Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 2, pp. 864–870, DOI 10.12694/scpe.v26i2.4047

INTELLIGENT OPTIMIZATION AND RECOMMENDATION SYSTEM DESIGN FOR PERSONALIZED TRAINING PROGRAMS FOR MARATHON ATHLETES BASED ON MACHINE LEARNING

LUXIA GUO*

Abstract. This research focuses on developing an innovative machine learning-based intelligent optimization and recommendation system for marathon runners' personalized training schemes. The system aims to provide accurate and dynamically adjusted training guidance for athletes through real-time monitoring, training effect evaluation and intelligent recommendation. First, the system uses advanced wearable technology to achieve real-time monitoring of multiple physiological and athletic data during athlete training, including but not limited to heart rate variability, lactate threshold, and gait analysis. This data forms the basis of a personalized training program. Secondly, the support vector machine (SVM) algorithm is used to evaluate the training effect of the collected data. Finally, the system combines individual characteristics and the historical performance of athletes and generates personalized training plans through optimization strategies such as support vector machines. This process not only considers the short-term training goals but also considers the long-term sports career planning. Through algorithm modeling and computer simulation, it is found that the system can realize the continuous optimization of the training scheme in the process of continuous iteration. The intelligent system significantly improves athletes' training efficiency and competition performance compared to traditional training methods. This study provides a new perspective and practice path for intellectualization in sports training.

Key words: Machine learning; Marathon; Personalized training; Training effect; Support vector machine.

1. Introduction. With the rapid development of science and technology, the application of machine learning in sports training is increasing daily, especially in optimizing and recommending personalized training programs for marathon runners. Marathon training not only requires athletes to have good physical strength and endurance but also needs scientific training methods and strategies to improve the training effect and competition performance. Traditional training methods often lack individualized adjustment for individual differences, and the introduction of machine learning technology makes the training program more accurately adapt to the specific situation of each athlete.

Real-time monitoring system plays a crucial role in marathon training. Through wearable devices such as GPS watches and heart rate monitors, coaches and athletes can obtain real-time training data, including key indicators such as movement trajectory, speed and heart rate [1]. These data help coaches understand the athletes' training status in time, adjust the training intensity, and prevent overtraining and sports injuries. Machine learning algorithms can process and analyze training data, extract useful information, and provide a scientific basis for athletes' training [2]. For example, by analyzing an athlete's historical training data, a machine learning model can predict an athlete's optimal training load and recovery time to optimize training schedules. Support vector machine (SVM) is a commonly used machine learning algorithm suitable for small samples of nonlinear and high-dimensional pattern recognition problems. In the marathon training effect evaluation, SVM can be used to distinguish the training data of elite athletes and ordinary athletes, identify the key factors affecting the training effect, and provide personalized training feedback for athletes.

There is a close relationship between marathon training real-time monitoring systems, machine learning and support vector machines [3]. Real-time monitoring systems provide the data source, machine learning algorithms process and analyze the data, and support vector machines are used to evaluate the training effects [4]. This combination can improve the science and effectiveness of training and realize the real-time monitoring and instant feedback of athletes' training status to achieve personalized training.

*School of Physical Education, Shanxi Vocational University of Engineering Science and Technology, Jinzhong 030600, China (Corresponding author, jtxytyz@163.com)

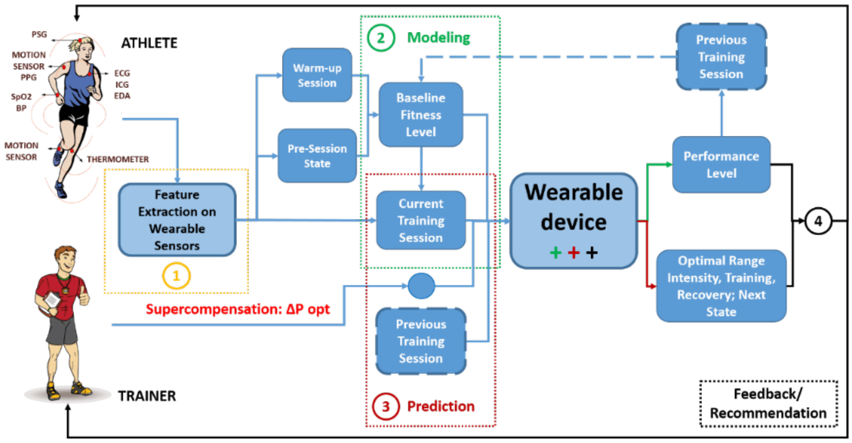

Fig. 2.1: *Architecture of real-time monitoring and guidance system for marathon runners during training.*

This paper aims to design an intelligent optimization and recommendation system based on machine learning for marathon runners' personalized training schemes [5]. The system will integrate the functions of real-time monitoring, data analysis and effect evaluation and verify the effectiveness and practicability of the system by establishing algorithm models and conducting simulation tests [6]. It is expected that the application of this system can significantly improve the training efficiency and competition performance of marathon runners while reducing the risk of sports injuries.

2. Design and establishment of a real-time monitoring system for marathon training.

2.1. Methods and basis for real-time monitoring of athletes' heart rate information. Heart rate is a physiological parameter often used in sports activities and a sensitive metabolic state of the body. Therefore, this topic takes heart rate as a "window" to conduct in-depth research on the body's metabolic state [7]. The study of the causal relationship between exercise load and the change in body function found that the heart rate significantly correlates with exercise intensity, oxygen uptake and energy metabolism. With the increase in load intensity, the body's oxygen consumption and oxygen uptake increase, and the heartbeat also increases, so when the exercise intensity changes, the heartbeat will also change [8]. If the athlete's heartbeat can be monitored in real-time and the heartbeat signal generated by the body for the exercise load can be transmitted to the athlete, then the intensity of the exercise can be adjusted according to their actual situation. This way, you get the best results. This paper now chooses a portable heartbeat monitor to realize wireless heartbeat monitoring. In this way, the athlete's heartbeat is fully integrated into the exercise state.

2.2. Realization of real-time monitoring of positioning during movement. In training for marathon and cross-country running, to accurately grasp the running position of the runners at a certain point in time and the geographical environment, the usual practice is to be driven by the coach. This makes it easy for vehicles to run during training, but it does not ensure the correct guidance of each athlete. There are no new monitoring methods in the training for marathon and cross-country running in our country. The use of GPS for real-time positioning, the specific position of the players displayed on the monitoring screen can facilitate the coach to understand the status of the players [9]. The system designed in this paper presents the player's position directly on the background of the electronic map and regards the player himself as a geographical target. By measuring, analyzing and simulating the spatial information that the players rely on, the training state of the players can be obtained indirectly. Figure 2.1 shows the basic framework for on-site monitoring of a marathon runner's training flow.

2.3. Methods of information exchange between athletes and coaches during training. The development trend of domestic competitive sports is how to realize instant communication with athletes during physical education [10]. According to the actual conditions of the competition, it determines whether the

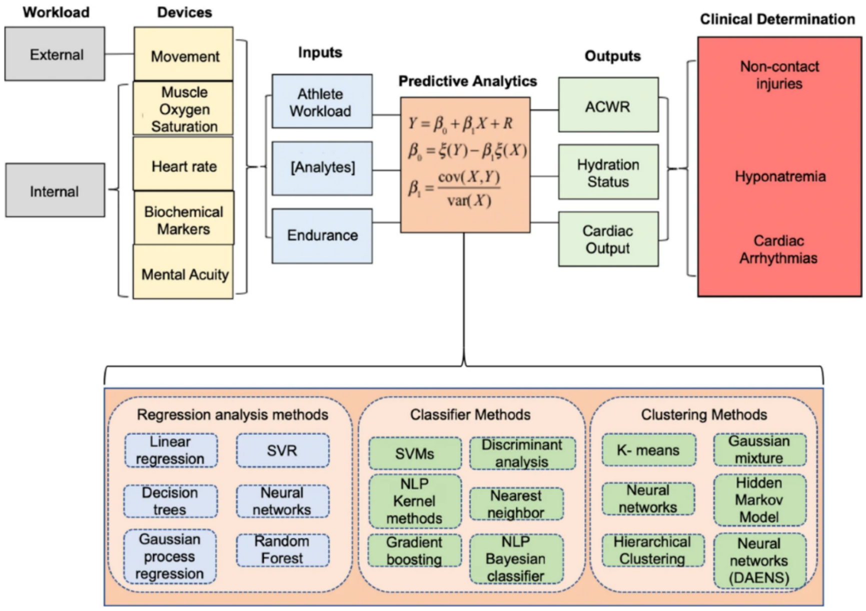

Fig. 2.2: *Sensor architecture of athlete physiological information monitoring system.*

wireless data transmission can be carried out efficiently, whether the communication coverage area and the function area can meet the actual needs, and whether the investment and operation cost of the communication equipment can be borne. The current mainstream mobile communication technology includes analog cluster, digital cluster, CDPD, GPRS, CDMA, and synchronous satellite communication. Trunking digital mobile communication is this system's most suitable communication mode, so this paper decides to use this method to realize the information exchange in physical exercise. By wearing small headsets, athletes can listen to the coach's instructions to achieve better exercise results and adapt to the game's needs.

2.4. Information Integration System Architecture. The athletes' heart rate, position, surrounding environment and other information are transmitted to the training monitoring center through the wireless network using the method of multi-sensor information integration and information fusion [11]. Using GIS or mobile geographic information system, mobile phone mapping, etc., the player's heart rate information, location information, traffic conditions and other information through the road classification method in an image to display on the monitoring terminal. The speed, heart rate and other parameters of the athletes were calculated. Figure 2.2 shows the sensor architecture of the athlete's physiological information monitoring system (see Wearable sensors for monitoring the physiological and biochemical profile of the athlete). In the training process of marathon runners, the information is integrated, the computer information analysis system is built, and the basic information of athletes is managed by database technology. At the same time, the athlete's initial state can be scientifically guided and portable monitoring devices can be used [12]. The real-time exchange of information between athletes and coaches is realized to achieve timely and efficient training guidance for coaches.

3. Evaluation of athlete training effect of machine learning algorithm.

3.1. Machine learning algorithm. If $G(c)$ is a probability metric located in space c, and the set of functions is $P(c, e)$ and $e \in \Theta$, then the least risk functional is the purpose of machine learning. Its expression is:

$$S(e) = \int P(c, e) dG(c) \tag{3.1}$$

The random variable $G(c)$ is uncertain, but there is a fixed random sample. Suppose u is the output of the training system, which can be o or 1. Suppose $g(v,e)$, where $e \in \Theta$ represents a set of indicator functions, and the expression to obtain the loss function is:

$$H(u, g(v,e)) = \begin{cases} 0, & u = g(v,e) \\ 1, & u \neq g(v,e) \end{cases} \tag{3.2}$$

The risk function is used to determine the index function $g(v,e)$ and the output probability of the trainer, and the function with the minor identification error is obtained from the given sample set. An evaluation problem is a nonlinear problem transformed into a high-dimensional feature space based on a given nonlinear function [13]. Then the optimal classification hyperplane is constructed.

By using $Z(v_i \cdot v_j)$ kernel function satisfying Mercer's condition, the nonlinear classification problem on the best classification plane can be transformed into a linear classification problem.

$$Q(\varphi) = \frac{1}{2} \sum_{i,j=1}^{I} \varphi_i \varphi_j u_i u_j Z(v_i \cdot u_j) - \sum_{i=1}^{I} \varphi_i \tag{3.3}$$

The optimal classification function is as follows:

$$g(v) = \operatorname{sgn}\left\{ \sum_{i=1}^{l} \varphi_i u_i Z(v_i \cdot v) + \varepsilon^* \right\} \tag{3.4}$$

A multivariable support vector machine method based on difference is proposed. A base-based method is selected to evaluate the athletes' training performance with physical kernel function quality. Its formula is:

$$Z(v, v_i) = \exp\left(-\frac{\|v - v_i\|^2}{2\sigma^2} \right) \tag{3.5}$$

3.2. Evaluation of athletes' training effect. $V = \{v_1, v_2, \cdots, v_n\}$ and $U = \{u_1, u_2, \cdots, u_n\}$ are a group of physiological parameters that can reflect the effectiveness of an athlete's training, including heart rate, oxygen uptake, hemoglobin and sarcosine. For the metric index set U, the matrix $E = (e_{ij})_{n\times m}$ represents the exponential matrix of the athlete training sample set V. For the measured index value u_j, the athlete training sample v_i is expressed as $e_{ij} = u_j(v_i)\,(i = 1, 2, \cdots, n; j = 1, 2, \cdots, m)$. The normalized exponential matrix E to $S = (S_{xj})_{n\times m}$ can effectively overcome the difference between the evaluation results of each dimension. Take the measured value vector $(s_{i1}, s_{i2}, \cdots, s_{mm})$ of the athlete's training sample v_i under the physiological metric index u_j as the input. The evaluation result z_i of the athlete's exercise effect v_i as the output unit, then there is a nonlinear correspondence between the normalized matrix S and the evaluation result, and it is denoted asG:

$$z_i = G(s_y) \tag{3.6}$$

A sample $v_i = (s_{il}, s_{i2}, \cdots, s_{im})$ of sports performance is taken as the input vector of the support vector machine [14]. The learning sample set is constructed by taking the training effectiveness sampling evaluation value as a regression index. The following regression function is obtained in $F = \{(v_i, z_i)\}_i^n$:

$$z = \sum_{k=1}^{s} (\varphi_k - \varphi_k^*) Z(v, v^k) + \varepsilon \tag{3.7}$$

where v^k and s are the number of support vectors and the number of support vectors, and φ_k is a Lagrange multiplier, $v^k = (s_{k1}, s_{k2}, \cdots, s_{kn}), k = 1, 2, \cdots, s$. Under the measured physiological index u_j, the index value $(s_{11}, s_{12}, \cdots, s_{mm})$ of the athlete training sample v_i and the evaluation value G of the training effect are nonlinear mapped [15]. Svm-based learning strategies were introduced into physical education teaching, and the quality of physical education was evaluated (Fig. 3.1). The steps to evaluate the effectiveness of an athlete's training are:

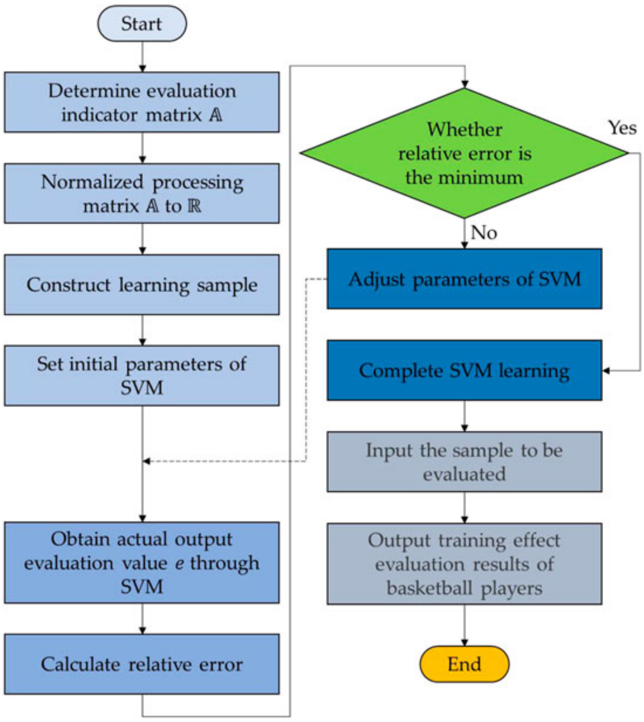

Fig. 3.1: *Structure of athletes' training effect evaluation.*

1. The evaluation index matrix E was established according to the physiological parameters related to physical education teaching effectiveness;
2. The index matrix E, which evaluates the training effectiveness of athletes, is transformed into a normalized matrix S;
3. Based on athletes' training effect sampling $v_i = (s_{i1}, s_{i2}, \cdots, s_{im})$ and evaluation index z_i, a learning sampling set C was constructed, and then SVM was trained and tested.
4. The corresponding regression equation is obtained using a radial basis as the kernel of the support vector machine. The selection index standard for evaluating athletes' exercise effectiveness is set up. These include:

$$\begin{cases} MAE = \frac{1}{h}\sum_{i=1}^{h}\left|z_p^i - z_{SVM}^i\right| \\ MSE = \frac{1}{h}\sum_{i=1}^{h}\left(z_p^i - z_{SVM}^i\right)^2 \end{cases} \tag{3.8}$$

   MAE and MSE are the mean absolute error and mean-variance of the test sample, z_p^i and z_{SVM}^h are the expert evaluation value of the test sample and the SVM calculation result of the confirmed sample, and h represents the total number of samples to be evaluated.
5. The learning process of the support vector machine terminates after the desired parameters are obtained [16]. The trained SVM was used to evaluate the training effect. By analyzing the physiological index vector $(s_{i1}, s_{i2}, \cdots, s_{im})$ of training effect sampling v_i of the evaluated athletes, the training effect evaluation result z_{STM}^i of SVM was obtained.

4. Case analysis. In this study, 10 marathon runners were selected as subjects. Let players participate in football, basketball, volleyball, swimming, running, 5 sports training [17]. Heart rate and lung capacity were collected during each exercise. Morning venous blood was collected at each stage through 10000 test data, including 2000 training samples and 8000 training samples. This program can be well applied to Windows and complete SVM classification and regression problems. Finally, the SVM was divided into 10 groups, each

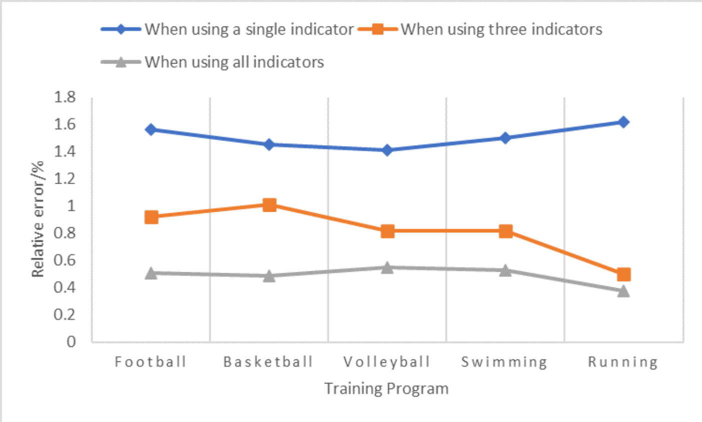

Fig. 4.1: *Relative error for different index quantities.*

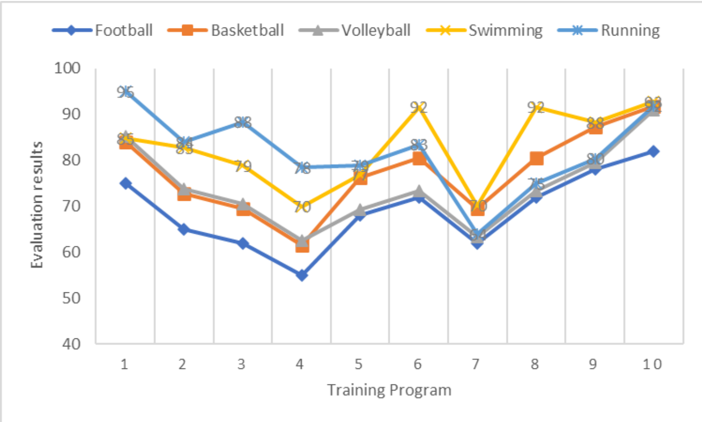

Fig. 4.2: *Evaluation results of athletes' training effect.*

group of 200, and finally, the parameters of the SVM with C=150 and $\sigma 2$=0.016 were obtained. It was used to train SVM, and the number of support vectors and parameters b=30 and -0.218 of the regression equation were obtained. Heart rate, maximal oxygen uptake, hemoglobin, creatine kinase and blood lactic acid were selected as physiological test items to evaluate the athletic performance of athletes [18]. When only heart rate was used as the physiological test index, 3 physiological test indexes, such as heart rate, maximal oxygen uptake and hemoglobin, were used, and the relative error of 5 physiological test indexes was applied to the evaluation of the training effect of each item was statistically treated. According to the experimental data in Figure 4.1, introducing more physiological parameters to evaluate training effects can improve the diversity of evaluation and thus improve the accuracy of evaluation results.

The evaluation method proposed in this paper is applied to the comprehensive evaluation of 10 athletes in 5 events. The results are shown in Figure 4.2.

The evaluation criteria for evaluating the training effect of athletes are selected. The relative deviation of the evaluation results is less than 1%. The relative deviation of physical fitness and professional skill methods in evaluating athletes' various sports skills is less than 1%. The comparison shows that the method proposed in this paper has a good performance in evaluating the athlete's performance. The method in this paper can help the coach to guide the training of athletes better to provide a reference for the training practice in the future.

5. Conclusion. The training effect is directly related to the final result of the competition, and the competition performance of the players is affected by many factors such as training mode and personal quality, and its performance has excellent fluctuations. The physical function of athletes is taken as the evaluation object, and the optimal SVM model is used to evaluate sports. The proposed evaluation model is empirically studied, taking marathon runners as an example. This provides a theoretical basis for improving the training level of athletes.

REFERENCES

[1] De Leeuw, A. W., van der Zwaard, S., van Baar, R., & Knobbe, A. (2022). Personalized machine learning approach to injury monitoring in elite volleyball players. European journal of sport science, 22(4), 511-520.
[2] Qiu, S., Hao, Z., Wang, Z., Liu, L., Liu, J., Zhao, H., & Fortino, G. (2021). Sensor combination selection strategy for kayak cycle phase segmentation based on body sensor networks. IEEE Internet of Things Journal, 9(6), 4190-4201.
[3] Rigamonti, L., Albrecht, U. V., Lutter, C., Tempel, M., Wolfarth, B., & Back, D. A. (2020). Potentials of digitalization in sports medicine: a narrative review. Current sports medicine reports, 19(4), 157-163.
[4] Nath, R. K., Thapliyal, H., Caban-Holt, A., & Mohanty, S. P. (2020). Machine learning based solutions for real-time stress monitoring. IEEE Consumer Electronics Magazine, 9(5), 34-41.
[5] Husnain, A., Hussain, H. K., Shahroz, H. M., Ali, M., & Hayat, Y. (2024). A Precision Health Initiative for Chronic Conditions: Design and Cohort Study Utilizing Wearable Technology, Machine Learning, and Deep Learning. International Journal of Advanced Engineering Technologies and Innovations, 1(2), 118-139.
[6] Jowitt, H. K., Durussel, J., Brandon, R., & King, M. (2020). Auto detecting deliveries in elite cricket fast bowlers using microsensors and machine learning. Journal of sports sciences, 38(7), 767-772.
[7] Rana, M., & Mittal, V. (2020). Wearable sensors for real-time kinematics analysis in sports: A review. IEEE Sensors Journal, 21(2), 1187-1207.
[8] Li, B., & Xu, X. (2021). Application of artificial intelligence in basketball sport. Journal of Education, Health and Sport, 11(7), 54-67.
[9] Nwachukwu, B. U., Beck, E. C., Lee, E. K., Cancienne, J. M., Waterman, B. R., Paul, K., & Nho, S. J. (2020). Application of machine learning for predicting clinically meaningful outcome after arthroscopic femoroacetabular impingement surgery. The American Journal of Sports Medicine, 48(2), 415-423.
[10] Worsey, M. T., Espinosa, H. G., Shepherd, J. B., & Thiel, D. V. (2020). An evaluation of wearable inertial sensor configuration and supervised machine learning models for automatic punch classification in boxing. IoT, 1(2), 360-381.
[11] Johnson, W. R., Mian, A., Robinson, M. A., Verheul, J., Lloyd, D. G., & Alderson, J. A. (2020). Multidimensional ground reaction forces and moments from wearable sensor accelerations via deep learning. IEEE Transactions on Biomedical Engineering, 68(1), 289-297.
[12] Giles, B., Kovalchik, S., & Reid, M. (2020). A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis. Journal of sports sciences, 38(1), 106-113.
[13] O'Driscoll, R., Turicchi, J., Hopkins, M., Horgan, G. W., Finlayson, G., & Stubbs, J. R. (2020). Improving energy expenditure estimates from wearable devices: A machine learning approach. Journal of Sports Sciences, 38(13), 1496-1505.
[14] Wilkens, S. (2021). Sports prediction and betting models in the machine learning age: The case of tennis. Journal of Sports Analytics, 7(2), 99-117.
[15] Incel, O. D., & Bursa, S. Ö. (2023). On-device deep learning for mobile and wearable sensing applications: A review. IEEE Sensors Journal, 23(6), 5501-5512.
[16] Ramkumar, P. N., Luu, B. C., Haeberle, H. S., Karnuta, J. M., Nwachukwu, B. U., & Williams, R. J. (2022). Sports medicine and artificial intelligence: a primer. The American Journal of Sports Medicine, 50(4), 1166-1174.
[17] Taj, I., & Zaman, N. (2022). Towards industrial revolution 5.0 and explainable artificial intelligence: Challenges and opportunities. International Journal of Computing and Digital Systems, 12(1), 295-320.
[18] Alexopoulos, K., Nikolakis, N., & Chryssolouris, G. (2020). Digital twin-driven supervised machine learning for the development of artificial intelligence applications in manufacturing. International Journal of Computer Integrated Manufacturing, 33(5), 429-439.

Edited by: Hailong Li
Special issue on: Deep Learning in Healthcare
Received: May 28, 2024
Accepted: Jul 18, 2024