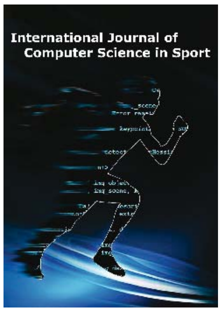
International Journal of Computer Science in Sport

Volume 18, Issue 3, 2019

Journal homepage: http://iacss.org/index.php?id=30

sciendo

DOI: 10.2478/ijcss-2019-0019



The effects of age and body weight on powerlifters: An analysis model of powerlifting performance based on machine learning

Vinh Huy Chau[1], Anh Thu Vo[1], Ba Tuan Le[2]

[1] Ho Chi Minh City University of Physical Education and Sport, Vietnam

[2] Institute of Research and Development, Duy Tan University, Vietnam

Abstract

As a up and coming sport, powerlifting is gathering more and more attetion. Powerlifters vary in their strength levels and performances at different ages as well as differing in height and weight. Hence the questions arises on how to establish the relationship between age and weight. It is difficult to judge the performance of athletes by artificial expertise, as subjective factors affecting the performance of powerlifters often fail to achieve the desired results. In recent years, artificial intelligence has made groundbreaking strides. Therefore, using artificial intelligence to predict the performance of athletes is among one of many interesting topics in sports competitions. Based on the artificial intelligence algorithm, this research proposes an analysis model of powerlifters' performance. The results show that the method proposed in this paper can predict the best performance of powerlifters. Coefficient of determination-R2=0.86 and root-mean-square error of prediction-RMSEP=20.98 demonstrate the effectiveness of our method.

KEYWORDS: POWERLIFTING PERFORMANCE, ANALYSIS MODEL, ARTIFICIAL INTELLIGENCE, KERNEL EXTREME LEARNING MACHINE, ARTIFICIAL BEE COLONY.

Introduction

Sports have many health benefits and are popular with the general public (Eime et al., 2016; Andronikos et al., 2016; Coker et al., 2018; Smith et al., 2018; Carey et al., 2018). In sports, powerlifting is loved by many people. This sport represents strength. Therefore, research on powerlifting is an interesting topic, especially the effect of age and body weight in powerlifting on athletes' performance. Anton et al. (2004) studied the powerlifting performance with increasing age in male and female athletes. The results show that age in both male and female athletes directly affect powerlifting ability and females powerlifting ability declines faster than males. Durguerian et al. (2016) investigated the weight loss, performance and state of mind of high-level powerlifters. In the experiment, they studied the diet intake of fast weight loss for 11 powerlifters. The findings indicate that food restrictions neither impair the powerlifting performance nor reduce energy and micronutrient intake to result in rapid body weight loss. Solberg et al. (2019) used the results of the 1998-2017 World Powerlifting Championships and the Olympic Games to quantify the best performances of powerlifters at peak ages. Buckingham et al. (2014) studied the influence of size-weight illusion on powerlifting. The research demonstrates that the size-weight illusion does not appear to affect exercise sports performance, which means that perceptual illusions are unlikely to affect one’s ability to persevere with powerlifting. James et al. (2018) considered the effects of power clean ability and training age on the performance of powerlifters. It turns out that the weak, inexperienced powerlifters have significantly increased their ability after concentrated training. If the coach pays more attention to the weaker, inexperienced athletes they can quickly improve the training effect. As far as we know, so far, no studies have been reported to predict the performance of powerlifters by age and weight.

For subjective factors affecting the performance of powerlifters, it is difficult to judge the performance of athletes by artificial expertise, often failing to achieve the desired results. In recent years, the advancement of artificial intelligence has promoted rapid development in all fields. Therefore, using artificial intelligence to predict the performance of athletes has become a very interesting topic in sports competitions (Lienhart, Einfalt, & Zecha, 2018). Huang et al. (2012) proposed a new artificial intelligence algorithm-kernel extreme learning machine (KELM). It is an improvement of the extreme learning machine (ELM) algorithm (Huang, Zhu, & Siew, 2006). The algorithm applies a kernel function and that does not need to determine the number of hidden layer nodes and activation functions. Consequently, KELM is much faster than traditional machine learning algorithms (such as support vector machines, BP neural networks). In addition, KELM has the advantages of fast convergence and good generalization. In recent years, the algorithm has been applied widely (Wang et al., 2019; Wang, Song, & Ma, 2017; Zhao et al., 2017).

Artificial bee colony algorithm (ABC) is an act of mimicking honey bees looking for nectar. The algorithm was proposed by Karaboga & Basturk (2007). The ABC algorithm has a simple structure, and the ABC search capability is stronger than the particle swarm optimization algorithm. Thus, it has been applied to many optimization problems (Karaboga et al., 2014; Mao et al., 2019; Xue et al., 2018; Sun, Chen, & Zhang, 2018). Since the regularization coefficient and kernel function matrix parameter in the KELM algorithm are determined by human's experience, this makes it difficult to determine the optimal combination. Therefore, this research will introduce the ABC algorithm to optimize the KELM network.

This paper will analyze the effects of age and body weight on the performance of male powerlifters and explore the relationship between them. Based on these relationships, the artificial intelligence method is used to establish a male powerlifting performance analysis model. The analysis was performed using MATLAB software to verify the validity of the model.

Methods

Dataset

This research collected information on 1,700 male powerlifters from the Kaggle website (https://www.kaggle.com/open-powerlifting/powerlifting-database). The data is information from different world powerlifting competitions in the period of 1972 to 2017, such as the World Open Powerlifting Championships and the World Masters Powerlifting Championships, including the athlete's age, body weight and best squat. The dataset is divided into 3 subsets which are training set, verification set, and test set.

Kernel Extreme Learning Machine

Suppose we have N samples $(\mathbf{x}_i, \mathbf{t}_i)$, $\mathbf{x}_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T \in \mathbf{R}^n$, $\mathbf{t}_i = [t_{i1}, t_{i2}, \cdots, t_{in}]^T \in \mathbf{R}^m$, the ELM neural network f_{ELM} is represented as below:

$$f_{\mathrm{ELM}}(\mathbf{x}) = \sum_{j=1}^{L} \beta_j g(\mathbf{w}_j \cdot \mathbf{x}_i + b_j) = \mathbf{t}_i\,, i = 1,\cdots,N; b_j, \beta_j \in \mathbf{R} \tag{1}$$

here, L is the number of cells; $\mathbf{w} = [w_{j1}, w_{j2}, \ldots, w_{jn}]^T \in \mathbf{R}^n$ is input weight, β_j is the output weight, b_j is the bias, and $g(x)$is the activation function.

Then, the f_{ELM}can be expressed as

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

here,

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

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

here, $\mathbf{H}$ is the output matrix, $\mathbf{T}$ is the expected output, and $\boldsymbol{\beta}$ is the output weight.

According to (Huang, Zhu, & Siew, 2006), we can calculate the output weight matrix $\boldsymbol{\beta}$.

$$\boldsymbol{\beta} = \mathbf{H}^{+}\mathbf{T}=\mathbf{H}^T(\frac{\mathbf{I}}{C} + \mathbf{HH}^T)^{-1}\mathbf{T} \tag{5}$$

Here, C is the regularization coefficient. Then, a kernel matrix $\boldsymbol{\omega}$ with kernel function K can be established by applying Mercer’s conditions on ELM.

$$\boldsymbol{\omega} = \mathbf{HH}^T \tag{6}$$

Finally, according to (5) and (6), the KELM can be calculated by equation (7).

$$f_{\text{K-ELM}}(\mathbf{x}) = \text{h(x)}\mathbf{H}^T(\frac{\mathbf{I}}{C} + \boldsymbol{\omega})^{-1}\mathbf{T} \quad (7)$$

Here,

$$\text{h(x)}\mathbf{H}^T = \begin{bmatrix} K(\mathbf{x}, x_1) \\ \vdots \\ K(\mathbf{x}, x_n) \end{bmatrix} \quad (8)$$

Improved artificial bee colony algorithm

There are 3 kinds of bees in the ABC algorithm: employed bees, onlooker bees, and scout bees. The position of the flower source represents a set of matrix D for the optimization problem (Karaboga & Basturk, 2007;). The amount of nectar represents the adaptability level of the optimization problem. Suppose that the number of nectar sources is *AN*, the fitness value of the i-th (i=1,2,...,*AN*) nectar source is f_i, and the maximum number of cycles is *Ki*. The employed bees search for a new nectar source according to equation (9).

$$v_{ij} = u_{ij} + \varphi(u_{ij} - u_{nj}) \quad (9)$$

Here, v_{ij} is a new nectar source; u_{ij} is an existing nectar source; φ is a random value; n is a random number between[1, *AN*], and $n\#i, 1 \leq i \leq AN, 1 \leq j \leq d$; d is the dimension of the matrix D.

When all the employed bees complete the search, they return to the hive and share the nectar sources with the onlooker bees. Then, the onlooker bees select a nectar source among those. Finally, they continue searching according to equation (9) to find the best flower location.

When the number of cycles reaches *limit* times (*limit*<*Ki*), the nectar source is not updated, and the scout bees update the nectar source according to equation (10).

$$u_{ij} = u_{\min,j} + \varphi(u_{\max,j} - u_{\min,j}) \quad (10)$$

In this equation, $u_{\min,j}$ and $u_{\max,j}$ are the upper and lower limits, respectively, of the j element in matrix D.

Since ABC runs in serial mode, the search speed will be slow. For this shortcoming, this research proposes an improved ABC algorithm (IABC). The ABC determines the maximum fitness value $f(i)_{max}$ of the nectar source before the onlooker bees starts searching, then ABC starts searching from this nectar source, where $f(i)_{max}$ is calculated by equation (11) and f is a set of fitness values for all nectar sources.

$$f(i)_{max} = \text{MAX}(f) \quad (11)$$

Optimizing the KELM network based on IABC algorithm

In the KELM network, the regularization coefficient C and the kernel function matrix parameter Y mainly depend on the human's experience, this makes it difficult to determine the optimal combination. Therefore, we use the IABC algorithm to find the optimal values of these two parameters to optimize the KELM network, called the IA-KELM algorithm. Figure 1 is a flow chart of the IA-KELM algorithm. Below we introduce the implementation process of the algorithm.

Algorithm 1: IA-KELM algorithm

1. Input data; initializes the IABC algorithm, where *d=2*, the number of nectar sources is *2*AN*, and the number of employed bees, onlooker bees, and scout bees are all *AN*.

for i=1:*Ki*

2. IABC's employed bees search for nectar sources and calculate the best fitness value through the KELM network.

3. Determine the best fitness value according to equation (11) and find the corresponding nectar source.

4. IABC's onlooker bees search for nectar sources and calculate the best fitness value through the KELM network.

5. If i=*limit* and the nectar source position does not change, the scout bee updates the nectar source position.

end

6. Output: the optimal C and Y.

Results and discussion

The data analysis

According to data about male powerlifters, we draw the relationship between age, body weight and best squat. As shown in Figure 2, we can see that the best results of the powerlifters are achieved between the ages of 20-40. When the age is less than 15 years old, the performance is not satisfactory; from the age of 16 years, the performance increase significantly; after entering adulthood, although the sports performance is improved, it is not significant. This is because the athletes begin to develop more rapidly at the age of 16. Their ability level has entered a stage of rapid development, then the performance of sports has increased sharply, and the level of power has increased substantially. As the age increases (after the age of 40), the performance begins to decline, in line with the laws of nature. For the body weight, since the body weight becomes higher, the sports performance is also improved. This is the relationship between age, body weight and best squat. Below we will build a powerlifting performance analysis model based on the relationship.

Powerlifting performance analysis model

Based on the IA-KELM algorithm, we establish a powerlifting performance analysis model (IA-KELM model). In this section, we use two model performance evaluation indicators, coefficient of determination (R2) and root-mean-square error of prediction (RMSEP) (Esfe et al., 2018). The definitions of these two indicators are expressed by the following equations (12) and (13).

$$R2=1-\frac{\sum_{i=1}^{n_{test}}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n_{test}}(y_i-\bar{y})^2} \tag{12}$$

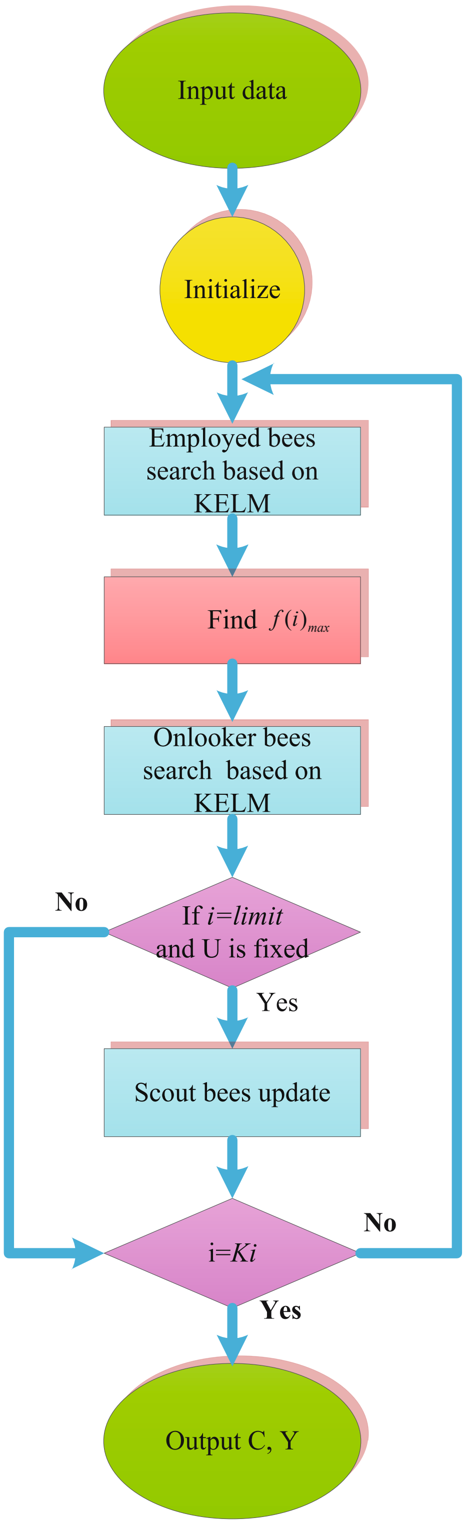


Figure 1 IA-KELM algorithm

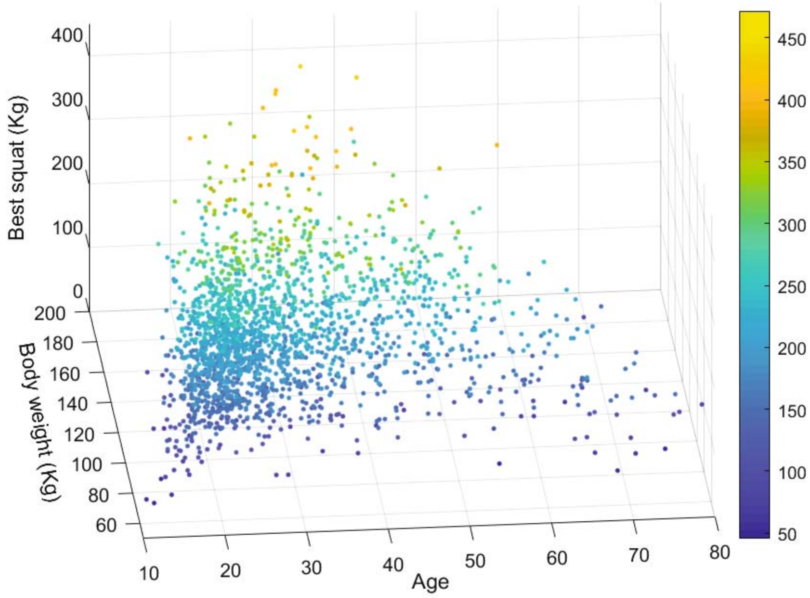
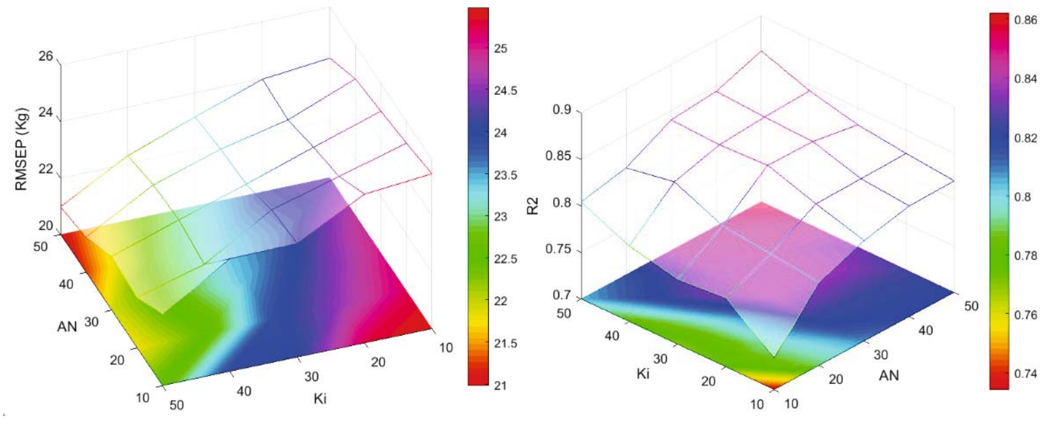


Figure 2 The relationship between age, body weight, and best squat

$$\text{RMSEP}=\sqrt{\frac{\sum_{i=1}^{n_{test}}(y_i-\hat{y}_i)^2}{n_{test}}} \tag{13}$$

In the equations, n_{test} is the number of samples in the test set; y_i is the actual value; $\bar{y}$ is the mean value of the actual values; and $\hat{y}_i$ is the predictive value.

Figure 3 shows the effect of different nectar source numbers *AN* and maximum iteration number *Ki* on the performance of the IA-KELM model. We can see that the RMSEP and R2 values of the model are also affected as the values of *AN* and *Ki* change, the larger the AN or Ki value, the smaller the RMSEP value and the closer R2 is to 1, the better the model performance. Figure 4 shows the results of random selection of 100 powerlifters. It can be seen that the output of the IA-ELM model is close to the actual value, especially in the range of 150-250 Kg, and the model performance is the best. The results show that the model can well predict the performance and level of the weightlifter. Moreover, it also proves the effectiveness of our method.



(a) RMSEP (b) R2

Figure 3 The effect of different AN and Ki values on IA-KELM powerlifting performance analysis model

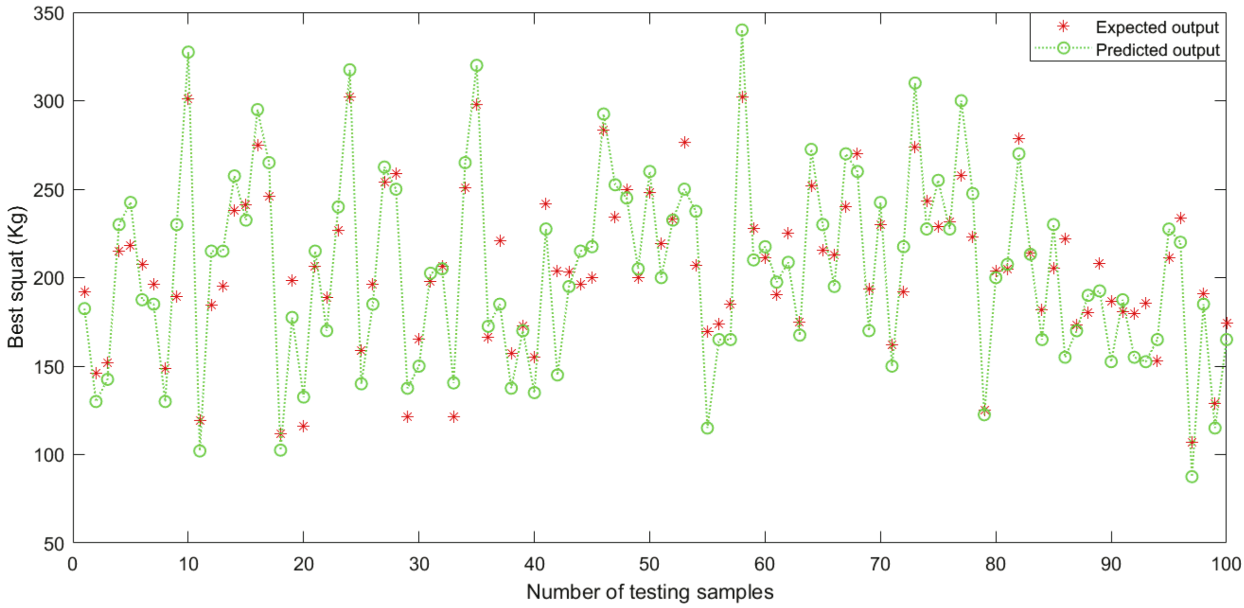


Figure 4 The prediction results of the best performance of powerlifters

Comparison of different methods

Support vector machine (SVM) is a machine learning method based on statistical theory (Cortes, & Vapnik, 1995). SVM can effectively solve small, complex, non-linear data. Therefore, it is widely used in many classification and regression problems (Li et al., 2018; Pai, Chang, & Lin, 2017; Geng, et al., 2016). Wu & Shen (2018) proposed an improved particle swarm optimization algorithm to optimize support vector machine (IPSO-LSSVM algorithm), and the algorithm obtains good prediction results on the regression problem. This study compared the IPSO-LSSVM, KELM, and IA-KELM algorithms. The comparison results are shown in table 1 and figure 5. We can see that the IA-KELM has the best prediction result, the smallest RMSEP value and the largest R2 value, the second is the IPSO-LSSVM model, and the third is the KELM model. The outcomes prove the effectiveness of the IA-KELM algorithm again. The IABC can effectively improve the KELM network, making the performance of the KELM model more stable and predictive.

Table 1 Comparison of different methods

Methods	**RMSEP**	**R2**
IPSO-LSSVM	22.8745	0.8247
KELM	27.4512	0.7114
IA-KELM	20.9811	0.8620

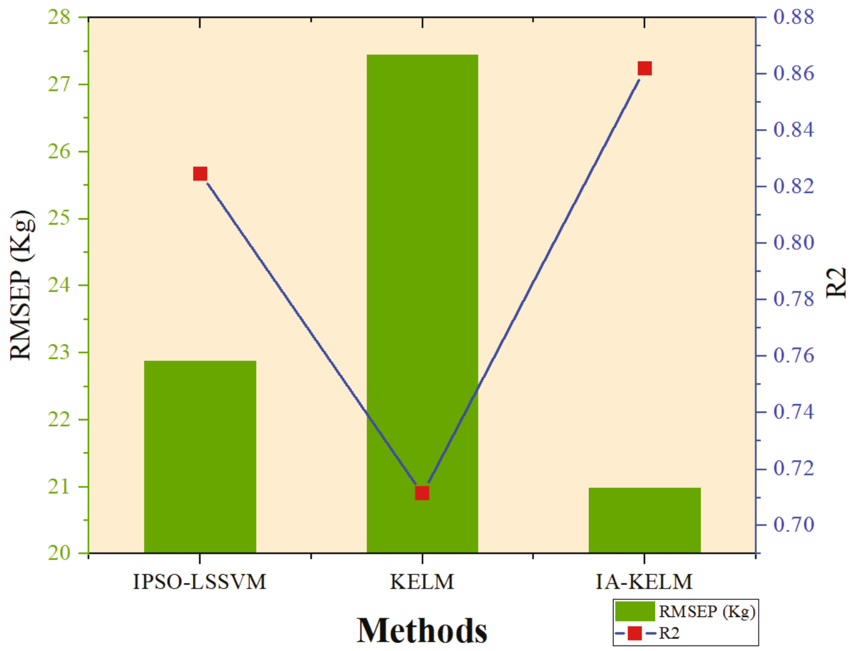


Figure 5 Performance comparisons of different methods

Conclusion

Powerlifters vary in their strength levels and performances at different ages, with height and weight, which has been proven. How to establish the relationship between age, weight and powerlifting is a very interesting study. In recent years, artificial intelligence has achieved groundbreaking development. Therefore, based on the artificial intelligence algorithm, this research proposes an analysis model for predicting the performance of powerlifters. The results show that our method can effectively predict the best performance of powerlifters, and the method provides a promising analytical tool for powerlifting.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Acknowledgements

This work was supported by the Ho Chi Minh City University of Physical Education and Sport, and the Duy Tan University, Viet Nam.

References

Andronikos, G., Elumaro, A. I., Westbury, T., & Martindale, R. J. (2016). Relative age effect: implications for effective practice. Journal of sports sciences, 34(12), 1124-1131.

Anton, M. M., Spirduso, W. W., & Tanaka, H. I. R. O. F. U. M. I. (2004). Age-related declines in anaerobic muscular performance: weightlifting and powerlifting. Medicine and science in sports and exercise, 36(1), 143-147.

Buckingham, G., Byrne, C. M., Paciocco, J., van Eimeren, L., & Goodale, M. A. (2014). Weightlifting exercise and the size–weight illusion. Attention, Perception, & Psychophysics, 76(2), 452-459.

Carey, D. L., Ong, K., Whiteley, R., Crossley, K. M., Crow, J., & Morris, M. E. (2018). Predictive modelling of training loads and injury in Australian football. International Journal of Computer Science in Sport, 17(1), 49-66.

Coker, N. A., Varanoske, A. N., Baker, K. M., Hahs-Vaughn, D. L., & Wells, A. J. (2018). Predictors of competitive success of national-level powerlifters: a multilevel analysis. International Journal of Performance Analysis in Sport, 18(5), 796-805.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

Durguerian, A., Bougard, C., Drogou, C., Sauvet, F., Chennaoui, M., & Filaire, E. (2016). Weight loss, performance and psychological related states in high-level weightlifters. International journal of sports medicine, 37(03), 230-238.

Eime, R. M., Harvey, J. T., Charity, M. J., Casey, M. M., Westerbeek, H., & Payne, W. R. (2016). Age profiles of sport participants. BMC sports science, medicine and rehabilitation, 8(1), 6.

Esfe, M. H., Tatar, A., Ahangar, M. R. H., & Rostamian, H. (2018). A comparison of performance of several artificial intelligence methods for predicting the dynamic viscosity of TiO2/SAE 50 nano-lubricant. Physica E: Low-dimensional Systems and Nanostructures, 96, 85-93.

Geng, Y., Chen, J., Fu, R., Bao, G., & Pahlavan, K. (2016). Enlighten wearable physiological monitoring systems: On-body rf characteristics based human motion classification using a support vector machine. IEEE transactions on mobile computing, 15(3), 656-671.

Huang, G. B., Zhou, H., Ding, X., & Zhang, R. (2012). Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 42(2), 513-529.

Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: theory and applications. Neurocomputing, 70(1-3), 489-501.

James, L. P., Comfort, P., Suchomel, T. J., Kelly, V. G., Beckman, E. M., & Haff, G. G. (2018). The impact of power clean ability and training age on adaptations to weightlifting-style training. Journal of Strength & Conditioning Research. Publish online.

Karaboga, D., & Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. Journal of global optimization, 39(3), 459-471.

Karaboga, D., Gorkemli, B., Ozturk, C., & Karaboga, N. (2014). A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artificial Intelligence Review, 42(1), 21-57.

Li, Y., Zhu, Z., Hou, A., Zhao, Q., Liu, L., & Zhang, L. (2018). Pulmonary nodule recognition based on multiple kernel learning support vector machine-pso. Computational and mathematical methods in medicine, 2018.

Lienhart, R., Einfalt, M., & Zecha, D. (2018). Mining automatically estimated poses from video recordings of top athletes. International Journal of Computer Science in Sport, 17(1), 94-112.

Mao, Y., Le, B. T., Xiao, D., He, D., Liu, C., Jiang, L., ... & Liu, X. (2019). Coal classification method based on visible-infrared spectroscopy and an improved multilayer extreme learning machine. Optics & Laser Technology, 114, 10-15.

Pai, P. F., ChangLiao, L. H., & Lin, K. P. (2017). Analyzing basketball games by a support vector machines with decision tree model. Neural Computing and Applications, 28(12), 4159-4167.

Smith, K. L., Weir, P. L., Till, K., Romann, M., & Cobley, S. (2018). Relative age effects across and within female sport contexts: a systematic review and meta-analysis. Sports Medicine, 48(6), 1451-1478.

Solberg, P. A., Hopkins, W. G., Paulsen, G., & Haugen, T. A. (2019). Peak Age and Performance Progression in World-Class Weightlifting and Powerlifting Athletes. International journal of sports physiology and performance, Publish online.

Sun, L., Chen, T., & Zhang, Q. (2018). An Artificial Bee Colony Algorithm with Random Location Updating. Scientific Programming, 2018.

Wang, H., Li, J., Cui, W., Lu, X., Zhang, Z., Sheng, C., & Liu, Q. (2019). Mobile Robot Indoor Positioning System Based on K-ELM. Journal of Sensors, 2019.

Wang, J., Song, Y. F., & Ma, T. L. (2017). Mexican Hat Wavelet Kernel ELM for Multiclass Classification. Computational intelligence and neuroscience, 2017.

Wu, Y. H., & Shen, H. (2018). Grey-related least squares support vector machine optimization model and its application in predicting natural gas consumption demand. Journal of Computational and Applied Mathematics, 338, 212-220.

Xue, Y., Jiang, J., Zhao, B., & Ma, T. (2018). A self-adaptive artificial bee colony algorithm based on global best for global optimization. Soft Computing, 22 (9), 2935-2952.

Zhao, D., Huang, C., Wei, Y., Yu, F., Wang, M., & Chen, H. (2017). An effective computational model for bankruptcy prediction using kernel extreme learning machine approach. Computational Economics, 49(2), 325-341.