performance:

$$x_i'' = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{2.12}$$

   In the formula, the maximum value of the athlete's score is $x_{\max}$, and the minimum value of the athlete's score is $x_{\min}$.
2. On this basis, a support vector machine model based on PSO is proposed [10].
3. Train athletes according to each set of parameters and implement learning through support vector regression machine.
4. If the number of iterations exceeds the set maximum value, the algorithm ends; If the number of iterations is less than the set maximum value, adjust the flight speed and position of the particle swarm.

Table 3.1: Data related to the badminton team training in a certain place

Type	Tid	Item	Avg
Data	5401	114	38

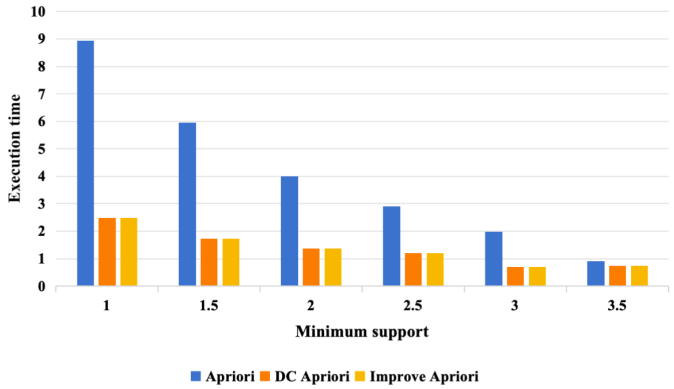

Fig. 3.1: Comparison of execution times at minimum support.

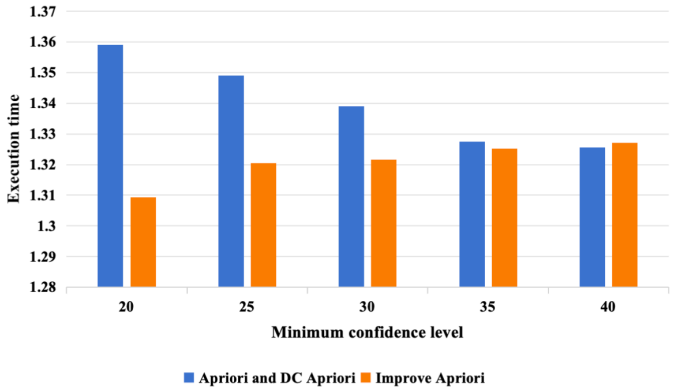

Fig. 3.2: Comparison of execution times at minimum confidence.

5. Increase the number of iterations of particle swarm optimization algorithm.
6. Retrain athlete performance through the optimal parameters of the support vector regression machine, and obtain the optimal parameters of the support vector regression machine through the optimal solutions pbest and gbest. Construct an athlete performance estimation model based on support vector regression machine.
7. Test and analyze the performance of the athlete performance estimation model through athlete performance test samples, and output the final athlete performance estimation results [3].

3. Simulation experiment analysis. On this basis, Apriori algorithm, DC Apriori algorithm and modified Apriori algorithm are compared, the results show that the algorithm is feasible. Most of the time during the experiment, programming was done in the Java language. In this paper, a badminton team in the relevant training data as the research object.

The “Tid”, “item”, and “quantities” in Table 3.1 represent the specific types of training items, the total

number of data items, and the average level per training.

In Figures 3.1 and 3.2, the change in system runtime is shown as minimum support and minimum confidence increase.

Figure 3.1 shows the reaction speed of the modified Apriori algorithm proposed in this article under the minimum support, which shows that it is a more efficient method [9, 13, 8]. As you can see in Figure 3.2, the improved Apriori algorithm has better performance when the minimum confidence is low. As the minimum confidence level increased, the advantages of Apriori faded and the same effect was achieved.

4. Conclusion. In order to improve the competitive level of athletes, we must constantly improve the competitive level. Based on information technology, this paper studies the problem of decision support in sports teaching mode. On this basis, a method of physical education based on network is proposed. Secondly, using Apriori, DC Apriori, Apriori and other classical algorithms to test and verify the Apriori algorithm, the Apriori algorithm can better support training decisions, with high practical value.

REFERENCES

[1] N. E. Barczak-Scarboro, E. Kroshus, B. Pexa, J. K. R. Mihalik, and J. DeFreese, *Athlete resilience trajectories across competitive training: The influence of physical and psychological stress*, Journal of Clinical Sport Psychology, 1 (2022), pp. 1–19.

[2] H. Carrera, J. Romero, I. Berganza, J. Dunavant, G. A. Abascal-Ponciano, J. D. Starkey, C. Starkey, and S. Cho, *192 evaluation of textural characteristics (texture profile analysis and shear force) on commercial training pet treats*, Journal of Animal Science, 101 (2023), pp. 98–99.

[3] Y. Chen and K. Ye, *A wireless network based technical and tactical analysis of volleyball game based on data mining techniques*, Wireless Networks, 29 (2023), pp. 161–172.

[4] J. G. Conn, J. W. Carter, J. J. Conn, V. Subramanian, A. Baxter, O. Engkvist, A. Llinas, E. L. Ratkova, S. D. Pickett, J. L. McDonagh, et al., *Blinded predictions and post hoc analysis of the second solubility challenge data: Exploring training data and feature set selection for machine and deep learning models*, Journal of Chemical Information and Modeling, 63 (2023), pp. 1099–1113.

[5] ———, *Blinded predictions and post hoc analysis of the second solubility challenge data: Exploring training data and feature set selection for machine and deep learning models*, Journal of Chemical Information and Modeling, 63 (2023), pp. 1099–1113.

[6] F. S. De Jesus and L. M. L. Fajardo, *The benefits of training and development programs for lending organization personnel: Basis for development of training program*, IRA-International Journal of Management & Social Sciences (ISSN 2455-2267), 18 (2022), pp. 13–32.

[7] B. Dessirier, K. M. Sharma, J. Pedersen, C.-F. Tsang, and A. Niemi, *Channel network modeling of flow and transport in fractured rock at the äspö hrl: Data-worth analysis for model development, calibration and prediction*, Water resources research, (2023), p. e2022WR033816.

[8] Y. Ge et al., *Intelligent analysis and evaluation method of athletics running data based on big data statistical model*, Mathematical Problems in Engineering, 2022 (2022).

[9] D. P. Golden and J. N. Hertel, *Clinician impact on athlete recovery and readiness in a 24-hour training cycle*, International Journal of Athletic Therapy and Training, 1 (2022), pp. 1–6.

[10] A. R. M. T. Islam, S. C. Pal, R. Chakrabortty, A. M. Idris, R. Salam, M. S. Islam, A. Zahid, S. Shahid, and Z. B. Ismail, *A coupled novel framework for assessing vulnerability of water resources using hydrochemical analysis and data-driven models*, Journal of Cleaner Production, 336 (2022), p. 130407.

[11] T. W. Jones, H. P. Lindblom, M. S. Laaksonen, and K. McGawley, *Using multivariate data analysis to project performance in biathletes and cross-country skiers*, International Journal of Sports Physiology and Performance, 1 (2023), pp. 1–12.

[12] L. P. Koziris, *Ncaa student-athlete training during covid-19 stay-at-home restrictions*, Strength & Conditioning Journal, 44 (2022), pp. 128–130.

[13] Z. Li, Y. Zhou, C. Zhao, Y. Guo, S. Lyu, J. Chen, W. Wen, and Y. Huang, *Design of a cargo-carrying analysis system for mountain orchard transporters based on rgb-d data*, Applied Sciences, 13 (2023), p. 6059.

[14] Z. Mei, *3d image analysis of sports technical features and sports training methods based on artificial intelligence*, Journal of Testing and Evaluation, 51 (2023), pp. 189–200.

[15] L. Yin, W. Zheng, H. Shi, and D. Ding, *Ecosystem services assessment and sensitivity analysis based on ann model and spatial data: A case study in miaodao archipelago*, Ecological Indicators, 135 (2022), p. 108511.

Edited by: Hailong Li
Special issue on: Deep Learning in Healthcare
Received: Jan 4, 2024
Accepted: Jan 25, 2024