Prediction method of silicon content in blast furnace hot metal based on IPSO-HKELM

Kaifei Hu¹, Qinghe Hu¹,⁴, Chongmin Liu¹ and Shuang Zhang²
¹ Information Science & Engineering School, Northeastern University, Shenyang 110819, China
² School of Software, Northeastern University, Shenyang 110819, China
⁴ Corresponding author: huqinghe@ise.neu.edu.cn

Abstract. Aiming at the problem that the silicon content of molten iron can not be detected online, a model for predicting silicon content in molten iron based on Hybrid Kernel Extreme Learning Machine optimized by Improved Particle Swarm Optimization Algorithm (IPSO-HKELM) is proposed. Firstly, the input variables are reduced by PCA, and then the prediction model of molten iron content based on HKELM is established. In this paper, PSO is used to optimize the kernel parameters of HKELM. Aiming at the problem that PSO is easy to fall into local optimum, the Inertia weight reduced with the number of iterations and the random back-based learning mutation operation are introduced, so that PSO can jump out of the local minimum point more easily and get the optimal result. Experiments show that the prediction model of silicon-based silicon content based on IPSO-HKELM has high prediction accuracy and short time, which can meet the actual production needs.

1. Introduction
Blast furnace ironmaking is an important part of steelmaking production. At present, there is still a significant gap between China's blast furnace ironmaking automation level and foreign countries. In the iron making process, many ironmaking operations rely mostly on artificial experience. When the raw material batch or place of production is different, the quality parameter of molten iron fluctuates greatly. Due to the limited measurement methods, blast furnace operators generally use manual periodic sampling and testing methods to detect the quality parameters of molten iron, resulting in serious lag in the quality information of molten iron, thus affecting the judgment of furnace conditions and further affecting the operation of blast furnaces. Since the silicon content is the main indicator of the quality parameter of molten iron, it can indirectly reflect the furnace temperature of the blast furnace. Therefore, there is a need for a fast and accurate method of predicting the silicon content of molten iron [1-2].

At present, there are many researches on the modeling of silicon content in molten iron, including rule-based reasoning models [3], mechanism-based mathematical models [4] and data-based statistical models [5]. In recent years, with the rapid development of computers, data-driven modeling methods have been widely used to solve the problem of hot metal quality. The data-driven modeling method is a black box modeling method. Its main idea is to use some mathematical tools and intelligent algorithms to directly approximate the input-output relationship of the process based on the process data without any prior knowledge [6]. Existing methods for modeling silicon content in molten iron include artificial neural network [7] (ANN), partial least squares [8], support vector regression[9]
The extreme learning machine [10] (ELM) is a single hidden layer feedforward neural network. Compared with other neural network algorithms, it has the advantages of fast learning speed and good generalization performance. However, the ELM model is unstable due to the ELM randomly selecting input weights and deviation vectors. In order to solve this problem, the literature proposed the KELM model [11], use the kernel function instead of the display map of the hidden layer node. Based on the above research, this paper combines the polynomial kernel function with the Gaussian kernel function to construct a hybrid kernel extreme learning machine (HKELM) model for the prediction of silicon content in blast furnace hot metal. The improved particle swarm optimization algorithm (IPSO) is used to optimize the parameters of the kernel function. In the traditional PSO, the inertia weight with the reduced evolution time and the random-based reverse learning mutation operation are added, which makes the PSO easier to jump out of the local optimum. Principal component analysis (PCA) [12] is used to reduce the dimensionality of the input variables before building the predictive model, which simplifies the model structure.

2. Description of blast furnace ironmaking system
The typical blast furnace ironmaking process is shown in Figure 1. The entire blast furnace ironmaking system is divided into several subsystems: blast furnace body, feed system, hot air system, pulverized coal injection system, blast furnace gas treatment system and tapping system. At the time of production, the treated iron ore raw materials, coke and solvent are fed into the furnace from the top of the blast furnace according to a certain ratio. Then, the oxygen-enriched and hot air heated by the hot blast stove is blown into the blast furnace from the tuyere at the bottom of the blast furnace, and mixed with the fuel such as pulverized coal and oil. The blast furnace raw material slowly descends from top to bottom by gravity, and a series of complicated physical reactions and chemical reactions occur with the rising gas flow, and finally the liquid molten iron is discharged from the tap hole. In iron ore, unreduced impurities are combined with a solvent such as limestone to form slag. The slag floats on the surface of the molten iron and is periodically discharged from the slag port from a higher position in the blast furnace hearth. The gas discharged from the top of the furnace can be used as fuel recycling for hot blast stoves, boilers, etc. after being treated by gravity dust removal and TRT [13-14].

Figure 1 Schematic diagram of blast furnace ironmaking process

3. Algorithm principle and analysis
3.1. Hybrid Kernel Extreme Learning Machine
HKELM borrows the idea of support vector machine and uses hybrid kernel function to replace the feature map of ELM hidden layer nodes. Compared with traditional ELM, HKELM does not need to artificially determine the number of hidden layer nodes, only need to select appropriate kernel
parameters, you can get the output weight [15].

Given \( N \) training sample data sets \((x_i, t_i)\), \( x_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T \in \mathbb{R}^n \) is the input data of the sample, \( t_i = [t_{i1}, t_{i2}, \cdots, t_{im}]^T \in \mathbb{R}^m \) is the sample output value. For SLFNs with \( L \) number of hidden layer nodes, if the excitation function is \( g(x) \), then the output of the network is:

\[
y_j = \sum_{i=1}^{L} \beta_i g_i(\omega_i \cdot x_j + b_i), \quad j = 1, 2, \cdots, N
\]  

Where, \( \beta_i \) is the weight vector between the \( i \)-th hidden layer node and the output layer node, \( \omega_i \) is the weight vector between the \( i \)-th hidden layer node and the input layer node, \( b_i \) is the bias of the \( i \)-th hidden layer, the output value of the network is \( y_i \). When the activation function is able to approximate any \( N \) samples with zero error, \( \sum_{i=1}^{N} \| y_i - t_i \| = 0 \), at the same time,

\[
t_j = \sum_{i=1}^{L} \beta_i g_i(\omega_i \cdot x_j + b_i), \quad j = 1, 2, \cdots, N
\]  

Therefore, the above \( N \) equations can be expressed as: \( H\beta = T \), where,

\[
H = \begin{bmatrix}
g_i(\omega_1 \cdot x_j + b_1) & \cdots & g_i(\omega_L \cdot x_j + b_L) \\
\vdots & \ddots & \vdots \\
g_i(\omega_1 \cdot x_N + b_1) & \cdots & g_i(\omega_L \cdot x_N + b_L)
\end{bmatrix}
\]  

\( H \) is the hidden layer output matrix, \( T \) is the desired output vector. Determining the output weight of the network by least squares method is:

\[
\beta = H^+ T = H^T(I/C + HH^T)^{-1}T
\]  

Where, \( H^+ \) is the Moore-Penrose generalized inverse matrix of the hidden layer output matrix, \( C \) is a penalty factor, \( I \) is the unit diagonal matrix. Adding \( I/C \) to matrix \( HH^T \) can make its characteristic root deviate from zero, it can improve the stability and generalization of results.

For the HKELM algorithm, the expression of the output function is as follows:

\[
f(x) = h(x)H^T(I/C + HH^T)^{-1}T
\]  

The calculation formula for defining the kernel function is as follows:

\[
\Omega_{ELM} = HH^T, \quad \Omega_{i,j} = h(x_i)h(x_j) = K(x_i, x_j)
\]  

Combining the polynomial kernel function with the radial basis kernel function to obtain a mixed kernel function:

\[
K(x_i, x_j) = \lambda \cdot (m \cdot (x_i, x_j) + n)^d + (1-\lambda) \cdot \exp(-\gamma \| x_i - x_j \|^2)
\]  

So HKELM's output function expression is:

\[
f(x) = \begin{bmatrix}
K(x, x_1) \\
\vdots \\
K(x, x_N)
\end{bmatrix} (I/C + \Omega_{ELM})^{-1}T
\]  

Where, \( \gamma, \theta, \lambda \) represent the kernel function parameter. In the HKELM algorithm, it is not necessary to know the feature mapping function of the hidden layer node. As long as the specific form of the kernel function is known, the value of the output function can be found. Compared with ELM,
HKELM does not need to set the number of hidden layer nodes, nor does it need to randomly set input weights and bias vectors, so it has better function approximation ability and classification ability.

3.2. Improvement of particle swarm optimization algorithm

In view of the large number of parameters of HKELM model, this paper proposes particle swarm optimization for parameter optimization. In order to improve the optimization ability of particle swarm optimization algorithm, the particle swarm optimization algorithm is easy to fall into the local minimum problem. Therefore, it is necessary to improve the traditional particle swarm optimization algorithm[16]. Due to the traditional PSO algorithm, the inertia weight $\omega$ represents the ability to inherit the previous particle velocity. When $\omega$ is large, it is beneficial to the global search ability of the algorithm. When $\omega$ is small, it is beneficial to the local search of the algorithm. $\omega$ is too large or too small. It is easy to cause the PSO algorithm to fall into local optimum or not to find the optimal solution. Therefore, the inertia weight with decreasing number of iterations is added, and the inertia weight is as shown in equation (11):

$$\omega(k) = \omega_i - (\omega_i - \omega_e)(k/T_{max})^2$$ \hspace{1cm} (9)

Where, $\omega_i$ is the initial inertia weight, $\omega_e$ is the inertia weight when iterating to the maximum number of times, $k$ is the current number of iterations, $T_{max}$ is the maximum number of iterations, $\omega$ decreases with the increase of the number of iterations, which can ensure the strong global search ability of the particle swarm algorithm in the early stage, and ensure the local search ability of the particles in the later iteration. Finally, a random-based reverse learning mutation operation is introduced in the particle swarm optimization algorithm. The probability of mutation is as shown in equation (12):

$$P = (4 \cdot P_s/T_{max}^2) \cdot k^2 + (4 \cdot P_z/T_{max}) \cdot k + P_e$$ \hspace{1cm} (10)

Where, $P$ is the probability of mutation, and when $P > 0.3, P = 0.5$. $P_s$ and $P_e$ are constants. In the experiment, taking $P_s = 0.35, P_e = 0.1$. The mutation probability increases first and then decreases with the number of iterations, so that the excellent particles are retained to some extent while enhancing the particle diversity. The variation formula is a stochastic-based reverse learning formula, as shown in equation (13):

$$X = X_{min} + \text{rand} \cdot (X_{max} - z)$$ \hspace{1cm} (11)

Where, $z$ is the global optimal particle.

4. Prediction model of silicon content in molten iron based on IPSO-HKELM

The operating environment of blast furnace ironmaking is complex. We need to find a way to quickly detect the silicon content of molten iron, react to the complex operating state inside the blast furnace in time, and then make corresponding operations to ensure the smooth operation of blast furnace ironmaking. The HKELM model does not need to set the number of nodes in the hidden layer, and the training time is short, the prediction accuracy is high, and the generalization performance is good. Since the introduction of nuclear parameters by HKELM leads to very sensitive changes to parameters, an IPSO is used to optimize the parameters of the kernel function. The specific steps of the prediction model of silicon content in blast furnace hot metal based on IPSO-HKELM are as follows:

1. Perform PCA on the sample data, save the result set for backup.
2. Set $m, n, \beta, \lambda$ to particle($d = 2$) and then randomly initialize the position and velocity of the particle.
3. Calculate the output of the training set hidden layer node according to equation (1), and add the L2 regular term.
4. Calculate the hidden layer output weight matrix.
5. Calculate the predicted value of the test set based on the output of the hidden layer node.
(6) Calculate individual extremum and population extremum by using the mean square error (MSE) as the fitness of the particle.

(7) Update particle position and speed based on update formula.

(8) Perform mutation operations on particles based on the mutation probability formula.

(9) Calculate the fitness of the updated particles, update the individual extremum and the population extremum. If the maximum number of iterations is met, return to step (7), otherwise continue (10).

(10) Preserving the particles corresponding to the optimal fitness of the population, that is, the optimal mixed kernel parameters, substituting the parameters into the HKELM model and testing with the test set.

Due to the more parameters of HKELM, however, the traditional particle swarm optimization algorithm optimizes HKELM that can not find the prediction model under the optimal parameters, adding the inertia weight with the decreasing number of iterations and the variation formula based on the reverse learning, which improves PSO in the optimization ability and convergence speed. Compared with the display mapping of the traditional ELM activation function, the hybrid kernel function takes into account the locality and globality of the data under the superior parameters, and has both strong learning ability and good generalization ability. Therefore, finding the optimal kernel function parameters through IPSO and obtaining the HKELM model under the optimal nuclear parameters can theoretically be used as a predictive model for the silicon content of blast furnace ironmaking iron.

5. Prediction and analysis of silicon content in molten iron

19 variables affecting the silicon content of molten iron collected and calculated at the No. 6 blast furnace of Bin Gang, including cold air flow, air supply ratio, hot air pressure, top pressure, differential pressure, top pressure air ratio, gas permeability, drag coefficient, actual wind speed, blast kinetic energy, blast humidity, theoretical combustion temperature, standard wind speed, belly gas index, belly gas volume, hot air temperature, oxygen-rich flow rate, oxygen enrichment rate, set coal injection. The above 19 parameters are used as input variables affecting the silicon content of blast furnace molten iron, and the silicon content of molten iron is used as the output variable.

The above 100 sets of data samples were selected, and the first 70 sets of data were used as the training set, and the last 30 sets of data were used as the test set. The test set was predicted by ELM, SVM and KELM models, and the predicted results were compared. Since there are many parameters of the blast furnace body and there is a strong correlation between the variables, if all the variables participate in the modeling of molten silicon content, the computational complexity will increase, which will affect the accuracy and effectiveness of the model prediction. The input variables are dimension reduced using a principal component analysis (PCA) algorithm. The results of principal component analysis are shown in Figure 2. The dimension of the principal component is taken as 10 (accumulated contribution rate 95%) and will be used as input to the model. The input variable is reduced from 19 to 10 dimensions, which greatly simplifies the model structure and saves training time.

![Figure 2. Principal component analysis result](image_url)
In order to compare the performance of the prediction model, this paper uses ELM, SVM, KELM, IPSO-HKELM to establish the prediction model of silicon content in molten iron. The maximum number of iterations of the particle swarm algorithm is 200, the number of population particles is 20, the initial inertia weight $\omega_s$ is 1.2, and the inertia weight $\omega_e$ is 0.4 when iterating to the maximum number of times. The kernel parameters in the hybrid kernel extreme learning machine are optimized by the improved particle swarm optimization algorithm. The performance of the predicted model of molten silicon content in molten iron was evaluated by comparing the root mean square error, prediction time and fitting coefficient of the test set.

Simulate with MATLAB2016 environment. Figure 3 shows the simulation results. The simulation results are shown in Figure 3. The comparison results of the four models of ELM, SVM, KELM and IPSO-HKELM for the test samples of molten iron silicon content are given. Table 1 compares the four models from three aspects: root mean square error, prediction time and fitting coefficient. The specific analysis results are as follows.

![Figure 3. Comparison of silicon content in molten iron](image)

Table 1. Comparison of performance of molten iron content prediction model

| Model type   | RMSE   | R      | Time consuming(s) |
|--------------|--------|--------|-------------------|
| SVM          | 0.0380 | 0.8045 | 0.0146            |
| ELM          | 0.0318 | 0.8676 | 0.0052            |
| KELM         | 0.0310 | 0.9339 | 0.0186            |
| IPSO-HKELM   | 0.0298 | 0.9738 | 0.0315            |

It can be found from Table 1. that the improved model has a longer running time and higher complexity of the algorithm, but considering its prediction accuracy is more obvious, the prediction result is more stable, and the silicon content of molten iron can be quickly predicted in practical applications. Therefore, this paper proposes that the IPSO-HKELM prediction model is feasible in predicting the silicon content of molten iron, and the prediction accuracy is very high.

6. Conclusion

This paper proposes a prediction model of silicon content in molten iron based on IPSO-HKELM algorithm. Firstly, in order to improve the modeling efficiency and reduce the computational complexity, the principal component analysis method is used to extract the 10 key controllable variables with the strongest correlation with the silicon content of molten iron from many related variables as the input variables of modeling. Then based on the actual industrial field data after processing, the IPSO-HKELM algorithm is used to establish a data-driven prediction model for molten silicon content of molten iron. The performance of this model is better than the traditional ELM, SVM
and KELM prediction models with high precision. Industrial experiments show that the established model can not only accurately estimate the silicon content of molten iron, but also can be used to control the quality parameters of molten iron.

References
[1] UEDA S, NATSUI S, NOGAMI H, et al. Recent progress and future perspective on mathematical modeling of blast furnace [J]. ISIJ International, 2010, 50(7): 914-923.
[2] JIAN L, GAO C H. Binary coding SVMs for the multiclass problem of blast furnace system [J]. IEEE Transactions on Industrial Electronics, 2013, 60(9): 3846-3856.
[3] Zarandi M H F, Ahmadian P. Fuzzy agent-based expert system for steel making process. Expert Systems with Applications, 2009, 36(5): 9539-9547.
[4] Das S K, Kumari A, Bandopadhay D, Akbar S A, Mondal G K. A mathematical model to characterize effects of liquid hold-up on bosh silicon transport in the dripping zone of a blast furnace. Applied Mathematical Modeling, 2011, 35(9): 4208-4221.
[5] Chao Y C, Su C W, Huang H P. The adaptive autoregressive models for the system dynamics and prediction of blast furnace. Chemical Engineering Communications, 1986, 44(1-6): 309-330.
[6] Zhou P, Lu S W, Chai T Y. Data-driven soft-sensor modeling for product quality estimation using case-based reasoning and fuzzy-similarity rough sets. IEEE Transactions on Automation Science and Engineering, 2014, 11(4):992-1003.
[7] Zhou P, Yuan M, Wang H, Wang Z, Chai T Y. Multivariate dynamic modeling for molten iron quality using on-line sequential random vector functional-link networks with self-feedback connections. Information Sciences, 2015, 325: 237-255.
[8] Shi L, Li Z L, Yu T, Li J P. Model of hot metal silicon content in blast furnace based on principal component analysis and partial least square. Journal of Iron and Steel Research, International, 2011, 18(10): 13-16.
[9] Tang X L, Zhuang L, Jiang C J. Prediction of silicon content in hot metal using support vector regression based on chaos particle swarm optimization. Expert Systems with Applications, 2009, 36(9): 11853-11857.
[10] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks [J]. International Joint Conference on Neural Networks, 2004, 2: 985-990 vol.2.
[11] Huang G B, Zhou H M, Ding X J, et al. Extreme learning machines for regression and multiclass classification. IEEE Transaction on Systems Man & Cybernetics Part B Cybernetics, vol.42, no.2, pp.513-529, Apr.2012.
[12] Zhou Ping, Zhang Li, Li Wen-Peng, Dai Peng, Chai Tian-You. Autoencoder and PCA based RVFLNs modeling for multivariate molten iron quality in blast furnace ironmaking. Acta Automatica Sinica, to be published.
[13] Song He-Da, Zhou Ping, Wang Hong, Chai Tian-You. Non-linear subspace modeling of multivariate molten iron quality blast furnace ironmaking and its application. Acta Automatica Sinica, 2016, 42(21): 1664-1679.
[14] Gao C H, Jian L, Luo S H. Modeling of the thermal state change of blast furnace hearth with support vector ma-chines. IEEE Trans. Industrial Electronics, 2011, 59(2):1134-1145.
[15] Liu X W, Wang L, Huang G B, et al. Multiple kernel extreme learning machine[J]. Neurocomputing, 2015, 149 (PA): 253-264.
[16] Zhang Yong, Gong Dun-Wei, Zhang Wan-Qiu. A simplex method based improved particle swarm optimization and analysis on its global convergence. Acta Automatica Sinica, 2009, 35(3): 289–298.