Silicon content prediction of hot metal in blast furnace based on attention mechanism and CNN-IndRNN model

WANG Gao-peng^{1,1*}, YAN Zhen-yu¹, Zhai Hai-peng¹, Zheng Rui-ji¹

¹School of software engineer, Chongqing university of post and communication, Chongqing 400065

^{1*}Correspondence author of this paper

Abstract. The stability of blast furnace temperature is an important condition to ensure the efficient production of hot metal. Accurate prediction of silicon content in hot metal is of great significance to the control of blast furnace temperature in iron and steel plants. At present, the accuracy of most silicon prediction models can only be guaranteed when the furnace condition is stable. However, due to many factors affecting the silicon content in hot metal of blast furnace, and there are large noises, large delays and large fluctuations in the data, the previous prediction results are of limited guiding significance to the actual production. In this paper, combined with the actual situation, the convolution neural network is used to extract the furnace condition characteristics, and then combined with the attention mechanism and the IndRNN model to get the prediction results, so that the prediction can better adapt to the fluctuating data set. The experimental results show that the prediction error of this model is lower than that of other models, which provides a new solution for the research of silicon content in hot metal of blast furnace.

1 Introduction

In iron and steel production, the blast furnace is generally selected as the reaction vessel to produce molten iron. When the blast furnace system works, complex physical and chemical reactions occur in different areas from the top to the bottom of the furnace, which has the characteristics of high temperature, high pressure, multiphase coupling and coexistence of multiple physical forms^[1]. When the temperature in the furnace is too high, the reaction in the furnace will be unbalanced and the materials will be wasted. At the same time, the operation time will be longer and the production efficiency will be reduced. When the temperature in the furnace is too low, the normal steelmaking of the blast furnace will be impossible. Therefore, the key to maintain high-efficiency hot metal production is to ensure the stability of blast furnace temperature. Due to the linear relationship between furnace temperature and silicon content in furnace, silicon content in molten iron has been regarded as the main index of blast furnace temperature prediction^[2,3]. However, in hot metal production, silicon content is difficult to monitor in real time. More and more silicon content prediction models are applied in hot metal production of blast furnace.

At present, the main silicon content prediction models are divided into mechanism model and data-driven model^[4]. The mechanism model is mainly based on the physical and chemical knowledge to analyze the internal reaction of blast furnace production. However, some parameters are difficult to obtain, so the overall effect is

not good. Therefore, in recent years, data-driven models are more and more used in silicon content prediction. For the data-driven model, according to experts' experience, in order to take timely and appropriate control measures, the predicted silicon content in hot metal is generally required to be three hours ahead of the chemical analysis value. Under the three hour delay, the accuracy of the model will be greatly reduced. Wang et al. Used the method of combining PCI with the least square method, Sun Jie et al.^[5] used the method of combining genetic algorithm with extreme learning machine. These two methods can achieve good results when the temperature of blast furnace is stable, but the effect is not ideal in the case of fluctuation; Jian et al.^[6] used SVM to predict silicon content, and the result has large deviation, Ye Fei^[7] With Yang et al.^[8], BP neural network was applied to silicon content prediction, while Li Ze-long^[9] proposed LSTM to predict silicon content, and both methods achieved good results.

With the enhancement of computer computing ability, deep learning has become one of the most popular machine learning methods. In the past, the common defects of furnace temperature prediction were poor performance during large fluctuations and large delays. In response to this problem, this paper proposes a blast furnace hot metal silicon content prediction model that combines convolutional network, attention mechanism, and cyclic neural network (Attention-CNN-IndRNN, Attention-CNN-IndRNN), the model combines the respective advantages of CNN and RNN models, uses convolutional layers to extract the internal state information of the blast furnace implicit in the data, uses the attention mechanism to identify data noise, and the

Corresponding author's e-mail: wanggp@cqupt.edu.cn

IndRNN model extracts time implied laws in dimensions. After experimental comparison, this model shows better prediction performance.

2 Convolution is used to extract the internal state of blast furnace

The reaction process of blast furnace is very complex, and it is difficult to predict the situation in the furnace. The current silicon prediction model generally regards the prediction of silicon content in hot metal as a conventional regression problem or time series problem. Few people mention this point about the internal situation of blast furnace, which leads to large fluctuation and inaccurate prediction results. In this paper, the internal state of blast furnace is considered: in different internal states of blast

furnace, even if the input data has been detected in the historical data, the predicted results may be different. One possibility of this situation is that the historical input variables affect the input variables at the next moment. For example: A large amount of coal is added to the blast furnace at t1. Even if a small amount of air is blown by the blower at t2, a small amount of coal is added. Because of the influence of coal in the furnace at t1, the furnace temperature will still be maintained at a higher temperature. Based on this idea, a new block data is generated from the original data through the time window, and the internal state characteristics of the blast furnace are extracted by using convolution neural network. After verification, the accuracy of the blast furnace model is greatly improved by the extracted information. The specific network model is shown in the figure 1:

Figure 1. Model structure

3 Related work

3.1 Convolutional Neural Network

Convolution neural network is a kind of feedforward neural network, in which convolution layer is the core of CNN. The convolution operation of local connection and weight sharing is used to process the data, and the deep features of data can be obtained. The convolution formula is as follows:

$$C = f(X \otimes W + b) \quad (1)$$

Where C is the output characteristic graph of convolution, X is the input data, $f()$ is the nonlinear demerit function, $\otimes$ is the convolution operation, W is the weight vector of convolution kernel, and b is the bias term.

The feature of convolution neural network is that it can extract hidden features in data and generate abstract high-level features.

3.2 Independently recurrent neural network-Recurrent neural network

The traditional multilayer perceptron and convolutional neural network think that there is no relationship between the input data, but RNN is a neural network which considers the relationship between the sequence before and after. The RNN neural network is connected between the hidden layers, and across the time point, the information of the hidden layer can be transferred to the hidden layer at the next moment.

Independently recurrent neural network (IndRNN)^[10] is a variant of RNN model, Different from the traditional RNN model, IndRNN replaces the traditional matrix multiplication with Hadamard product, and chooses to use Relu activation function instead of sigmoid activation function.

The calculation formula of IndRNN can be expressed as follows:

$$h_t = \sigma(wx_t + u \odot h_{(t-1)} + b) \quad (2)$$

The calculation method of the nth neuron is as follows:

$$h_{(n,t)} = \sigma(w_n x_t + u_n h_{(n,t-1)} + b_n) \quad (3)$$

Compared with the traditional RNN model, IndRNN transforms the weight matrix U into an independent weight vector u . In the above formula, $\odot$ denotes the product of matrix elements. That is, at time t , each neuron only receives the input at the moment and its own state at $t-1$ as input. Therefore, each neuron can process its own output independently and is not interfered by the neurons in the past. Moreover, because Relu activation function is added, it can effectively avoid gradient explosion and gradient disappearance, and ensure stable convergence even in the face of longer sequences.

3.3 Attention mechanism

Attention mechanism is a special mechanism that imitates the characteristics of human observation. When people observe something, they usually don't allocate all their energy to all the objects they see, but they focus their attention on some important observation points. The parts that are allocated to the attention tend to carry more

valuable information than those that are not paid attention to. Paying attention to the details of key targets and suppressing useless information will bring better fitting effect to the model. Nowadays, attention mechanism has become one of the core technologies in deep learning technology.

In this paper, attention mechanism is introduced into the prediction of blast furnace temperature. By paying attention to the importance of each variable to the overall data, attention mechanism allocates the weight of variables and distinguishes the importance of variables, which can reduce the impact of data noise on model prediction to a certain extent and make the model more stable.

In this paper, the calculation formula of attention mechanism weight matrix is as follows:

$$e = a(vi) = \sigma(wx^T + b)^T \quad (4)$$

Where σ is the softmax activation function.

Figure 2. Schematic diagram of attention mechanism

4 Experimental simulation

In order to verify the prediction of this algorithm in the actual production of blast furnace, this paper first uses expert experience and time window to generate samples of data, and compares it with multi-layer perceptron and LSTM model to verify the effectiveness of the model. The specific process of the experiment is as follows.

4.1. Data preprocessing

In the production of hot metal in blast furnace, there are many kinds of influencing variables of silicon content in

hot metal, which are mainly divided into two types: state parameters and control parameters. Among them, the main state parameters are gas utilization rate, heat loss of blast furnace cooling system, feeding speed; the main control parameters are theoretical coke ratio, blast temperature, blast humidity, theoretical coal injection rate. Based on the experience of experts, this paper makes a preliminary selection of forecast parameters. In order to ensure that the silicon content data and the predictive variables have the same frequency, the selected data is subjected to segmented three Hermite interpolation processing^[12,13]. Finally, seven input variables for silicon content prediction were determined. The details are shown in Table 1.

Table 1. Features used in this model

Symbol	Quantity	unit
X1	feeding batch number per hour	1/h
X2	Gas utilization	%
X3	Thoery coke rate	kg/t
X4	Thoery PCI rate	kg/t
X5	heat loss of BF cooling system	MJ/h
X6	Blast temperature	°C
X7	Previous Si-content	%

4.2 Data standardization

Because the magnitudes of different features in the blast furnace are too different, the data needs to be unified into the same dimension, the data standardization formula is shown below

$$x_{train}^* = \frac{x_{train} - \bar{x}_{train}}{\sigma_{train}} \quad (5)$$

$$x_{test}^* = \frac{x_{test} - \bar{x}_{train}}{\sigma_{train}} \quad (6)$$

Where $\bar{x}_{train}$ is the mean value of training set data, σ_{train} is the standard deviation of training set data, x_{train}^* is the standardized training set data, x_{test}^* is the standardized test set data.

5 Experimental results and analysis

The data in this paper are collected from the real-time production data of a steel plant from August to December 2014, and 300 test data are generated by rolling through the sliding window, and the multi-layer perceptron and

LSTM are established for comparative experiments. Multilayer perceptron is a classical neural network with multiple hidden layers, which can be highly parallel and widely used. LSTM network has been introduced into the field of prediction of silicon content in hot metal of blast furnace in recent two years, and has a good performance in solving the problem of blast furnace with time series method. The evaluation criteria of this paper are MSE and accuracy, and the specific formula is as follows.

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_t^{(i)} - \hat{y}_p^{(i)})^2 \quad (7)$$

$$HR = \frac{1}{m} \sum_{f=1}^m H_f \times 100\% \quad (8)$$

$$H_f = \begin{cases} 1, & |y_p - y_t| \leq 0.1 \\ 0, & |y_p - y_t| > 0.1 \end{cases} \quad (9)$$

Where m is the total number of samples, y_t is the true value of the silicon content in the test set, $\hat{y}_p$ is the prediction value of the model for silicon content of the test set, and HR is the hit rate.

The experimental results are as follows:

Figure 3. Comparison of Dense with Att-CNN-IndRNN

Figure 4. Comparison of LSTM with Att-CNN-IndRNN

Table 2 MSE and accuracy

Evaluation standard	Experimental results		
	Multilayer perceptron	LSTM	Att-CNN-RNN
MSE	0.00809	0.00713	0.00453
Accuracy	75%	79%	86%

In the above figure, the blue "x" represents the actual measured silicon value, the red solid line represents the actual effect curve of the Att-CNN-RNN model, and the black solid line represents the actual effect curve of the multilayer perceptron and the LSTM model.

In Figure (3), because the multilayer perceptron model does not have the memory advantage of the RNN model, the overall model performance fluctuates greatly, and the performance is not as good as the latter two. In Figure (4), although LSTM has a memory advantage, the extraction of the internal conditions of the blast furnace is not ideal, and the fluctuation data cannot be well grasped. Both of these two models are more susceptible to noise because they do not have an attention mechanism, causing relatively large prediction deviations.

The experimental results show that the Att-CNN-RNN model can better predict the trend of the furnace temperature, indicating that the model that adds the attention mechanism and considers the furnace conditions can better grasp the furnace temperature fluctuations.

6 conclusion

According to the characteristics of blast furnace molten iron engineering, this paper proposes a time window plus convolution to extract the internal state of the blast furnace; the model is verified by actual production data with complex changes; aiming at the large noise in the blast furnace molten iron data, a hybrid network combining convolutional network, attention mechanism and IndRNN is used in the prediction of the silicon content of the blast furnace molten iron. After data comparison, this model has a better prediction effect in the data set with fluctuation and noise, and provides a new research idea for the prediction of the silicon content of molten iron.

Reference:

1. Yang, Y. L., Zhang, S. and Yin, Y. X. (2016). A modified elm algorithm for the prediction of silicon content in hot metal, *Neural Comput & Applic* 27(1): 241–247. DOI 10.1007/s00521-014-1775-x.
2. Jiang Chaohui, Dong Menglin, Gui Weihua, et al. Two-dimensional prediction of silicon content in blast furnace hot metal based on Bootstrap. *Acta Automatica Sinica*, 2016, 42(5):715–723.
3. Chen, J. (2001). Predictive system for blast furnaces by integrating a neural network with qualitative analysis, *Engineering Applications of Artificial Intelligence* 14(1): 77–85. DOI 10.1016/S0952-1976(00)00062-2.
4. Yin Juping, Jiang Chaohui. Data-based prediction of

silicon content in blast furnace hot metal[J]. *Design and Research of Nonferrous Metallurgy*,2015,36(03):36-38+41.

5. Sun Jie, Cui Tingting, Xu Bin, Wang Xingnan. Prediction of silicon content in molten iron based on IGA-ELM[J]. *Journal of North China University of Technology* (Natural Science Edition),2020,42(01):25-29.
6. Jian, L., Song, Y. Q., Shen, S. Q., Wang, Y. and Yin, H. Q. (2015). Adaptive least squares support vector machine predictor for blast furnace ironmaking process, *ISIJ International* 55(4): 845–850. DOI 10.2355/isijinternational.55.845.
7. Ye Fei. Prediction method of blast furnace silicon content based on improved BP neural network [D]. *Anhui University of Technology*,2019.
8. Yang, Y. L., Zhang, S. and Yin, Y. X. (2016). A modified elm algorithm for the prediction of silicon content in hot metal, *Neural Comput & Applic* 27(1): 241–247. DOI 10.1007/s00521-014-1775-x.
9. Li Zelong, Yang Chunjie, Liu Wenhui, Zhou Heng, Li Yuxuan. Prediction of silicon content in molten iron based on LSTM-RNN model[J]. *CIESC Journal*,2018,69(03):992-997.
10. Li S, Li W, Cook C, et al. Independently recurrent neural network (IndRNN): Building a longer and deeper rnn[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018: 5457-5466.
11. Vaswani A ,Shazeer N ,Parmar N , et al. Attention Is All You Need[J]. *arXiv*, 2017.
12. Dhar, S. and Dash, P. K. (2017). Adaptive threshold based new active islanding protection scheme for multiple PV based microgrid application, *Iet Generation Transmission & Distribution* 11(1): 118–132. DOI 10.1049/iet-gtd.2016.0679.
13. Acar, Y., Dogan, H., Basar, E. and Panayirci, E. (2016). Interpolation based pilot-aided channel estimation for STBC spatial modulation and performance analysis under imperfect CSI, *Iet Communications* 14(10): 1820–1828. DOI 10.1049/iet-com.2016.0151.