CHARACTERISTICS OF PHYSIOLOGICAL CHANGES IN ATHLETE TRAINING BASED ON THE DATA MINING ALGORITHM

ORIGINAL ARTICLE
ARTIGO ORIGINAL
ARTÍCULO ORIGINAL

CARACTERÍSTICAS DAS ALTERAÇÕES FISIOLÓGICAS NO TREINAMENTO DE ATLETAS COM BASE EM ALGORITMO DE MINERAÇÃO DE DADOS

CARACTERÍSTICAS DE LAS ALTERACIONES FISIOLÓGICAS EN EL ENTRENAMIENTO DE LOS ATLETAS BASADAS EN EL ALGORITMO DE MINERÍA DE DATOS

Xiaofeng Tan[1]
(Public health)
Mingfang Song[2]
(Public sports)

1. Physical Education Institute of Hunan University of Technology, Zhuzhou, Hunan, China.
2. Sports Department, Hunan Railway Professional Technology College, Zhuzhou, Hunan, China.

Correspondence:
Mingfang Song
Sports Department, Hunan Railway Professional Technology College, Zhuzhou, Hunan, China.
30791647@163.com

ABSTRACT

Objective: In the competition of athletic training, it is imperative to use various physiological and biochemical indicators to study the changes they can bear. Methods: In this paper, national tennis players' physiological and biochemical indicators are taken as samples, and Artificial Neural Network (ANN) in data mining algorithm is used to classify and predict the sample data. Based on this, to solve the BP neural network's failure in easily falling into a local minimum, the ant colony optimization (ACO) algorithm was introduced to train the changes in the neural network. Finally, the improved BP neural network technology of the ant colony optimization algorithm is used in the model to analyze the physiological changes in tennis players. Results: The research results show that the model successfully predicted the physiological change in athletes and could provide coaches with a basis for decision-making. Conclusions: The physiological change in athletes is combined with the neural network algorithm to establish a connection between the two, which provides an effective and reliable method for detecting the physical function of sports transportation with unique guidance in athletes' training and competition. ***Level of evidence II; Therapeutic studies - investigation of treatment results***.

Keywords: Data Mining; Physiological Monitoring; Athletic Performance.

RESUMO

Objetivo: Na competição do treinamento atlético, é imperativo usar vários indicadores fisiológicos e bioquímicos para estudar as alterações que eles podem suportar. Métodos: Neste trabalho, os indicadores fisiológicos e bioquímicos dos tenistas nacionais são tomados como amostras, e a Rede Neural Artificial (ANN) no algoritmo de mineração de dados é usada para classificar e prever os dados da amostra. Com base nisso, para solucionar a falha que a rede neural da BP tem em cair facilmente num mínimo local, o algoritmo de otimização da colônia de formigas (ACO) foi introduzido para treinar as alterações na rede neural. Finalmente, a tecnologia melhorada da rede neural BP do algoritmo de otimização da colônia de formigas é usada no modelo de análise das alterações fisiológicas nos tenistas. Resultados: Os resultados da pesquisa mostram que o modelo previu com sucesso a alteração fisiológica dos atletas e pôde fornecer aos treinadores uma base para a tomada de decisões. Conclusões: A alteração fisiológica dos atletas é combinada com o algoritmo da rede neural para estabelecer uma conexão entre os dois, o que fornece um método eficaz e confiável para a detecção da função física do transporte esportivo com orientação singular no treinamento e competição dos atletas. ***Nível de evidência II; Estudos terapêuticos – investigação de resultados de tratamento***.

Descritores: *Mineração de Dados; Monitoramento Fisiológico; Desempenho Atlético.*

RESUMEN

Objetivo: En la competición del entrenamiento atlético, es imperativo utilizar varios indicadores fisiológicos y bioquímicos para estudiar los cambios soportables. Métodos: En este trabajo se toman como muestra los indicadores fisiológicos y bioquímicos de los tenistas nacionales, y se utiliza la Red Neural Artificial (ANN) en el algoritmo de minería de datos para clasificar y predecir los datos de la muestra. En base a esto, para resolver el fallo que tiene la red neuronal BP en caer fácilmente en un mínimo local, se introdujo el algoritmo de optimización de colonias de hormigas (ACO) para entrenar los cambios en la red neuronal. Por último, la tecnología de red neuronal BP mejorada del algoritmo de optimización de colonias de hormigas se utiliza en el modelo para analizar los cambios fisiológicos de los tenistas. Resultados: Los resultados de la investigación muestran que el modelo predijo con éxito el cambio fisiológico en los atletas y podría proporcionar a los entrenadores una base para la toma de decisiones. Conclusiones: El cambio fisiológico de los atletas se combina con el algoritmo de la red neuronal para establecer una conexión entre ambos, lo que proporciona un método eficaz y fiable para detectar la función física del transporte deportivo con una orientación única en el entrenamiento y la competición de los atletas. ***Nivel de evidencia II; Estudios terapéuticos – investigación de resultados de tratamiento***.

Descriptores: *Minería de Datos; Monitoreo Fisiológico; Rendimiento Atlético.*

DOI: http://dx.doi.org/10.1590/1517-8692202228052021_0533
Article received on 12/11/2021 accepted on 12/22/2021

INTRODUCTION

In today's sports field, the use of science and technology is more and more extensive. The use of science and technology to assist athletes in training practice and promote the scientific dataization of sports training management is the mainstream in today's world.[1] However, at present, when developing training programs for athletes, coaches generally rely on empirical knowledge. They do not make full use of athletes' physiological index data, lack quantitative analysis of data, and cannot control athletes' physiological load in time.[2] Therefore, in this paper, the physiological load characteristics of athletes are studied, and the data mining algorithm is used to establish an ANN based ant colony optimization algorithm for load forecasting.[3] This research can effectively improve the scientific nature of athlete training planning, better play the athlete's physical function, and provide technical guidance for the training of sports competitions.[4]

In the research method, the physiological and biochemical indexes of the physiological state of the athletes are comprehensively analyzed, and the basic principles of the ANN algorithm in the data mining algorithm are introduced. Aiming at the problems of slow convergence, unstable network learning and easy to fall into local optimum, an artificial ant colony optimization algorithm is proposed to enhance its global locating ability and accelerate convergence.[5] Finally, the athlete's physiological and biochemical indicators, ANN algorithm and ant colony algorithm are used to form the prediction model.

The research in this paper has certain innovative significance. At present, there are many researches on athletes' load status and physiological and biochemical indicators. However, most of them do not make full use of this physiological information, and the rules existing in the data are mined to form a prediction system that can guide athletes' training. In this study, artificial intelligence technology was applied to the field of sports training. By using data mining algorithms, hidden layer information and rules in a large amount of data are fully extracted, which is suitable for athletes' training management.[6]

This research is mainly divided into three parts: In the first part, related research on data mining technology in the field of sports is elaborated. In the second part, ANN algorithm and ant colony optimization algorithm are introduced. Based on this, an improved data mining algorithm is formed for the determination of athlete load. In the third part, the established athlete load forecasting model is tested and the experimental results are obtained.

Model construction

Neural network and ant colony algorithm introduction

ANN is a new type of artificial intelligence technology that connects one or more neurons to each other to simulate the working mode of the human brain. ANN has self-learning capabilities and is suitable for prediction. It is capable of distributed storage and associative memory of information and is highly robust. Its processing of information is hierarchical, massively parallel, and runs faster. The working principle of the ANN is to process the information through different connection weights of the neurons of each node, and then transfer to the next node. The working process is mainly divided into two parts: the learning phase and the work phase. In the learning phase, the deviation between the expected value of the input learning sample and the actual value is calculated by the ANN and the weight is adjusted. And this step is repeated repeatedly to complete the learning. In the working phase, a stable connection weight is obtained, and the output value can be directly obtained according to the input content.

In Figure 1, assume that the input layer has 1 unit, the hidden layer has J units, and the output layer has K units. The weight of the input layer to the hidden layer is vij, and the weight of the hidden layer to the output layer is wjk. Assuming the input vector is $X = (x_0, x_1, x_2, \cdots, x_{1\text{-}1})$, then, its expected output vector is $D^P = (d_0, d_1, d_2, \cdots, d_{K\text{-}1})$. Therefore, in the forward calculation output, the total error of the N samples is as in equation (1), among which, d_k^p is the expected output of sample P and O_k^p is the actual output of the output layer.

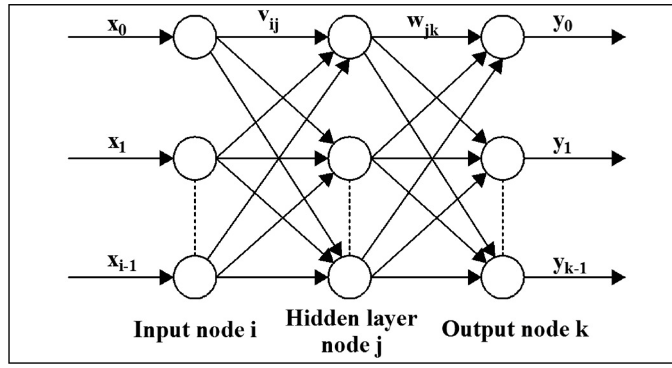

Figure 1. Schematic diagram of three-layer back propagation neural network.

$$E = \frac{1}{2N}\sum_{p=0}^{N-1}\sum_{k=0}^{K-1}\left(d_k^p - O_k^p\right)^2 \quad (1)$$

In the inverse error propagation, the optimal gradient descent method is used to solve the unconstrained optimization equation. Therefore, the back-propagation errors of the output layer and the hidden layer are δ_k and δ_j, respectively, among which, $f(net_k) = O_k$ is the output of the kth node of the output layer, and $f(net_j) = O_j$ is the output of the jth node of the hidden layer.

$$\delta_k = (d_k - O_k) f'(net_k) \quad (2)$$

$$\delta_j = f(net_j)\sum_{k=0}^{K-1}\delta_k w_{jk} \quad (3)$$

When $f(net) = \frac{1}{1+e^{-net}}$ is used for each layer, the weight corrections Δw_{jk} and Δv_{ij} of the output layer and the hidden layer can be calculated:

$$\Delta w_{jk} = \eta O_j (d_k - O_k) O_k (1 - O_k) \quad (4)$$

$$\Delta v_{ij} = \eta O_j (1 - O_j)\sum_{k=0}^{K-1}\delta_k w_{jk} \cdot O_i \quad (5)$$

The working process of the BP neural network is shown in Figure 2. The weights and thresholds are initialized, and P learning samples are sequentially input to calculate the output of each node of the hidden layer and the output layer. The back propagation error of each layer is calculated according to equation (2) (3). Until the P samples are calculated, the weights and thresholds of each layer are weighted. The operation ends when each sample and its output neurons satisfy the mean square error less than the expected value.

Neural network (ACO-BPNN) training based on ant colony algorithm optimization

For BP neural network, it is not suitable for multi-peak function minimum point search, which may fall into local optimal solution. The common way is to add a momentum term to the standard BP neural network to change the learning rate, but only the problem of too long

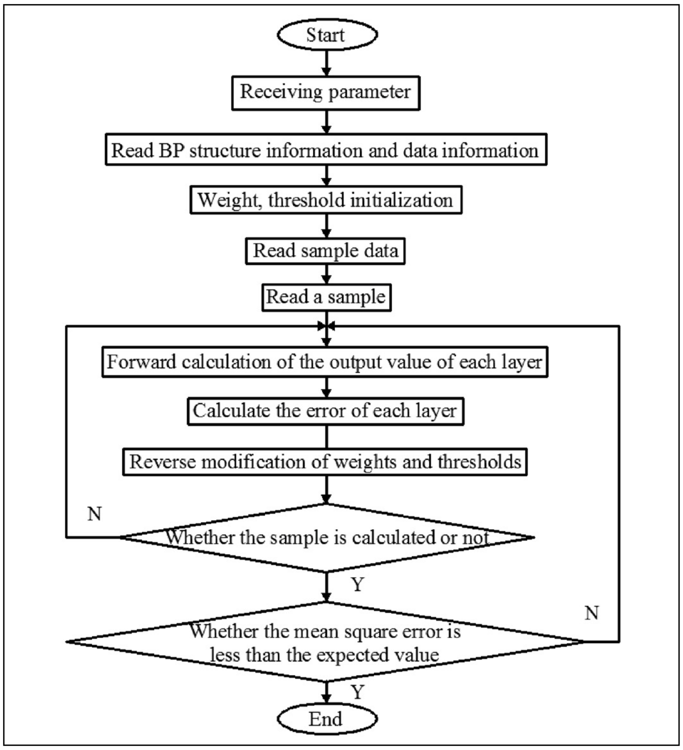

Figure 2. Algorithm flow chart of back propagation neural network.

convergence time is solved, and there is still no possibility of getting rid of the local optimal solution. Therefore, in the above, it is proposed to join the ant colony algorithm to train the BP neural network, and construct the ant colony neural network model (ACO-BPNN) as the algorithm for the study of athletes' physiological load characteristics.

In order to avoid the defect of network performance degradation caused by too many hidden layers, the neural network structure of n-r-1 form is adopted. This BP neural network has a hidden layer, as shown in Figure 3. In this paper, the number of input nodes of the neural network is determined by the professional training period of the athletes and physiological and biochemical indicators. The output node is set to one and the value range [-1, 1] of the output node is determined according to the measured value of the load amount. The closer the value of the output node is to 1, the less the load of the athlete is. The number of hidden layer nodes is m. According to the empirical formula $m = \sqrt{n+l} + \alpha$, α is a constant, the value range is [1, 10], n is the number of neurons in the input layer, and l is the number of neurons in the output layer. The training flow chart of the whole ant colony algorithm for BP neural network is shown in Figure 4.

Firstly, the weight interval in the ant colony optimization algorithm is defined as $[W_{mim}, W_{max}]$, and the value interval is evenly divided into r sub-regions, according to which pheromone tables of different weight points are established, as shown in Table 1. Assume that at the initial time, the time and the number of cycles NC is 0, so the value of the pheromone quantity τ_0 of each point is the same, ρ is the pheromone volatilization coefficient, and the end condition is that the maximum number of cycles NC_{max} is reached.

Experimental Design and Analysis

Experimental indicator selection and sample data processing

When training athletes, it is necessary to keep up with the changes in the physical function of the athletes to determine the intensity of the exercise load they are subjected to during training. However, with reference to a large number of literatures, it is found that the factors affecting the training load are not only the athletes' biochemical indicators, but also many different levels and aspects affect the athlete's training load intensity. If only a single biochemical indicator is used for evaluation, a certain error will occur. Therefore, the blood index, urine index and professional training period are used to evaluate the training load of the athletes. The blood index can reflect the effect of exercise on body function over a period of time. In this study, creatine kinase (CK enzyme), blood urea nitrogen (BUN), hemoglobin (HGB), blood urea (BU), serum testosterone/cortisol (T/C) and serum lactate dehydrogenase (LDH) are selected. CK enzyme can indicate muscle cell damage and reflect the athlete's fatigue status. The BUN level can comprehensively reflect the athlete's load bearing capacity and help the coach to regulate the amount of exercise and training rhythm. HGB is related to the endurance of the human body.

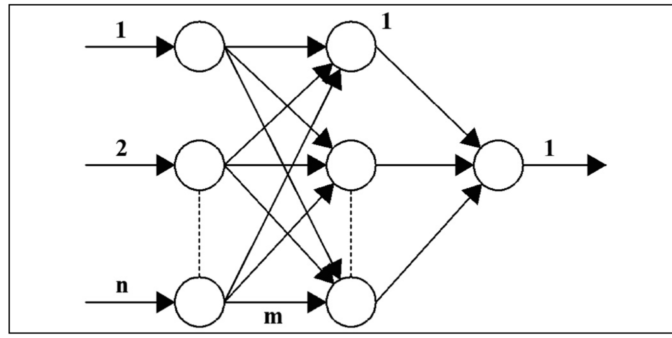

Figure 3. Back propagation neural network structure.

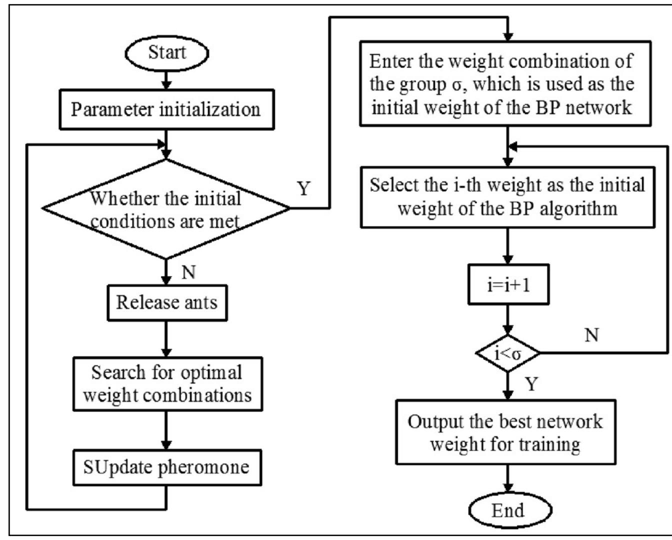

Figure 4. ACO-BPNN training process.

Table 1. Pheromone table with different parameters.

Grade	1	2	3	4	……	r+1
Dividing scale	a1	a2	a3	a4	……	ar+1
Pheromone value	τ(1)	τ(2)	τ(3)	τ(4)	……	τ(r+1)

Experimental results

In this study, the physiological indexes of 187 athletes are selected as sample data, and 170 samples are extracted into the ant colony optimization BP network model for training, and another 17 samples are used as test samples. A number of experts and doctors jointly determine the criteria for the output, and finally determine that the output is 1, indicating that the athlete can fully adapt to the current exercise load, and can appropriately increase the amount and intensity of training. The output 0 means that the athlete just adapts to the current exercise load and can maintain the current training amount. The output -1 indicates that the athlete's body cannot withstand the current exercise load, and the training amount should be appropriately reduced to adjust the training plan.

In this study, the specific parameter settings for the ant-optimized BP network algorithm are shown in Table 2, and the scale M is set to 20. The maximum number of iterations QQQ is 200. The value interval of the weight w is [-2, 2]. The domain is divided equally into 60 sub-regions, namely, r=60. Only one optimal solution is retained. To ensure that the weight correction gradient falls, the minimum value of QQQQQ is 0.003.

The standard BP neural network algorithm, the commonly used BP neural network algorithm with the momentum item and the BP neural network algorithm trained in the ant colony optimization algorithm are compared. The results of these three algorithms after computer simulation test are shown in Table 3. According to the data in Table 3, the BP neural network model trained by the ant colony optimization algorithm has better performance and stronger stability in the convergence process. The experimental data in the comparison table can be seen. Compared with the other two algorithms, the ant colony BP network algorithm in this study converges faster and the mean square error is smaller.

CONCLLUSION

Scientific technical means have been effectively applied in many aspects of the sports field. When coaches train and guide athletes, it is able for them to develop a more scientific and safe and standardized training plan according to the comprehensive indicators of athletes, which can help athletes adjust the intensity and level of training, and effectively promote the improvement of athletes' performance. In this paper, aiming at the relationship between the exercise load that athletes can bear and their physiological indexes, a predictive model based on data mining algorithm is constructed. The ant colony optimization algorithm is trained on the BP neural network to solve the defects of the standard BP neural network which are easy to fall into the local minimum and the convergence speed is slow. The BP neural network model based on ant colony optimization algorithm is compared with the standard BP neural network and the BP neural network with momentum item. The experimental results show that the ant colony neural network has faster convergence speed and higher computational efficiency, as well as better search performance. The problem that BP neural network can only be localized is solved, and the calculation result is stable. The model can be applied to the athlete's load forecasting, effectively improve the quality of the coach's training plan and promote the athlete's exercise level. However, there are still some shortcomings in this research. It is necessary to quantify the fuzzy concept, and combine qualitative and quantitative analysis into the evaluation system to enhance the accuracy of the prediction model. The ant colony algorithm is still not mature, and the parameter selection in this study needs to be further strengthened. It needs to be improved by a large number of relevant researches by more scholars. In order to adapt to the computational needs of large-scale complex networks, the convergence of ant colony algorithms also needs to be improved.

Table 2. Parameter setting of ACO-BPNN and Standard BP neural network.

algorithm	M	NACO	Wmax	Wmin	r	Q	σ	η	Eo	NRP
ACO-BPNN	20	200	2	-2	60	0.005	1	0.003	0.005	12000
Standard BP neural network			0.1	-0.1				0.003	0.005	20000

All authors declare no potential conflict of interest related to this article

AUTHORS' CONTRIBUTIONS: Each author made significant individual contributions to this manuscript. Xiaofeng Tan, Mingfang Song: writing and execution.

REFERENCES

1. Ma J, Liu G. Application of Data Mining Technology in Analysis of Volleyball Biochemical Indexes. Zhejiang Sports Science. 2016;38(5):96-9.
2. Gao S. Athlete performance prediction model based on chaos theory and machine learning algorithm. Modern Electronic Technology. 2018;41(510):152-5.
3. Gao L. Construction and Simulation of Athlete's Error Action Recognition Model Based on Computer Vision. Microcomputer Applications. 2018;34(302):62-5.
4. Li Z, Mao L, Sheng B et al. Research on Physical Fitness Video Tracking Technology of Football Players Based on KCF Algorithm. Journal of Shanghai Physical Education Institute. 2018;42(209):79-84.
5. Yang H, Luo T, Li L et al. Monitoring method of athlete's respiratory state based on inertial measurement unit. Chinese Journal of Inertial Technology. 2018;26(5).
6. Ma D, Li S, Zhang J. Design of Athletes' Special Performance Prediction System Based on Hybrid Genetic Neural Network. Modern Electronic Technology. 2018;41(8):183-6.