Research article

Powerlifting score prediction using a machine learning method

Vinh Huy Chau*

Ho Chi Minh City University of Physical Education and Sport, Ho Chi Minh City 700000, Vietnam

* **Correspondence:** Email: huyvcc@upes.edu.vn.

Abstract: This research discusses an interesting topic, using artificial intelligence methods to predict the score of powerlifters. We collected the characteristics of powerlifters, and then used the reservoir computing extreme learning machine to build a predictive model. In order to further improve the prediction results, we propose a method to optimize the reservoir computing extreme learning machine using the whale optimization algorithm. Experimental results show that our proposed method can effectively predict the score of powerlifters with the coefficient of determination value is 0.7958 and root-mean-square error of prediction value is 16.73. This provides a reliable basis for experts to judge the results before the competition.

Keywords: artificial intelligence; machine learning; powerlifting; score prediction

1. Introduction

In recent years, along with the development of economy and society, sport is more and more interested in people. Sports help people to increase resistance, reduce work stress and enhance solidarity among people, etc. [1–3]. According to the World Health Organization, each year about 2 million people die from lack of exercise. Lack of exercise will reduce the body's immunity and make adolescents develop abnormally [4,5].

Powerlifting is a popular sport. In competition, the impact on powerlifters' performance is mainly due to age, weight, fitness and psychology [6–8,13,14,17]. Therefore, the training methods of coaches for powerlifters are extremely important, and studying the factors that influence athletes' performance is an inseparable task in training process [9–11]. Solberg et al. [12] based on the data powerlifting data in the international competitions; they calculated the score of powerlifters at their peak performance, thereby giving the development trend of athletes, helping experts to evaluate more correctly about the athletes' abilities before playing. Latella et al. [13] analyzed the long-term strength adaptation of

powerlifters through 15 years of data. They evaluated the daily strength growth rate of 1897 powerlifters. The results show that men and women have similar strength adaptability, and the rate of strength gain differs by strength quartile for females. These findings provide new evidence for strength adaption and provide a reliable basis for powerlifting coaches. Latella et al. [14] pointed out that there is little understanding of the factors that affect the strength of powerlifters, and studied the strength of powerlifters through age and weight. The results provide important information for the factors affecting athletes' strength performance and provide basis for coaches in long-term training plans. Anton et al. [15] analyzed the strength performance of world-class powerlifters with age. They concluded that powerlifters of different ages performed differently, and that women's abilities declined faster. Jame et al. [16] studied a method to improving the performance of weaker and inexperienced powerlifters. After testing this method, the performance was significantly improved. Pearson et al. [17] investigated the influence of competition frequency on the strength of powerlifters within 12 months. The results show that powerlifters who participate in multiple competitions each year are more likely to achieve higher overall results. Coker et al. [18] reported that the success or failure of each lift attempt has a correlation with Wilks points. The limitation in this study is that they did not relate their findings to competitive outcomes.

Age, weight and psychology are critical to the performance of powerlifters. In a competition, the score of the previous competition content will directly affect the next content. Therefore, studying and predicting the performance of powerlifters is extremely important for calculating the individual scores. It can be used as a basis for choosing the right weight in the next competition. Traditional prediction methods usually have the disadvantages of low accuracy and unable to immediately predict the performance of athletes in the competition [19]. In recent years, artificial intelligence has developed dramatically and is applied in many fields such as performance fitness and error metrics [20,21]. So, the application of artificial intelligence in predicting the score of athletes has profound implications. Extreme learning machine (ELM) algorithm [22,23] is a novel method of artificial intelligence. This algorithm has very fast training speed and high analysis accuracy, and has attracted the attention of many researchers. Huang et al. [24] proposed the kernel ELM to further improve the generalization ability of ELM and it perform better in classification and regression problems. Tang et al. [25] brought the unsupervised learning algorithm into the ELM theory and proposed a multilayer ELM algorithm. The algorithm does not require fine-tuning, and has better learning efficiency than typical deep learning algorithms. Compared with the current popular multi-layer learning method, this algorithm can achieve better and faster convergence. Deng et al. [26] proposed a regularized ELM algorithm, which added the structural risk minimization principle and the weighted least squares algorithm. The experiments show that the generalization performance of ELM has been significantly improved. Zheng et al. [27] used the maximum mixture correntropy criterion as an optimization algorithm to improve the kernel ELM, which overcomes the shortcoming of the minimum mean square error of the Gaussian assumption. Cao et al. [28] improved the ELM algorithm based on the maximum entropy criterion in order to alleviate the poor effect when processing with large outliers, and verified the effectiveness of their proposed method on more benchmark data sets. Ertugrul et al. [29] proposed a reservoir computing-ELM (RC-ELM) algorithm based on the reservoir computing theory. RC-ELM does not need to fine-tune any parameters. In most experimental results, RC-ELM performance is better than ELM.

Based on the age, weight and field score of powerlifters, we study the best performance of powerlifters and develop a model for predicting the best score. The model is based on the RC-ELM

algorithm, but the parameters of RC-ELM are initialized randomly, which leads to unstable model performance. Therefore, in order to further improve the prediction accuracy of the model, this study uses a novel whale optimization algorithm to optimize the parameters. Finally, we use the MATLAB software for modeling and testing to verify the performance of the proposed method.

2. Methods

2.1. Reservoir computing extreme learning machine

The function of a hidden layer neural network is shown in Eq (1)

$$f(\mathbf{x}) = \sum_{j=1}^{L} \beta_j g(\mathbf{w}_j \cdot \mathbf{x}_i + b_j) = \mathbf{t}_i \tag{1}$$

where, $\beta_j \in \mathbf{R}$ is the output weight, $\mathbf{w} = [\mathrm{w}_{j1}, \mathrm{w}_{j2}, ..., \mathrm{w}_{jn}]^T \in \mathbf{R}^n$ is the input weight, $b_j \in \mathbf{R}$ is the bias, $\mathbf{x}_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T \in \mathbf{R}^n$ is input data, $\mathbf{t}_i = [t_{i1}, t_{i2}, \cdots, t_{in}]^T \in \mathbf{R}^m$ is expected output; $g(\cdot)$ is the activation function, and L is number of hidden layer nodes.

The $f(\mathbf{x})$ is equal to Eq (2).

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \tag{2}$$

where,

$$\mathbf{H}(\mathbf{w}_1, \cdots, \mathbf{w}_L, b_1, \cdots, b_L, \mathbf{x}_1, \cdots, \mathbf{x}_N)$$
$$= \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L} \tag{3}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \qquad \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N \times m} \tag{4}$$

N is number of samples.
According to the proof of ELM, we can calculate the $\boldsymbol{\beta}$.

$$\boldsymbol{\beta} = \mathbf{H}^+ \mathbf{T} \tag{5}$$

where, $\mathbf{H}^+$ is the Moore-Penrose generalized inverse of $\mathbf{H}$.

In the reservoir computing extreme learning machine (RC-ELM) network, the ith neuron in the reservoir can be obtained as Eq (6) in Figure 1.

$$\mathbf{S}_i = \mathbf{w}\mathbf{x}_i + \mathbf{p}_{1..m,i}\mathbf{S}_{1..m} + b \tag{6}$$

where $\mathbf{S}$ is the summation of inputs of a neuron, $\mathbf{p}$ is the weights of the neurons in the reservoir, and m

is the number of neurons in the reservoir. The summation of inputs of a neuron in the reservoir can be obtained as Eq (7).

$$\mathbf{S}_i - \mathbf{p}_{1...m,i}\mathbf{S}_{1...m} = \mathbf{w}\mathbf{x}_i + b \tag{7}$$

Each of the summations of inputs of a neuron can be calculated by obtained as Eqs (8) and (9).

$$\mathbf{S}(1-\mathbf{p}) = \mathbf{w}\mathbf{x} + b \tag{8}$$

$$\mathbf{S} = inv(1-\mathbf{p}) * \mathbf{w}\mathbf{x} + b \tag{9}$$

Finally, the $\boldsymbol{\beta}$ can calculated by Eq (10).

$$\boldsymbol{\beta} = g(\mathbf{S})^+ \mathbf{T} \tag{10}$$
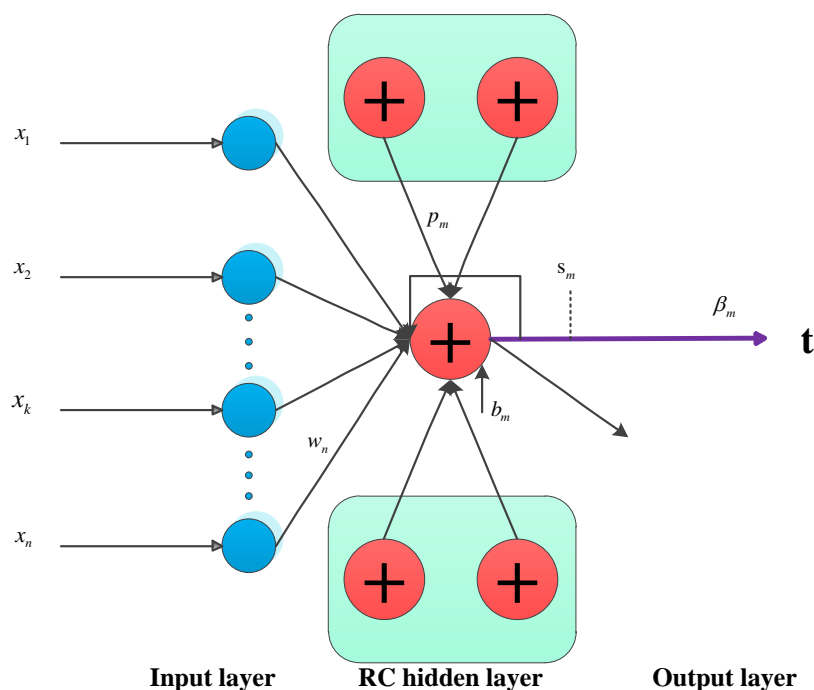
Figure 1. Summation of the ith neuron in the reservoir.

2.2. Whale optimization algorithm

Whale optimization algorithm (WOA) [30,31] is a new swarm intelligence optimization algorithm based on whale's unique hunting strategy. The WOA optimization process is divided into three phases: foraging, bubble-net preying and searching for food.

(a) Foraging

The whales find food by working with the swarm. When the food is detected, all individuals in the swarm will gradually approach the food. Assuming that the location of the food is the best target, the steps to approach this target is shown in Eq (11).

$$X_{i+1} = X_i - A\left|CX_i^* - X_i\right| \tag{11}$$

where i is the current number of iterations, X_i is the current position, X_{i+1} is the position after iteration i, X_i^* is the current optimal position of the all whales, and A and C are coefficients, these coefficients are calculated according to the Eqs (12) and (13).

$$A = 2at_1 - a \tag{12}$$

$$C = 2t_2 \tag{13}$$

where t_1 and t_2 are random numbers in the range of [0,1]. As the number of iterations increases, the value of a decreases linearly from 2 to 0.

(b) Bubble-net preying

At this phase, the predation behavior of whales includes two processes: shrink wrapping and spiral uprising. The probability of choosing between these two processes is 0.5. Through these two processes, we can achieve local optimization. This phase is represented mathematically according to Eq (14).

$$X_{i+1} = \begin{cases} X_i - A\left|CX_i^* - X_i\right|, & p < 0.5 \\ \left|X_i^* - X_i\right|e^{bl}\cos(2\pi l) + X_i^*, & p \geq 0.5 \end{cases} \tag{14}$$

where l is a random number in the range (−1, 1), p is a random number in the range (0,1), and b is a constant of the spiral.

(c) Searching for food

At this phase, each whale is assigned a random location for global search (when A > 1). The calculation of the random location is shown in Eq (15).

$$X_{i+1} = X_{rand} - A\left|CX_{rand} - X_i\right| \tag{15}$$

where X_{rand} is the random location.

2.3. Propose an optimal method for RC-ELM by WOA

As mentioned above, the weights and bias of the RC-ELM are randomly generated, which will greatly affect the model performance. Therefore, to improve the efficiency and increase the stability of the model, we propose a method to optimize the weights and bias of RC-ELM by using WOA (this method is called WOA- RC-ELM). The WOA-RC-ELM algorithm flowchart is showed in Figure 2.
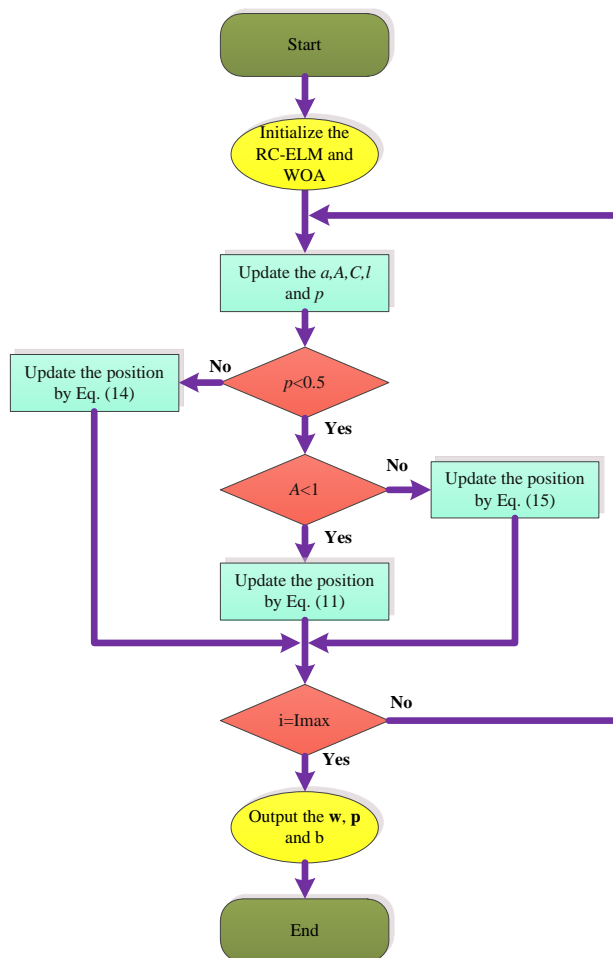
Figure 2. WOA-RC-ELM algorithm flowchart.

Algorithm 1: WOA-RC-ELM algorithm
Initialize the WOA-RC-ELM algorithm
for i = 1: Imax
    for j = 1: N (where N is the number of whale groups)
        Update a, A, C, l and p
        if $p < 0.5$
          if $A < 1$
            Update the position by Eq (11)
          else
            Update the position by Eq (15)
          end if
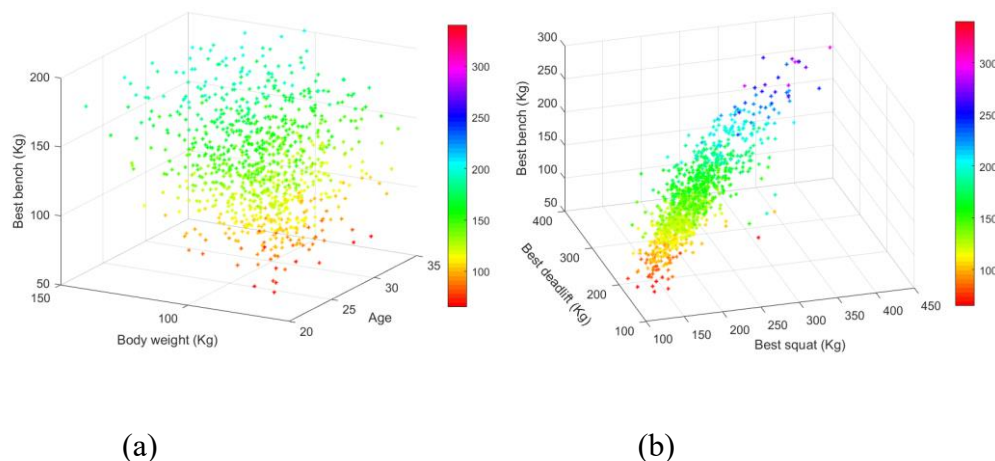   else
          Update the position by Eq (14)
        end if
    end for
end for
Output optimized w, p and b.

3. Results and discussions

3.1. Data collection and analysis

(a)   (b)

Figure 3. Relationship between data features.

We collected the scores of world-class powerlifters from 1972 to 2017 [32,33], such as the World Masters Powerlifting Championships, and the World Open Powerlifting Championships. There are a total of 5500 male powerlifters' features, and the age ranges from 20–35. These features include age, body weight, weight class, best squat, best bench, and best deadlift. We randomly selected 4500 samples as the training set and 1000 samples as the test set. The data is normalized to (0, 1). Choose the best bench as the output value, and the remaining features as the input value.

For the powerlifters, about 12 age is the initial stage of training development. Entering the age of 14 is a sensitive period, and proper training methods must be available at this stage. Around the age of 16, the powerlifters have developed basic skills. Over 20 years old is the peak period of performance [34].

Based on the features of powerlifters, we draw the relationship between them. Figure 3 shows this relationship. Figure 3(a) is the effect of age and weight on the best bench. We can see that as the weight value increases, the best bench value also increases. Powerlifters basically have the greatest strength and rich experience between the ages of 20 and 35. At this stage athletes can reach peak state and break their own records. For the mathematical modeling angle, the points in the Figure 3(a) are relatively discrete and nonlinear. Figure 3(b) is the relationship between best squat, best bench, and best deadlift. It can be seen that these features have a strong linear relationship. This implies that if the performance of the powerlifters is good at squat and deadlift, then the bench performance will also be good. Conversely, if the squat and deadlift performance is not good, then the bench is also performing poorly. Therefore, we can rely on these features to predict the performance of powerlifters during the competition, and it is necessary to build a mathematical model to calculate these scores.

3.2. Score prediction model

Based on the WOA-RC-ELM algorithm, we design a score prediction model for male powerlifters.

The coefficient of determination (R^2) and root-mean-square error of prediction (RMSEP) [35,36]0 were used to evaluate the model. The calculation of these two indicators is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n_{test}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{test}} (y_i - \overline{y})^2} \tag{14}$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_{test}} (y_i - \hat{y}_i)^2}{n_{test}}} \tag{15}$$

where y_i is the actual value, $\overline{y}$ is the mean value, and $\hat{y}_i$ is the predictive value, n_{test} is the number of samples. The closer the value of R^2 is to 1 and the smaller the value of RMSEP, the better the prediction effect of the model.

(a)   (b)
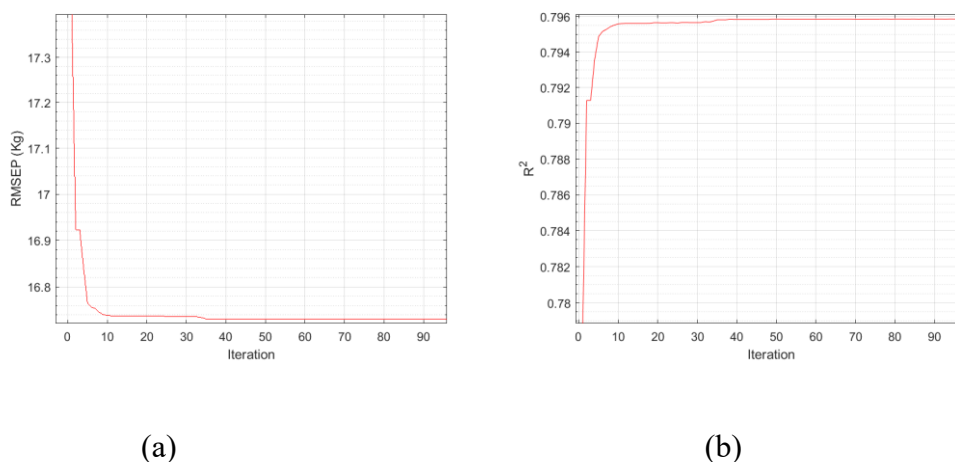
Figure 4. WOA-RC-ELM algorithm performance.

Figure 4 shows the performance of the WOA-RC-ELM algorithm. Where, the number of iterations is 100 and the parameter m of RC-ELM is 20. We can see that the WOA-RC-ELM algorithm converges quickly. When the number of iterations is 35, the global optimal state is reached. RMSEP decreased from 17.41 to 16.73, and R^2 increased from 0.7786 to 0.7958, the result proves the effectiveness of the proposed method. WOA can effectively optimize the multi-dimensional parameter matrix and improve the accuracy and stability of the RC-ELM algorithm. The results can be used to assist experts in predicting the score of powerlifters and lay the foundation for coaches' training methods.

3.3. Comparison of different methods

This study compares WOA-RC-ELM, RC-ELM and ELM, and Figure 5 shows the comparison results. We can see that the best performance is the WOA-RC-ELM method, with an

RMSEP of 16.73 and R^2 of 0.7958. The performance of RC-ELM is better than ELM method. The RMSEP and R^2 of RC-ELM method are 17.41 and 0.7786, respectively. The RMSEP of the ELM method is 17.65 and R^2 is 0.7728. The comparison results once again demonstrate the effectiveness of the proposed method. WOA can improve the performance of RC-ELM algorithm and proves that WOA is an effective optimization algorithm. The proposed model is a reliable method for predicting powerlifting performance, which can provide a reference for the best score of powerlifters before the competition.
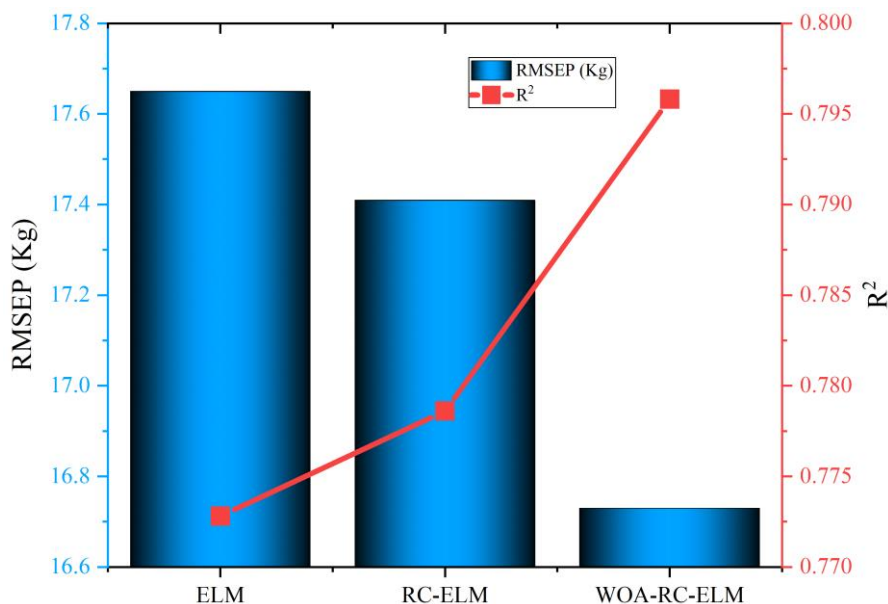
Figure 5. Comparison of different methods.

4. Conclusions

Male powerlifters of different ages and weights have different performances. This research established a score prediction model based on the characteristics of male powerlifters and a machine learning method. The model uses the age, body weight, weight class, best squat, best bench, and best deadlift features of powerlifters, and the WOA algorithm is used to improve the predictive ability of the model. Specifically, a 23 years old powerlifter with weight is 99 kg, our model predicts the best bench score is 141 kg, which is 1.5 kg different from the actual value. The results show that our proposed method can effectively predict the score of male powerlifters, which provides a reliable prediction method for powerlifting before competitions.

Acknowledgments

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. D. J. V. D. Windt, V. Sud, H. Zhang, A. Tsung, H. Huang, The effects of physical exercise on fatty liver disease, *J. Liver Res.*, **18** (2018), 89–101.

2. D. Song, D. S. F. Yu, P. W. C. Li, Y. Lei, The effectiveness of physical exercise on cognitive and psychological outcomes in individuals with mild cognitive impairment: A systematic review and meta-analysis, *Int. J. Nurs. Stud.*, **79** (2018), 155–164.

3. M. G. Benedetti, G. Furlini, A. Zati, G. L. Mauro, The effectiveness of physical exercise on bone density in osteoporotic patients, *Bio. Med. Res. Int.*, (2018), 4840531.

4. M. M. Ziaaldini, E. Marzetti, A. Picca, Z. Murlasits, Biochemical pathways of sarcopenia and their modulation by physical exercise: A narrative review, *Front. Med.*, **4** (2017), 167.

5. C. M. Hearing, W. C. Chang, K. L. Szuhany, T. Deckersbach, A. A. Nierenberg, L. G. Sylvia, Physical exercise for treatment of mood disorders: A critical review, *Curr. Behav. Neurosci. Rep.*, **3** (2016), 350–359.

6. L. W. Judge, L. J. Urbina, D. L. Hoover, B. W. Craig, L. M. Judge, B. M. Leitzelar, et al., The impact of competitive trait anxiety on collegiate powerlifting performance, *J. Strength Cond. Res.*, **30** (2016), 2399–2405.

7. M. Mottaghi, A. Atarodi, Z. Rohani, The relationship between coaches' and athletes' competitive anxiety and their performance, *Iran. J. Psychiatry Behav. Sci.*, **7** (2013), 68–76.

8. H. J. Pritchard, D. A. Tod, M. J. Barnes, J. W. Keogh, M. R. McGuigan, Tapering practices of New Zealand's elite raw powerlifters, *J. Strength Cond. Res.*, **30** (2016), 1796–1804.

9. M. Zawadka, M. S. Paszkowska, P. Gawda, E. Lukasik, J. Smolka, M. Jablonski, What factors can affect lumbopelvic flexion-extension motion in the sagittal plane?: A literature review, *Hum. Mov. Sci.*, **58** (2018), 205–218.

10. A. C. Fry, W. J. Kraemer, M. H. Stone, L. P. Koziris, J. T. Thrush, S. J. Fleck, Relationships between serum testosterone, cortisol and weightlifting performance, *J. Strength Cond. Res.*, **14** (2000), 338–343.

11. I. Sandau, M. Witt, The influence of stress factors on joint loading of lower extremities in weightlifting, *Ger. J. Sports Med./Deutsch. Z. Sportmed.*, **69** (2018), 1–10.

12. P. A. Solberg, W. G. Hopkins, G. Paulsen, T. A. Haugen, Peak age and performance progression in world-class weightlifting and powerlifting athletes, *Int. J. Sports Physiol. Perform.*, (2019), 1–24.

13. C. Latella, W. P. Teo, J. Spathis, D. V. D. Hoek, Long-term strength adaptation: A 15-year analysis of powerlifting athletes, *J. Strength Cond. Res.*, **34** (2020), 2412.

14. C. Latella, D. V. D. Hoek, W. P. Teo, Factors affecting powerlifting performance: an analysis of age-and weight-based determinants of relative strength, *Int. J. Perform. Anal. Sport*, **18** (2018), 532–544.

15. M. M. Anton, W. W. Spirduso, H. Tanaka, Age-related declines in anaerobic muscular performance: Weightlifting and powerlifting, *Med. Sci. Sports Exercise*, **36** (2014), 143–147.

16. L. P. James, P. Comfort, T. J. Suchomel, V. G. Kelly, E. M. Beckman, G. G. Haff, The impact of power clean ability and training age on adaptations to weightlifting-style training, *J. Strength Cond. Res.*, **33** (2019), 2936–2944.

17. J. Pearson, J. G. Spathis, D. J. V. D. Hoek, P. J. Owen, J. Weakley, C. Latella, Effect of competition frequency on strength performance of powerlifting athletes, *J. Strength Cond. Res.*, **34** (2020), 1213–1219.

18. N. A. Coker, A. N. Varanoske, K. M. Baker, D. L. H. Vaughn, A. J. Wells, Predictors of competitive success of national-level powerlifters: A multilevel analysis, *Int. J. Perform. Anal. Sport*, **18** (2018), 796–805.

19. S. K. Travis, M. C. Zourdos, C. D. Bazyler, Weight selection attempts of elite classic powerlifters, *Percept. Mot. Skills*, 2020.

20. M. Z. Naser, A. Alavi, Insights into performance fitness and error metrics for machine learning, preprint, arXiv:2006.00887.

21. A. Botchkarev, Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology, preprint, arXiv:1809.03006.

22. G. B. Huang, Q. Y. Zhu, C. K. Siew, Extreme learning machine: Theory and applications, *Neurocomputing*, **70** (2006), 489–501.

23. M. Rocca, C. Perna, Nonlinear autoregressive sieve bootstrap based on extreme learning machines, *Math. Biosci. Eng.*, **17** (2019), 636–653.

24. G. B. Huang, H. Zhou, X. Ding, R. Zhang, Extreme learning machine for regression and multiclass classification, *IEEE Trans. Syst., Man, Cybern., Part B (Cybern.)*, **42** (2011), 513–529.

25. J. Tang, C. Deng, G. B. Huang, Extreme learning machine for multilayer perceptron, *IEEE Trans. Neural Networks Learn. Syst.*, **27** (2015), 809–821.

26. W. Deng, Q. Zheng, L. Chen, Regularized extreme learning machine, *2009 IEEE Symp. Comput. Intel. Data Min.*, (2009), 389–395.

27. Y. Zheng, B. Chen, S. Wang, W. Wang, W. Qin, Mixture correntropy-based kernel extreme learning machines, *IEEE Trans. Neural Networks Learn. Syst.*, 2020.

28. J. Cao, H. Dai, B. Lei, C. Yin, H. Zeng, A. Kummert, Maximum correntropy criterion-based hierarchical one-class classification, *IEEE Trans. Neural Networks Learn. Syst.*, 2020.

29. O. F. Ertugrul, A novel randomized machine learning approach: Reservoir computing extreme learning machine. *Appl. Soft Comput.*, **94** (2020), 106433.

30. S. Mirjalili, A. Lewis, The whale optimization algorithm. *Adv. Eng. Software*, **95** (2016), 51–67.

31. L. L. Li, J. Sun, M. L. Tseng, Z. G. Li, Extreme learning machine optimized by whale optimization algorithm using insulated gate bipolar transistor module aging degree evaluation, *Expert Syst. Appl.*, **127** (2019), 58–67.

32. https://www.powerlifting.sport/championships/results.html

33. https://www.kaggle.com/open-powerlifting/powerlifting-database

34. T. Yang, Study on the growth rule of the olympic men's weightlifting champion and winning factors. *Sichuan Sports Sci.*, **38** (2019), 54–58.

35. B. T. Le, D. Xiao, Y. Mao, D. He, Coal analysis based on visible-infrared spectroscopy and a deep neural network, *Infrared Phys. Technol.*, **93** (2018), 34–40.

36. D. Xiao, B. T. Le, Rapid analysis of coal characteristics based on deep learning and visible-infrared spectroscopy, *Microchem. J.*, (2020), 104880.