Training Load Prediction in Physical Education Teaching Based on BP Neural Network Model

Danqing Liu, Shoubang Li, and Kun You
Xi’an Shiyou University, Xi’an, Shaanxi 710065, China

Correspondence should be addressed to Danqing Liu; liudanqing@xsyu.edu.cn

Received 31 December 2021; Revised 29 January 2022; Accepted 8 February 2022; Published 4 April 2022

1. Introduction

Monitoring training load is an important part of modern sports science practice. Load data is typically collected, interpreted, cleaned, analyzed, and disseminated with the goal of improving player performance and reducing injury risk. Practitioners seek to optimize training load at many points during the training process, such as planning and altering individual sessions, day-to-day, season periodization, and managing athletes with a long-term perspective to increase player performance and reduce injury risk [1]. Training load prediction can aid in determining whether an athlete is adapting to a training program and in minimizing the risk of developing nonfunctional overreaching, illness, and injury [2].

Training load consists of two parts, including load volume and load intensity. Load intensity is a physical quantity that reflects the value of exertion, functional tension, and force in a particular exercise unit and the concentration of training workload in a certain period. Load is an index reflecting a specific practice unit [3]. At present, there are many indicators to measure the size of exercise load, such as maximum oxygen uptake (VO2max), blood lactic acid, heart rate, and Rating of Perceived Exertion (RPE). [4]. Heart rate (HR) is a more effective indicator for monitoring exercise load. It is verified through experiments that there is a specific correlation between HR and load intensity [5].

Many people believe that monitoring an athlete’s training load is critical for determining if the athlete is adjusting to the training program and reducing the danger of nonfunctional overreaching (fatigue that lasts weeks to months), injury, and illness. Rago et al. [6] investigated the top Spanish football teams’ training load and cardiorespiratory health and reported that cardiopulmonary fitness improves moderately as the season advances. Moreover, the cumulative total training distance between early February and early May was negatively correlated with the average HR change percentage in the last 30 s of Submaximal Yo-Yo Intermittent Recovery Test Level 1 (Yo-Yo IR1SUB).
Orellana et al. [7] explored that no matter how the exercise intensity and duration change (average four milliseconds), the Root Mean Square of the Successive Differences (RMSSD) value drops significantly. Within 30 minutes, the recovery slope of HR variability was used as an indicator of internal training load. In addition, the RMSSD recovery is linear and varies according to exercise intensity. Nakamura et al. [8] analyzed the changes in HR variability of futsal players during preseason training. The authors in [9] studied the changes in baroreflex control HR of male volunteers during aerobic speed measurement exercise. Chen [10] developed a mathematical model that classifies volleyball players’ physical fitness. The 100-meter run that affects the HR during aerobic speed measurement exercise. The authors in [9] studied the changes in baroreflex control HR of male volunteers during preseason training. The authors in [9] studied the changes in baroreflex control HR of male volunteers during preseason training.

With the advancement of information technology, the model prediction has become popular in several sports. It can effectively increase the quality of training [11]. Song et al. [12] discussed the safety prediction and evaluation of sports injuries based on convolutional neural networks (CNN) of deep learning. The author in [13] improved the traditional BP neural network (BPNN) algorithm and used the adaptive BPNN to build a sports prediction model. Wang and Qu [14] analyzed the regression prediction model of sports based on support vector machine (SVM) and AI. Jiang and Chen [15] created a monitoring and analysis system that has four key functions: data entry and performance conversion, vertical and horizontal analysis, result output, and decision-making plans for an efficient athlete’s physical fitness and skills. Although a large number of scholars have explored sports prediction models, studies on the prediction of training load in physical education are limited. Currently, BPNN is one of the most widely used neural network models. It has powerful parallel processing capabilities, adaptive adjustment capabilities, and powerful mapping capabilities [16, 17]. Therefore, the BPNN algorithm explores the training load in physical education. This is of great significance for improving the ability of physical education teachers to control the training load.

In this study, a training model based on the back-propagation neural network is suggested for predicting the training load of middle school students in physical education. The training load is divided into type I, type II, and type III, and the average heart rate values of students in each grade during physical training are measured. Ninety students in the seventh, eighth, and ninth grades are selected, and principal component analysis (PCA) and BPNN are applied to predict the training load of middle school students. The score matrix corresponding to the main component is obtained and used as the construction parameter of the BPNN model for model training. After the training is successful, the average HR values of 5 students in the seventh, eighth, and ninth grades in the physical activity class are used as the test set for model prediction. Finally, the model’s accuracy is evaluated by comparing the actual and predicted values.

The rest of the manuscript is organized as follows: In Section 2, a detailed illustration is provided about data collection and proposed training load prediction method. The results are discussed in Section 3, and the conclusion is given in Section 4.

2. Materials and Methods

2.1. Experimental Objects and Measurement Tools. During data collection, ninety students in the seventh, eighth, and ninth grades (30 for each grade) of a middle school were selected. All volunteers had no visceral diseases and normal liver and kidney functions. Before the experiment, the volunteers were informed of the potential risks of participating in the investigation. Stopwatches and measuring tapes were used to measure the load of various physical training methods (number of times, number of sets, time, and distance). Students obtained immediate HR through Polar team 2HR monitoring equipment during physical training. Polar Team Pro is a powerful and simple-to-use tool for tracking athlete performance both in real-time and afterward. It analyzes everything on the go or after the workout to gain insight into individual and team physical performance and training load safely and securely.

2.2. Experimental Procedure. The seventh-, eighth-, and ninth-grade experiments were carried out in physical education classes on Mondays, Tuesdays, and Thursdays, respectively. Before physical training, volunteers wore HR equipment correctly and accomplished some simple preparatory activities. The volunteers’ immediate HR was obtained through the Polar team 2HR monitoring device and iPad during the physical training process. The HR data before and after each training set was saved. After the experiment, the HR data and the volunteers’ practice conditions were recorded. The experimental process is shown in Figure 1.

For data analysis, Microsoft Excel was used. The average HR of each training item of the volunteer was calculated to obtain the original signal for modeling and testing. According to the training load classification standard provided in [16, 18], the test value was divided into training load classification as an objective basis for modeling.

The average HR of each volunteer in each training item is calculated. Combined with the training load categories in Table 1, the training intensity is divided into low, moderate, and high.

2.3. Design of the Physical Training Program. Physical training mainly includes five aspects: speed, strength, endurance, flexibility, and agility. Flexibility and sensitivity training are improved when other aspects of practice are strengthened [19, 20]. Therefore, the two parts of flexibility and sensitivity are not studied separately for training load. In summary, the three elements of strength, speed, and endurance have been designed for training.

Programs. Nine training items, including push-ups, sit-ups, squats, fast leg lifts, short sprints, fast arm swings, constant speed, durable running, 400-meter repetition runs, and continuous rope skipping, were selected separately.
specific action essentials and training programs for each project are shown in Table 2.

<table>
<thead>
<tr>
<th>Types of training load</th>
<th>Average HR (times/minute)</th>
<th>Training intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>&lt;130</td>
<td>Lower</td>
</tr>
<tr>
<td>II</td>
<td>130–170</td>
<td>Moderate</td>
</tr>
<tr>
<td>III</td>
<td>&gt;170</td>
<td>Higher</td>
</tr>
</tbody>
</table>

2.4. Data Processing and Model Establishment

2.4.1. Data Preprocessing. The normalization process is adopted. Data normalization is a fundamental component of data mining. It converts the source data into a format that allows for efficient data processing. The primary goal of data normalization is to reduce or eliminate duplicate data. Normalization aims to map the data into the same interval after algorithm operation [21]. In this study, we employed the min–max method of normalization. The standardized processing method is shown in

$$x = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$$

where $x_i$ represents the current value of such data, $x_{\text{max}}$ is the maximum value of such data, and $x_{\text{min}}$ shows the minimum value of such data.

2.4.2. Principal Component Analysis. Principal component analysis (PCA) is a statistical analysis method. It has been widely used in data processing. It can transform multivariate statistical problems into a few representative variables and retain critical information. The idea of dimensionality reduction is used to convert multiple variables into several comprehensive variables under the premise of losing little information. This includes most of the information of the original variables and dramatically reduces the amount of computation, which improves the efficiency and speed of data computation [22, 23]. According to the variance contribution rate, the comprehensive indicators are sorted in descending order. The first complete variable has the most significant variance contribution rate and is called the first principal component. Except for the first wide variable, the second comprehensive variable has the most considerable variance contribution rate. It is not related to the first complete variable and is called the second principal component. By analogy, the $n$th comprehensive variable is called the $n$th central component. The selected comprehensive variables should ensure that at least 85% of the critical information is reflected in the complete evaluation results. Such evaluation results are scientific and reliable. The score matrices of several principal components whose cumulative contribution rate of main component variance exceeds 90% are selected as modeling parameters. The variance contribution rate is used to reflect the amount of information of each principal component [24], as shown in

$$F_1 = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i}$$

The cumulative contribution rate is used to reflect the comprehensive ability of the first $n$ principal components [25], as shown in

$$F_2 = \sum_{i=1}^{m} \left( \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i} \right)$$

In (2) and (3), $\lambda_i$ is the variance contribution rate, $p$ is the number of principal components, and $m$ is the current accumulated number.

2.4.3. BP Neural Network. BPNN has a wide range of applications. The learning is divided into two parts: forward propagation and backpropagation. It is a network with unidirectional propagation and multilayer forward characteristics and belongs to a feedforward neural network, and the error backpropagation algorithm is adopted. The algorithm has input layer nodes, output layer nodes, and one or more hidden layer nodes. Backpropagation corresponds to the direction of information transfer of the neural network from the input layer to the hidden layer to the output layer. When the gradient descent method solves the neural network parameters, the error transmission direction is from the output layer to the hidden layer to the input layer. During the backpropagation, the error signal is propagated from the output layer to the input layer in a layer-by-layer manner. During the backpropagation of the error signal, the
weight value of a network is regulated by the error feedback. The weight and offset values are continuously modified to get the network’s real output closer to the intended one. Backpropagation is a fast, simple, and easy learning algorithm. It is a flexible method as it does not require prior knowledge about the network. It is a standard method that generally works well. When the neural unit of each layer receives the input information of each unit of the previous layer, the weight and mode processing are used to receive it. Each neuron contains an activation function, which is used to calculate the output value of the unit [26].

The BPNN algorithm uses gradient search technology to minimize the mean square error (MSE) between the actual output value of the network and the expected output value. The specific steps of BPNN learning are given in Table 3.

If the precision is achieved, the algorithm ends. Otherwise, a new learning sample and the expected output are selected and transferred to a step to enter the next round of learning [27]. There are various parameters of the BPNN algorithm, which are explained as follows:

(i) The number of network layers: the structure of BPNN is composed of the input layer, hidden layer, and output layer. Among them, the hidden layer can be one layer or multiple layers. The three-layer BPNN can approximate continuous functions in any closed interval. It has nonlinear solid mapping capabilities and self-learning, self-organization, and self-adaptation [28]. Therefore, the basic three-layer structure is selected.

(ii) Number of neurons in the input and output layers: the number of input variables can be obtained; that is, the number of neurons can be obtained in the input layer. The prediction target of the model is the load result of the training volume, which is used to describe the athlete’s load to the training, so the number of neurons in the output layer is 3. The load results are classified, representing too minor, moderate, and excessive training volume.

(iii) Number of neurons in the hidden layer: the calculation of the number of hidden layer nodes is given in

\[ n_1 = \sqrt{n + m + \alpha}, \]  

where \( n_1 \) is the number of neurons in the hidden layer, \( n \) is the number of neurons in the input layer, \( m \) shows the number of neurons in the output layer, and \( \alpha \) is an integer between 1 and 10.

<table>
<thead>
<tr>
<th>Physical training content</th>
<th>Action essentials</th>
<th>Training program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength</td>
<td>Push-ups</td>
<td>Straighten your arms shoulder-width apart, keep your legs close together, use the palms of your hands and your toes on the ground, in a prone position, straight up your arms under your bent arms, and your body is in a straight line. Lie on your back, cover your ears with your hands, bend your knees at an angle of about 90 degrees, fix your feet on the ground, and bend your torso with the strength of your waist and abdomen, with your elbows touching your knees as the standard.</td>
</tr>
<tr>
<td>Sit-ups</td>
<td></td>
<td>Stand with your feet shoulder-width apart, with your toes facing forward, squat straight with your knees bent, and keep your upper body upright. When squatting, the upper and lower legs are at an angle of about 70–80 degrees.</td>
</tr>
<tr>
<td>Squats</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed</td>
<td>Quick leg raise</td>
<td>Stand upright and raise your legs quickly and alternately. When submitting, the thigh and calf are at a right angle. The supporting leg is straightened, and the toes are on the ground. The two legs are alternately counted up and down. With correct and coordinated running movements, quickly rush to the finish line and jog back to the starting point.</td>
</tr>
<tr>
<td>Short sprint</td>
<td></td>
<td>30-meter sprint, 45-second interval, six sets</td>
</tr>
<tr>
<td>Fast arm swings</td>
<td></td>
<td>One group of 40 times, 60-second interval, three groups</td>
</tr>
<tr>
<td>Endurance</td>
<td>Constant speed and durable running</td>
<td>Run at a constant speed with the correct running posture according to the rhythm of the music. Repeat running exercises with correct running posture.</td>
</tr>
<tr>
<td>400-meter repetition runs</td>
<td></td>
<td>400-meter interval, 90-second interval, three groups</td>
</tr>
<tr>
<td>Continuous rope skipping</td>
<td></td>
<td>For 180 seconds, measure the HR at 30 seconds, 60 seconds, 90 seconds, 120 seconds, and 180 seconds</td>
</tr>
</tbody>
</table>

Table 2: Design of the physical training program.
The average HR data of volunteers of each grade in each training program was standardized, and PCA was performed. Several principal component score matrices with cumulative paramount component variance contribution rates exceeding 90% were selected as the characteristic parameters of each grade for modeling. The flowchart of modeling is shown in Figure 2.

The experimental data of 25 volunteers were constructed using a BPNN model. The experimental data of 5 volunteers were used as the test set data to predict the modeling results. Finally, the accuracy is obtained of the judgment prediction model.

### 3. Results and Discussion

#### 3.1. The Distribution of the Number of Students in Different Training Load Categories in Each Grade

According to the training program, the HR values of students in each grade were measured during training, and the average HR was obtained. The standardized data was divided according to the average HR value of the students during training and the training load type provided in the literature. The training load intensity was divided into type I, type II, and type III, respectively. The specific distribution of personnel in each grade is shown in Figure 3.

In Figure 3, in terms of the total number of people, type I is the least, followed by type III, and type II is the highest. The training intensity is moderate for most students in each grade. Comparing the distribution of personnel in the three grades, the number of people in the seventh grade is the minimum in the first class, and the number of people in the eighth and ninth grades is the same. The number of people in type II is in the ninth-grade > seventh-grade > eighth-grade order, respectively. Among class III, the ninth grade has the minimum number of people, and the seventh and eighth grades have the same number of people.

#### 3.2. Results of PCA

The variance contribution rate and the cumulative contribution rate of each principal component are calculated according to (3). The score matrix of the selected main elements is used as the modeling parameter. The contribution rates of the central features of each grade are shown in Figures 4–6, respectively. The abscissa represents the variance contribution rate, and the ordinate represents the principal component.

In Figures 4–6, the contribution rate of the principal component variance of the seventh, eighth, and ninth grades shows a decreasing trend. The variance contribution rate of the first principal component is about 60%. The variance contribution rate of the second principal component is about 15%. The variance contribution rate of the third principal component drops to about 10%. The variance contribution rate of the fourth main component is about 5%. Combined with the cumulative variance contribution rate curve, the cumulative contribution rate of the first four principal components in grades 7, 8, and 9 has reached 90%. To simplify the calculation, each grade chooses the score matrix of the first four principal components as the modeling parameters of the BPNN model.

#### 3.3. Predictive Training Load Type of the BPNN Model

In each grade, the training data of 25 students are selected as the training set. The training data of 5 students is used as the test set. The BPNN is trained with the principal component
feature parameters of the training set. The MSE of the trained prediction model drops to 10%–4%. The weight and threshold matrix in the current state is used as the final prediction model’s weight and threshold matrix. The test set data is predicted. The prediction results for each grade are shown in Figures 7–9.

In Figures 7–9, the predicted value and the actual value curve trend are roughly the same. The predicted values are concentrated in the vicinity of class II. This result is consistent with the objective classification result, indicating that the training load type prediction conforms to the actual training load type. In Figure 7, the predicted value and real value of sample no. 10 are quite...
different. This may be due to the small sample size of the experiment and the poor representativeness of the sample. Although there is a slight deviation in the model’s applicability, the other test results are generally satisfactory. The predictive model can predict the training load in physical education.

4. Conclusion

An appropriate training load can effectively improve students’ health in physical education. In this study, a training model is presented for the prediction of the training load of middle school students in physical education-based machine learning techniques such as BPNN. The HR data of the seventh-, eighth-, and ninth-grade students were collected in speed, strength, and endurance. Combined with the classification standard of training load categories, the training load categories of students are divided into each grade. For most students in all grades, the training load type belongs to type II, and the training load is divided into each grade. For most students in all grades, the training load type belongs to type II, and the training intensity is moderate. Secondly, the average HR value was analyzed using PCA for each training item $d$. According to the variance contribution rate and cumulative variance contribution rate, the cumulative contribution rate of the first four principal components in grades seven, eight, and ninth has reached more than 90%. Therefore, the BPNN model uses the score matrix corresponding to the four principal components. After the model was successfully trained, the physical training HR data of 5 people selected in each grade were used to test the prediction model. The predicted value is basically in line with the actual training load type. However, the expected and fundamental values of sample no. 10 in the seventh grade are different, which may be due to the poor representativeness of the samples. Although the prediction model has a slight deviation, the overall prediction results are relatively satisfactory. Therefore, the predictive model effectively predicts the training load in physical education. In future work, to improve the accuracy of the prediction model, experimental samples or improved algorithms will be added for related research.

Data Availability

The data used to support the findings of this study are included within the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Education Department of Shaanxi Provincial Government, Shaanxi Province Higher Education and Teaching Reform Research Project, “The Exploration and Practice of the Teaching Mode of Integration of Inside and Outside Class in Colleges and Universities under the Background of Internet +” (Project no. 17BY047).

References


