CHEMICAL ENGINEERING TRANSACTIONS
VOL. 46, 2015

A publication of
AIDIC
The Italian Association
of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Peiyu Ren, Yancang Li, Huiping Song

ISBN 978-88-95608-37-2; **ISSN** 2283-9216
DOI: 10.3303/CET1546069

Research on the Middle-Distance Running Achievement Prediction for College Students Based on LSSVMGAS

Xiaobing Yang, Suqiong Feng*

Sichuan Agricultural University
fengtingting0828@163.com

By the combination of the computer technology and the physical training, we can better predict the training achievement. Then we can make the targeted training for the college students. Therefore, the combination of the physical training for the college students and the computer technology has become a trend. Middle-distance running is one of the most important standards for the physical training of the college students. Improving the middle-distance running performance not only improves the physical quality of the college students, but also improves the perseverance of the college students. In this paper, we combine the LSSVM method with the GAS method and propose the improved LSSVMGAS method. In the experimental part, we use the LSSVMGAS method to predict achievement of the middle-distance running for the college students and achieve the better results.

1. Introduction

In June 30, 2015, health and family planning commission issues a report. The report shows that the average height of adult male and female is 167.1cm and 155.8cm in China. It is lower than the average level of Japan. In this report, we can see that the physical quality of Chinese people is not optimistic. In National Standard for Sports Training, the middle-distance running is easy to be neglected of the college students. To middle-distance running not only enhances the physical quality for the college students, but also hones the character and will. Qiuhe Huang, Lanyong Wei and Haifeng Huo (2014) researched the college teachers' teaching ability.
Tang Chifei (2008) studied the running train for the college students. The author thought that the running train for the college students must be in accordance with their own growth and make the most effective use of their energy. Chi Hua (2013) et al. researched the college students' running pain sensation scale. According to the item analysis, reliability and validity, the sensation scale could better reflect the degree and level of the feeling pain of running for the college students. Guan Qingli and Wang Haiyuan (2010) proposed that the long-distance running should take the physical education and the extra-curricular activities as the carrier. And it needed the participation of the teachers and the students. Zhao Fangye (2013) thought that the college physical education must adapt to the physical conditions of the contemporary college students. We should make the reform from the education philosophy, curriculum structure, teaching management link and the assessment indicator. In addition, Lin Zhenglan (2013), Pang Rong (2011) also studied the question.
Support vector machine (SVM) was a data mining method based on the statistical learning theory. Tao Lin (2015) et al. studies the SVM Least squares support vector machine (LSSVM) was a new support vector machine. The operation speed of the LSSVM was faster than other support vector machines. Therefore, it had the wide application. Now, the LSSVM had many applications in the industrial chemical by Mohammad Mesbah et al. (2015), Hossein, Safari et al. (2014) Hamidreza Yarveicy et al. (2014), energy by Xing Yan and Nurul A (2013) and weather by Yi Zhang et al.. Gravitational search algorithm (GSA) was an intelligent optimization algorithm which was proposed in 2009. The thought of GSA algorithm was derived from Newton's law of gravity. It guided the optimization search according to the group intelligence which generated by the gravity among the particles. The gravitational search concept was simple, easy to achieve and needed to adjust a few parameters. Now, Oliveria, Zikang Su, Beatriz Gonzalez and other people (2015) studied the gravitational search algorithm.

Please cite this article as: Yang X.B., Feng S.Q., 2015, Research on the middle-distance running achievement prediction for college students based on lssvmgas, Chemical Engineering Transactions, 46, 409-414 DOI:10.3303/CET1546069

In this paper, we combine the LSSVM with GAS and propose an improved LSSVMGAS method. Then, we use the method to predict the middle-distance running performance for the college students and achieve the good results. The structure of this paper is as follows. The first part is the introduction. In this part, we present the research status of the related content. The second part introduces the LSSVM. The third part is the basic principle of GSA. In the fourth part, we combine the LSSVM with GSA method and propose the improved LSSVMGAS method. The fifth part is the numerical analysis and the last part is the conclusion.

2. The LSSVM method

We suppose the training data set is $(x_i, y_i), i = 1, 2, \cdots, N$. x_i and y_i are the input vector and the corresponding output respectively. The input space R^d is mapped into a feature space Z though the nonlinear function $\varphi(x_i)$.

The linear function of the feature space defines as follows.

$$y = f(x, \omega) = \omega^T \phi(x) + b \tag{1}$$

$\omega \in Z, b \in R$

According to the principle of the risk minimization, we can get the objective function

$$\begin{aligned} &\min_{\omega,\xi} \frac{1}{2}\omega^T\omega + \frac{\gamma}{2}\sum_{i=1}^{N}\xi_i^2 \\ &s.t.\, y_i = \omega^T\varphi(x_i) + b + \xi_i \\ &\quad \gamma > 0 \\ &\quad i = 1, 2, \cdots, N \end{aligned} \tag{2}$$

γ is the penalty factor, ξ_i is an error variable to indicate an error between the true output value of the sample i and its estimated value.

We can construct the corresponding Lagrange function:

$$L(\omega, b, \xi_i, \beta) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\gamma\sum_{i=1}^{N}\xi_i^2 - \sum_{i=1}^{N}\beta_i\left[\omega^T\varphi(x_i) + b + \xi_i - y_i\right] \tag{3}$$

Where β_i is the Lagrange multiplier.

3. Basic principle of GSA

The gravitational search algorithm is a kind of group intelligence optimization algorithm which is proposed by Esmat professor. The magnitude of the gravitational force is proportional to the mass of the two particles. And it is inversely proportional to the distance of the two particles.

$$F = G\frac{M_1 M_2}{R^2} \tag{4}$$

The gravitational force F_{ij} between the particle that the quality is i and the particle that the quality is j is expressed by the following function.

$$F_{ij} = G\frac{M_{aj} \times M_{pi}}{R^2} \tag{5}$$

$$a_i = \frac{F_{ij}}{M_{ii}} \tag{6}$$

M_{aj} is the active gravitational mass of the particle j . M_{pi} is the passive gravitational mass of the particle i . M_{ii} is the inertia mass of the particle i .

In the search space of D dimension, we assume that there are N particles. We define the location of the i particle as $X_i = (x_i^1, \cdots, x_i^d, \cdots, x_i^n)$ and $i = 1, 2, \cdots, N$. Where, N is the population quantity. x_i^d is the location of the i particle in d dimension. Now, we assume that the gravity and the mass of the particle are determined by the target function value of the search space. We calculate the inertia mass of the particle according to the formula (7) and (8). and we get the normalized mass according to the formula (9).

$$M_{ai} = M_{pi} = M_{ii} = M_i, i = 1, 2, \cdots, N \tag{7}$$

$$m_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \tag{8}$$

$$M_i(t) = \frac{m_i(t)}{\sum_{j=1}^{N} m_j(t)} \tag{9}$$

Where, $fit_t(t)$ is the fitness degree value of the i particle at t time.

Then, we can calculate the gravity of each particle in each dimension according to the following formula. The gravity of the i particle and the j particle in d dimension is,

$$F_{ij}^d = G(t)\frac{M_i(t) \times M_j(t)}{R_{ij}(t)+\varepsilon}(x_j^d(t) - x_i^d(t)) \tag{10}$$

$x_j^d(t)$ represents the location of the j particle in d dimension. $x_i^d(t)$ represents the location of the i particle in d dimension. ε is a very small constant which is to prevent the denominator is zero.

In the gravitational search algorithm, in order to increase the random properties of the algorithm, we think that the force of the particle i is equal to the sum of other particles. The value is defined as follows.

$$F_i^d(t) = \sum_{j=1, j \neq i}^{N} rand_j F_{ij}^d(t) \tag{11}$$

Where, $rand_j$ is a random number which belongs to $[0,1]$.

According to Newton's second theorem, the acceleration of the particle i in d dimension at t time is defined as follows.

$$a_i^d(t) = \frac{F_i^d(t)}{M_i(t)} \tag{12}$$

Where, $M_i(t)$ is the inertia mass of the particle i at t time.

We update the speed and the location according to the following formula.

$$v_i^d(t+1) = rand_i \times v_i^d(t) + a_i^d(t) \tag{13}$$

$$x_i^d(t+1) = x_i^d(t) + x_i^d(t+1) \tag{14}$$

We assume that the number of the particles is $Nbest$. And the $Nbest$ changes with the time. Therefore, $Nbest$ decreases from the initial N_0. In the initial time, the particles have forces. As the time goes on, some particles that the masses are bigger begin to force other particles that the masses are smaller. To the end, there is only particle forcing other particles. Therefore, the formula can be changed as follows.

$$F_i^d(t) = \sum_{j \in Kbest, j \neq i}^{N} rand_j F_{ij}^d(t) \tag{15}$$

Where, $Kbest$ is the set of the particles that the inertia masses are bigger.

4. LSSVMGSA

This paper establishes the predicted model of LSSVM and GSA. Firstly, we improve the GSA algorithm. In addition, GSA will optimize the parameter of the LSSVM model. The linear combination of the intermediate nodes is the output and each intermediate node corresponding to a support vector.

In GSA algorithm, the expected particle is guided by the resultant. And it moves to the particle that the quality is heavy. At the same time, during the moving of the expected particle, it is also influenced by other particles. Therefore, when we calculate the resultant, there may be some deviations. Therefore, we introduce the correction factor to modify the resultant force in the algorithm.

$$x = |M(i) - M(j)| \tag{16}$$

$$y(x) = \begin{cases} \frac{\pi}{2} & x \geq C \\ k_2 x & 0 \leq x < C \end{cases} \tag{17}$$

$$\lambda = k_1 \cdot \cos(y(x))$$

Where, λ is the correction factor that the force needs to add. k_1 and k_2 are the adjustable parameters. And they can be adjusted according to the different locations. C is a coefficient which is related to the k_2.

$$C = \frac{\pi k_2^{-1}}{2} \tag{18}$$

$y(x)$ is a function. And the new formula is as follows.

$$F_i^d(t) = \sum_{j=1, j\neq i}^{N} \lambda \cdot F_{ij}^d(t) = \sum_{j=1, j\neq i}^{N} k_1 \cdot \cos(k_2 \cdot (|M(i) - M(j)|)) \cdot F_{ij}^d(t) \tag{19}$$

The fitness function is defined as:

$$fitness = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_p)^2} \tag{20}$$

N is the number of the training samples. y_i is the true value, y_p is the fitted value.

The flow chart of LSSVMGAS is as follows.

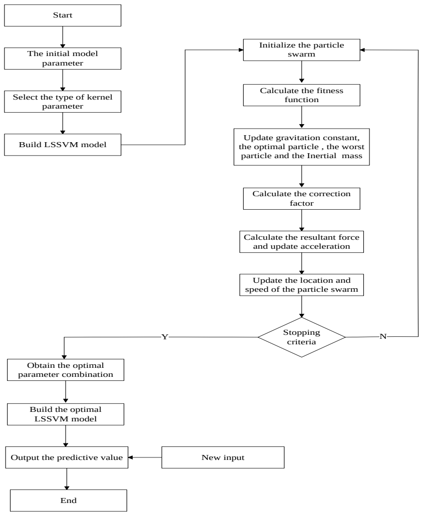

Figure 1: Flow chart of LSSVMGAS

5. Numerical analysis

In the numerical experiment, we predict the long-distance running performance for the college students. We choose 4 college students to run and predict their performances. There are two male college students and two female college students. The male college students run 1000 meters and the female college students run 800 meters. Firstly, we collect data of 30 groups' long-distance running. The first 20 groups are as the training sets and the last 10 groups are as the predicted sets. The data of the training sets is as follows.

Table 1: The data of the training set

No	Male1	Male2	Female1	Female2
1	224	210	214	225
2	226	209	218	226
3	218	203	211	217
4	226	208	208	220
5	230	205	205	221
6	217	214	210	214
7	214	212	213	218
8	210	210	211	219
9	215	208	216	225
10	223	213	213	227
11	225	211	218	225
12	228	214	223	224
13	219	207	220	221
14	220	194	213	218
15	223	205	211	217
16	216	203	206	219
17	224	206	214	223
18	225	218	215	224
19	233	220	218	221
20	228	216	220	227

Then, we predict the performance. The predicted performance and the actual performance are as follows.

Table 2: The actual values and the predicted values

No	Male1		Male2		Female1		Female2	
	Actual values	Predicted values	Actual values	Predicted values	Actual values	Predicted values	Actual values	Predicted values
21	225	227	215	213	217	219	224	223
22	223	224	217	215	213	215	224	224
23	220	220	208	210	214	214	221	223
24	218	220	210	211	216	215	218	220
25	221	218	213	211	218	217	220	221
26	224	222	214	214	217	215	218	219
27	223	224	212	213	213	214	223	222
28	225	224	208	209	211	212	226	225
29	224	223	207	209	208	210	227	226
30	224	224	210	211	207	208	225	224

We compare the actual performance with the predicted performance. The results are as follows.

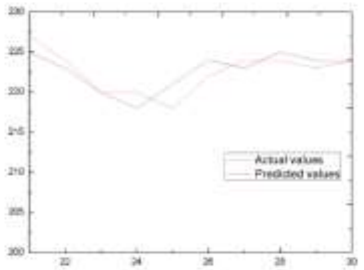

Figure 2: The experiment results of male1

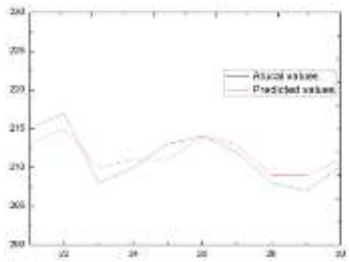

Figure 3: The experiment results of male2

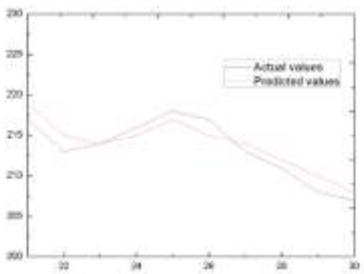

Figure 4: The experiment results of female1

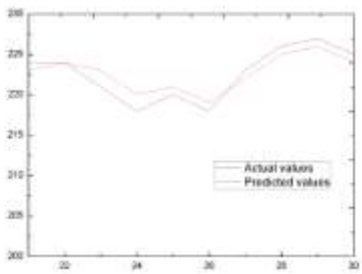

Figure 5: The experiment results of female2

From the above figure, we can see that the predicted value which is obtained by the LSSVMGAS method is very close to the actual values. The two curves of each training set are very similar. From the experiment, we can know that the LSSVMGAS method achieves the good results. The experiment shows that the LSSVMGAS method is feasible and effective.

7. Conclusions

The middle-distance running is popular by the teachers and the students. And the majorly of teachers pay more and more attention to the middle-distance running with the computer technology. In this paper, we propose the LSSVMGAS method to predict the middle-distance running performance and the targeted training. The main work of this paper is as follows. Firstly, we introduce the research status of the related content. Secondly, we introduce the LSSVM method and GAS method. Thirdly, we combine the LSSVM method with the GAS method and propose the improved LSSVMGAS method. Fourthly, we use the LSSVMGAS method to predict the middle-distance running performance for the college students. And it achieves the good experimental results. The content of this part has some reference value for the research of the LSSVM and the physical quality of the college students.

Acknowledgment

This research is supported by Undergraduate Theses Incubation Program of Sichuan Agricultural University.

References

Chi H., Fang X.C., Lu J.B., Du L., 2013, The establishment and study of college students' running pain sensation scale, Journal of Shenyang Sport University, 2: 141-142.
De Moura Oliveira P.B., Solteiro Pires E.J., Novais P., 2015, Design of Posicast PID control systems using a gravitational search algorithm, Neurocomputing, 167: 18-23.
Gorjaei R.G., Songolzadeh R., Torkaman M., Safari M., Zargar G., 2015, A novel PSO-LSSVM model for predicting liquid rate of two phase flow through wellhead chokes, Journal of Natural Gas Science and Engineering, 24: 228-237.
Gouthamkumar N., Sharma V., Naresh R., 2015, Disruption based gravitational search algorithm for short term hydrothermal scheduling, Expert Systems with Applications, 42(20): 7000-7011.
Guan Q.L., Wang H.Y., 2010, Exploration of running activities in public security colleges, Journal of physical education institute of Shanxi normal university, 8: 83-87.
Huang Q.H., Wei L.Y., Huo H.F., 2014, Research on independent college teachers’ teaching ability based on factor analysis in SPSS, Mathematical Modelling and Engineering Problems, 1(1): 25-28.
Lin T., Wu P., Gao F.M., Yu Y., 2015, Study on SVM temperature compensation of liquid ammonia volumetric flowmeter based on variable weight PSO, International Journal of Heat and Technology, 33(2): 151-156.
Lin Z.L., Deng N.C., Yang Y.H., 2003, Study on the division of year's training of middle- distance run player of college student, Journal of Nanjing Institute of Physical Education, 8: 103-105.
Mesbah M., Soroush E., Azari V., Lee M.Y., Bahadori A., Habibnia S., 2015, Vapor liquid equilibrium prediction of carbon dioxide and hydrocarbon systems using LSSVM algorithm, The Journal of Supercritical Fluids, 97: 256-267.
Pang R., 2011, Discussion on cultivating students' interest in middle distance running, China Education Innovation Herald, 32: 212-215.
Safari H., Shokrollahi A., Jamialahmadi M., Ghazanfari M.H., Bahadori A., Zendehboudi S., 2014, Prediction of the aqueous solubility of BaSO4 using pitzer ion interaction model and LSSVM algorithm, Fluid Phase Equilibria, 374: 48-62.
Shayeghi H., Ghasemi A., 2013, Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme, Energy Conversion and Management, 74: 482-491.
Su Z.K., Wang H.L., 2015, A novel robust hybrid gravitational search algorithm for reusable launch vehicle approach and landing trajectory optimization, Neurocomputing, 162: 116-127.
Tang C.F., 2008, Discussion on the comprehensive quality training of College Students' running, Scientific and technological information, 12: 221-223.
Yan X., Chowdhury N.A., 2013, Mid-term electricity market clearing price forecasting: A hybrid LSSVM and ARMAX approach, International Journal of Electrical Power & Energy Systems, 53: 20-26.
Yarveicy H., Moghaddam A.K., Ghiasi M.M., 2014, Practical use of statistical learning theory for modeling freezing point depression of electrolyte solutions: LSSVM model, Journal of Natural Gas Science and Engineering, 20, September, 414-421
Zhang Y., Li H., Wang Z.H., Zhang W.B., Li J., 2015, A preliminary study on time series forecast of fair-weather atmospheric electric field with WT-LSSVM method, Journal of Electrostatics, 75, 85-89.
Zhao F.Y., 2013, Research on reform of college physical education curriculum-the problems of “The College Students ' Physical Fitness Test”. Bulletin of Sport Science & Technology, 8: 60-63.