

Research Article

Soft Sensor Model Based on Improved Elman Neural Network with Variable Data Preprocessing and Its Application

Hai-bo Zhu  and Yong Zhang 

School of Electronics and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, China

Correspondence should be addressed to Yong Zhang; 2550345485@qq.com

Received 2 September 2018; Accepted 14 November 2018; Published 13 December 2018

Academic Editor: Jesús Lozano

Copyright © 2018 Hai-bo Zhu and Yong Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to solve the problems of strong coupling, nonlinearity, and complex mechanism in real-world engineering process, building soft sensor with excellent performance and robustness has become the core issue in industrial processes. In this paper, we propose a new soft sensor model based on improved Elman neural network (Elman NN) and introduce variable data preprocessing method to the soft sensor model. The improved Elman NN employs local feedback and feedforward network mechanism through context layer to accurately reflect the dynamic characteristics of the soft sensor model, which has the superiority to approximate delay systems and adaption of time-varying characteristics. The proposed variable data preprocessing method adopts combining Isometric Mapping (ISOMAP) with local linear embedding (LLE), which effectively maintains the neighborhood structure and the global mutual distance of dataset to remove the noises and data redundancy. The soft sensor model based on improved Elman NN with variable data preprocessing method by combining ISOMAP and LLE is applied in practical sintering and pelletizing to estimate the temperature in the rotary kiln calcining process. Comparing several conventional soft sensor model methods, the results indicate that the proposed method has more excellent generalization performance and robustness. Its model prediction accuracy and anti-interference ability have been improved, which provide an effective and promising method for the industrial process application.

1. Introduction

Soft sensor is an inferential prediction virtual technique which adopts easily measured variables to predict the process variables which are difficult to measure directly because of technological and economical limitations or complex environment. The soft sensor tries to establish a regression prediction model between easily measured variables and difficultly measured variables, which is adopted to solve the problem that prevents the measurements from being employed as feedback signals for quality control systems. And the soft sensor methods have been employed more widely in the industrial process and become a major developing trend in both academia and industry [1].

In view of the model prediction in the industrial production process, early scholars put forward some model

predictive control, such as generalized predictive control [2, 3], dynamic matrix predictive control [4], and model algorithm control [5]. However, these soft sensor prediction methods suffer from common shortcomings. One is that the prediction control of nonlinear systems cannot be effectively solved. Another is that the stability and robustness of multivariable predictive control algorithm need to be solved and it is very difficult to build accurate principle models for complex processes. At present, with the deep research of soft sensor control theory and the continuous progress of engineering technology, some artificial intelligence and machine learning methods based on data-driven technologies are proposed to solve the problems which are difficult to measure key process and quality variables for soft sensor models, such as artificial neural network (ANN) [6–8], rough set [9], partial least squares [10], support vector machine

(SVM) [11–13], and hybrid methods [14, 15]. Some new works on advanced predictive control based on neural networks are referred to, for example, Ławryńczuk proposed MPC algorithms based on double-layer perceptron neural models [16]. And these data-driven soft sensors have been implemented successfully in many industrial application fields, like chemical and metallurgical engineering industries.

However, the complexity of the industrial model will increase and bring about the problem of dimensionality with variables increasing. In addition, there are strong correlations and considerable redundancy between variable data. And the collected variable data may be contaminated by random noises or abnormal working condition data, which indicates that the process variables for soft sensor are stochastic. If those variable data samples are used directly for establishing the soft sensor model, the performance of the soft sensor model may not be guaranteed, which results in terrible prediction accuracy or poor estimation. Hence, it is significant to discover the general trends of data by latent variable models before soft sensor modeling. And the randomness of variable sample data should be taken into consideration and the redundant variable sample data need to be discarded selectively through the use of appropriate variable selection techniques. The core is to exploit the essential information behind the effective sample data to establish a soft sensor model that has excellent performance and strong robustness. The shortcoming of variable sample data has advanced the development of the dimension reduction methods. So many data preprocessing methods for the soft sensor model have been researched in recent years. Sun et al. [17] presented the principal component analysis (PCA) method to analyze the correlation of input variables, which is effective to reduce the dimensions of input variables from the chemical oxygen demand prediction model based on LS-SVM. Najafi et al. [18] proposed a new method called “Path-Based Isomap” to find the low-dimensional embedding. The low-dimensional embedding is computed by exploiting geodesic path-mapping algorithm. Besides, some manifold learning methods have been proposed including Laplacian eigenmaps [19, 20], locally linear embedding [21], and Isometric Mapping method [22, 23]. However, these linear and nonlinear data preprocessing methods have their own limitations, which can only maintain a single feature of the variable dataset and cannot take into account other valuable characteristics. In order to overcome this problem, a data preprocessing method based on combining Isometric Mapping (ISOMAP) with local linear embedding (LLE) for input variable dataset is proposed. Hence, this paper establishes a soft sensor model based on improved Elman NN with input variable data preprocessing method which integrates ISOMAP with LLE under the kernel framework to estimate the quality variable and makes it have better performance and robustness.

The rest of this paper is organized as follows. Section 2 describes the improved Elman NN and data preprocessing method. Section 3 proposes a soft sensor model based on improved Elman NN with variable data selection method by combining ISOMAP with LLE and describes it in detail. In Section 4, an industrial case study on the prediction of

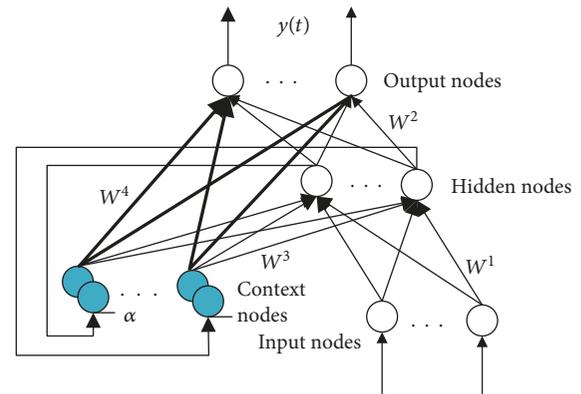


FIGURE 1: Architecture of improved Elman NN.

temperature in a rotary kiln calcining process in pellet sintering is given. Finally, conclusions are made in Section 5.

2. Background Theory

For convenience, some definitions that will be used throughout this paper are given: NN: neural network; ISOMAP: Isometric Mapping; LLE: local linear embedding.

2.1. Improved Elman Neural Network. Elman neural network (Elman NN) [24–27] is a typical dynamic neural network with recurrent feedback layer based on BP neural network. The Elman NN stores the internal state through the context layer, which has the performance of mapping dynamic characteristics. The self-connected mode makes it sensitive to the value of the historical state and it has better ability to approximate nonlinear mapping than the BP NN. However, this feedback only aims at the hidden layer to adjust the connection weights of the internal feedback, which lacks the ability of global adjustment and dynamic stability. And the learning identification ability has some limitations.

To further improve the dynamic performance and nonlinear approximation capability of Elman NN, the output of context layer as a feedback signal should be fully taken advantage of. Figure 1 shows the architecture of the improved Elman NN that is consisted of an input layer, an output layer, a hidden layer, and a context layer which is connected to hidden layer neurons and output layer neurons.

From Figure 1 we can see that the improved Elman NN adds input signal from context layer nodes to the output layer, which are connected through a novel connection weight matrix W_{ij}^4 between the context layer and output layer. After the hidden layer process, the output of the hidden layer returns to the context layer, then the intermediate state value of the context layer is transmitted to the output layer and hidden layer through the connection weight matrices W_{ij}^4 and W_{ij}^3 separately, which constitutes a feedforward and local feedback network. The feedback loop is formed in the hidden layer and context layer, and the path of feedforward is consisted of the input layer, hidden layer, context layer, and output layer, as shown in Figure 2.

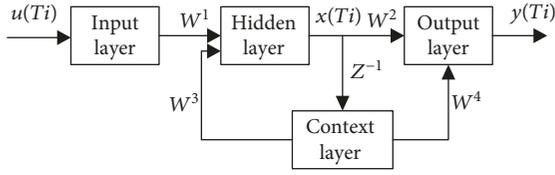


FIGURE 2: Structure block diagram of improved Elman NN.

We define the following symbols: the number of input layer nodes, output layer nodes, and hidden layer nodes is m , n , and l , respectively. The network consisted of input data u , intermediate state oh , and response output y . The activation function of structure from the input layer to the hidden layer and from the hidden layer to the output layer is defined f_θ and g_θ , respectively.

The mathematical model in which the input layer is mapped to the hidden layer through the mapping activation function is

$$oh_j(k) = f_\theta(u_i(k)) = s \left(\sum_{i=1}^m w_{i,j}^1 u_i(k) + \sum_{i=1}^m w_{i,j}^3 oc_i(k) + b_j^h \right), \quad (1)$$

where s represents the *sigmoid* function. Equation (1) is parameterized through $\theta = \{w_{ij}^1, w_{ij}^3, b_j^h\}$, in which w_{ij}^1 is the weight of input layer node i to the hidden layer node j ; w_{ij}^3 is the weight of context layer node i to the hidden layer node j and b_j^h is the bias vector of the hidden layer.

The hidden layer oh_j is mapped to the output space, y is the output vector of oh_j , and the activation function is

$$y = g_\theta(oh_j, oc_i) = \sum_{i=1}^n w_{ij}^2 oh_j(k) + \sum_{i=1}^l w_{ij}^4 oc_i(k) + b^o. \quad (2)$$

Equation (2) is decided by $\theta' = \{w_{ij}^2, w_{ij}^4, b^o\}$, where w_{ij}^2 is the weight of latent layer node i to the output layer node j ; w_{ij}^4 is the connection weight of context layer node i to the output layer node j and b^o represents the bias vector of the output layer.

In the training process, the output of the context layer $oc_i(k)$ is the sum of output of the hidden layer $oh_j(k-1)$ and α times output of context layer at a previous time.

$$oc_i(k) = oh_j(k-1) + \alpha \cdot oc_i(k-1). \quad (3)$$

The loss function is defined as the squared error:

$$L(\tilde{y}_d, y) = \|\tilde{y}_d(k) - y(k)\|^2, \quad (4)$$

where $y_d(k)$ is the predicted output value at time k ; $y_i(k)$ is the actual output value at time k .

A novel set of parameter vector of the improved Elman NN can be obtained and optimized to minimize the expected average squared error:

$$\begin{aligned} (\theta^*, \theta'^*) &= \arg \min_{\theta, \theta'} \frac{1}{n} \sum_{i=1}^n L(\tilde{y}_{d,i}(k), y_i(k)) \\ &= \arg \min_{\theta, \theta'} \frac{1}{n} \sum_{i=1}^n L(g_{\theta'}(f_\theta(u_i(k))), y_i(k)). \end{aligned} \quad (5)$$

2.2. Input Variable Data Selection and Dimension Reduction Method. Considering the original input variables of improved Elman NN exist with redundant variable data and abnormal condition variable data, a novel variable data selection methodology by integrating Isometric Mapping (ISOMAP) [28] and local linear embedding (LLE) [29, 30] is proposed to handle measurement uncertainty and determine the dimensionality of the latent variable. The purpose is to combine two manifold learning algorithms effectively under the kernel framework and maintain various characteristics of data set, which establishes a dimensionality reduction method based on manifold learning for high-dimensional nonlinear data sets to effectively solve the network topology and the complex training problems caused by the oversize input vector data. The novel methodology can be divided into three stages. At the first stage, we construct the Mercer kernel matrix K_{ISOMAP} of Isomap. Second, we extract the kernel matrix K_{LLE} of LLE. At the third stage, we obtain the kernel matrix $K_{ISOMAP-LLE}$ under the kernel framework.

The general steps of the data dimensionality reduction process are the following.

Step 1. Use K neighborhood method or ϵ radius method to find neighborhood values.

Step 2. Use the shortest path method to calculate the geodesic distance d_{ij} and to form the geodesic distance matrix D^2 , in which $D_{ij}^2 = |d_{ij}^2| \in R^{N \times N}$. The matrix $K(D^2)$ is constructed on the basis of approximate geodesic distance matrix D^2 .

$$K(D^2) = -\frac{1}{2} H D^2 H, \quad (6)$$

where H represents a centralization matrix, $H = I - (1/N)ee^T$ and $e = [1 \ 1 \ \dots \ 1]^T \in R^{N \times 1}$.

Step 3. The matrix expression is constructed according to the matrix $K(D^2)$.

$$\begin{bmatrix} 0 & 2K(D^2) \\ -1 & -4K(D) \end{bmatrix}. \quad (7)$$

Step 4. Obtain the maximum eigenvalue c^* of the matrix and construct the Mercer kernel matrix K_{ISOMAP} according to the maximum eigenvalue c^* .

$$K_{\text{ISOMAP}} = K(D^2) + 2cK(D) + \frac{1}{2}c^2H. \quad (8)$$

In order to ensure that K_{ISOMAP} is a real symmetric semi-definite matrix, we need to satisfy $c \geq c^*$.

Step 5. For each data point of the LLE algorithm $x_i (i = 1, 2, \dots, N)$, solve its cost function $\varepsilon = |x_i - \sum_{j=1}^k w_{ij}x_{ij}|^2$.

The weight coefficient w_{ij} of data points is determined, and the weight matrix W is constructed. At the same time, a real symmetric semipositive definite matrix M is obtained based on the weight matrix W , and the kernel matrix K_{LLE} of the LLE algorithm is extracted.

$$M = (I - W)^T(I - W), \quad (9)$$

$$K_{\text{LLE}} = \lambda_{\max}I - M, \quad (10)$$

where λ_{\max} is the largest eigenvalue of the weight matrix W .

Step 6. Obtain the kernel matrix of $K_{\text{ISOMAP-LLE}}$ by the kernel matrix K_{ISOMAP} of the ISOMAP algorithm and the kernel matrix K_{LLE} of the LLE algorithm. And the kernel matrix regulator β is introduced to adjust the weight of the ISOMAP and LLE algorithms, and a new kernel matrix $K_{\text{ISOMAP-LLE}}$ is constructed based on the kernel function.

$$K_{\text{ISOMAP-LLE}} = K_{\text{ISOMAP}} + (1 - \beta)K_{\text{LLE}}. \quad (11)$$

Step 7. Calculate the eigenvalue matrix P and eigenvector matrix Q of the new kernel matrix $K_{\text{ISOMAP-LLE}}$ and obtain the low-dimensional embedded coordinates Y based on MDS:

$$Y = \Lambda^{1/2}Q, \quad (12)$$

where Λ is the diagonal matrix of P .

3. Soft Sensor Model Based on Improved Elman NN with Dimension-Reduction

3.1. Soft Sensor Model Based on Improved Elman NN with Variable Data Preprocessing. Soft sensor model based on improved Elman NN with variable data selection is aimed at exploiting the essential information behind the process data and filter redundant variable data in the soft sensor model. This paper proposes an input variable data selection and dimension-reduction methodology by integrating ISOMAP and LLE manifold learning algorithm, which is effective to solve redundant variable data and abnormal working data of the oversized sample data. This structure of the soft sensor model based on improved Elman NN with variable data selection is shown in Figure 3. The soft sensor model consisted of two parts: an input data selection (ISOMAP and LLE) and a supervised learning (improved Elman NN). In this paper, the variable screening method adopted the manifold learning variable selection method based on the

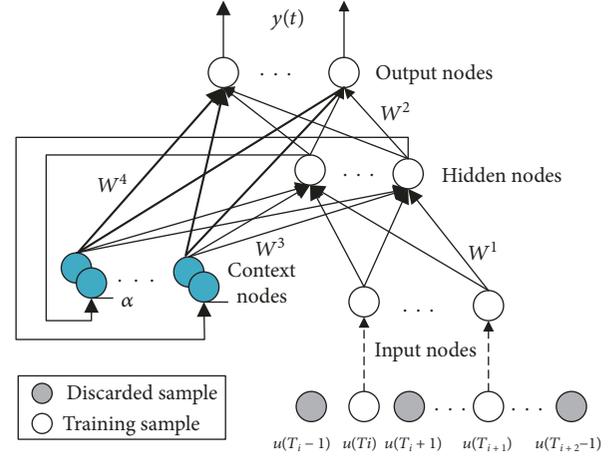


FIGURE 3: The structure of the soft sensor model based on improved Elman NN with dimension-reduction.

fusion of LLE and ISOMAP under the kernel framework, which is mainly reflected in two aspects. (1) The initial input variables should be chosen reasonably to make the “quality” of initial input variables be ensured. (2) The useful information of input variables should not be prematurely eliminated. And adopting supervised criterion trains each layer to obtain the weights which are set as the initial weights instead of the random initial weights.

In Figure 3, W^1 , W^2 , W^3 , and W^4 are the weights of improved Elman NN, respectively, which are adopted as the weights of corresponding layers in the improved Elman NN optimized with respect to a supervised training criterion. The input data of the soft sensor model consisted of training sample data and discarded sample data: the discarded sample data U_{discard} , i.e., noise and redundant values of the primary input variables, and the training sample data (U_{train} , Y), i.e., pairs of input-output sample data without redundant and abnormal working data, are used to train improved Elman NN to have better accuracy for the soft sensor model.

3.2. The Training Based on Supervised Learning. In the training process of improved Elman NN, each layer in the feedforward path is considered as one-layer BP network which can update weight parameters by gradient descent. The representation of i th hidden layer is regarded as the input for the $i + 1$ th context layer which feeds the previous time state of the hidden layer nodes back to the hidden layer. The error is adopted to refresh the weights.

The loss function can be the traditional squared error as in equation (4). Equation (5) can be reformulated by optimizing the parameters to minimize the average squared error. The gradient descent method is used to update parameters of the improved Elman NN and amend the weights $w_{i,j}^1$, $w_{i,j}^2$, $w_{i,j}^3$, and $w_{i,j}^4$ to achieve self-adaptive learning. At time k , we can get

$$\Delta w_{ij}^1 = \eta_1 \sum_{k=1}^T \sum_{i=1}^m \left[(\hat{y}_{d,i}(k) - y_i(k)) w_{ij}^2(k) \right] f'(\cdot) \cdot u_i(k), \quad (13)$$

where $u_i(k)$, $i = 1, 2, \dots, m$, is the signal of input layer nodes. k , $k = 1, 2, \dots, T$, is the current operation iterations. η_1 is the learning rate of the weights of $w_{i,j}^1$.

$$\Delta w_{ij}^2 = \eta_2 \sum_{k=1}^T \sum_{i=1}^n [(\tilde{y}_{d,i}(k) - y_i(k)) \text{oh}_j(k)], \quad (14)$$

where η_2 is the learning rate of the weights of $w_{i,j}^2$.

$$\begin{aligned} \Delta w_{ij}^3 &= \eta_3 \left(\frac{\partial L}{\partial \text{oh}_i(k)} \cdot \frac{\partial \text{oh}_i(k)}{\partial w_{ij}^3(k)} + \frac{\partial L}{\partial \text{oc}_i(k+1)} \cdot \frac{\partial \text{oc}_i(k+1)}{\partial w_{ij}^3(k)} \right) \\ &= \eta_3 \left\{ \sum_{i=1}^l [(\tilde{y}_{d,i}(k) - y_i(k)) w_{ij}^2(k)] f'(\cdot) \text{oc}_i(k) \right. \\ &\quad \left. + \sum_{i=1}^l [(\tilde{y}_{d,i}(k+1) - y_i(k+1)) w_{ij}^4(k+1)] f'(\cdot) \text{oc}_i(k) \right\}, \end{aligned} \quad (15)$$

where η_3 is the learning rate of the weights of $w_{i,j}^3$. oc_i is obtained as in equation (3).

$$\Delta w_{ij}^4 = \eta_4 \sum_{k=1}^T \sum_{i=1}^n [(\tilde{y}_{d,i}(k) - y_i(k)) \text{oc}_i(k)], \quad (16)$$

where η_4 is the learning rate of the weights of $w_{i,j}^4$.

A new set of weights W_{ij} can be updated by

$$\tilde{w}_{ij}^{k+1} = w_{ij}^k + \Delta w_{ij}^{k+1}. \quad (17)$$

Then, the prediction model of y is obtained by the new weight parameters (θ^*, θ'^*) .

$$\begin{aligned} \tilde{y}_{\theta^*, \theta'^*} &= g \left(\sum_{i=1}^n \tilde{w}_{ij}^2 f \left(\sum_{i=1}^m \tilde{w}_{i,j}^1 u_i(k) + \sum_{i=1}^l \tilde{w}_{i,j}^3 \text{oc}_i(k) + b_j^h \right) \right. \\ &\quad \left. + \sum_{i=1}^l \tilde{w}_{ij}^4 \text{oc}_i(k) + b_j^o \right). \end{aligned} \quad (18)$$

3.3. Formulated the Criterion of Evaluating. After the variable selection and training, the prediction value of the soft sensor model based on improved Elman NN as in equation (18) needs be further evaluated. In this paper, the evaluation criterion, the Bayesian information criterion (BIC), is adopted as a measure between the complexity of the soft sensor model and the accuracy.

$$\text{BIC}_{\theta^*, \theta'^*} = n \cdot \log_e \left(\frac{1}{2} \sum_{i=1}^n (\tilde{y}_i - y_i)^2 \right) + m \cdot \log_e n, \quad (19)$$

where \tilde{y}_i is the prediction value by modified weight parameters as in equation (18). y is the actual value. n represents the

TABLE 1: Candidate variables of the soft sensor model for temperature prediction in the rotary kiln calcining process.

Variables	Description	Unit
u_1	The speed of the chain grate	m/min
u_2	The speed of the rotary kiln	rin/min
u_3	The speed of the ring refrigerator	m/min
u_4	The temperature of rotary kiln head	°C
u_5	The temperature of rotary kiln tail	°C
u_6	The quantity of coal injection	t
u_7	I section temperature of the cooler	°C
u_8	II section temperature of the cooler	°C
u_9	Material thickness in the chain grate	mm
u_{10}	Preheating I section temperature of the smoke hood in chain grate	°C
u_{11}	Preheating II section temperature of the smoke hood in chain grate	°C

observational number of the soft sensor model. m is the number of selected variables through integrating ISOMAP and LLE.

The optimal parameters can be obtained by the minimization of the average BIC values:

$$(\theta^*, \theta'^*) = \arg \min \left(\frac{1}{N} \sum_{n=1}^N \text{BIC}(\theta, \theta') \right). \quad (20)$$

The procedure of the soft sensor model based on improved Elman NN with selected variable and dimension-reduction is summarized as follows.

Step 1. Initialization. Train a new improved Elman NN by the training dataset and obtain initial weight parameters of the network.

Step 2. Selection of input variable dataset.

Step 2.1. Initialize variable data preprocessing method based on combining ISOMAP with LLE.

Step 2.2. For the current variable dataset, calculate approximate geodesic distance to obtain the maximum eigenvalue c^* of the matrix and construct the Mercer kernel matrix K_{ISOMAP} by equation (8).

Step 2.3. Obtain the real symmetric semipositive definite matrix M according to cost function ε and extract the kernel matrix K_{LLE} by equation (10).

Step 2.4. Calculate $K_{\text{ISOMAP-LLE}}$ by equation (11) based on Step 2.2 and Step 2.3.

Step 2.5. Calculate Y by equation (12).

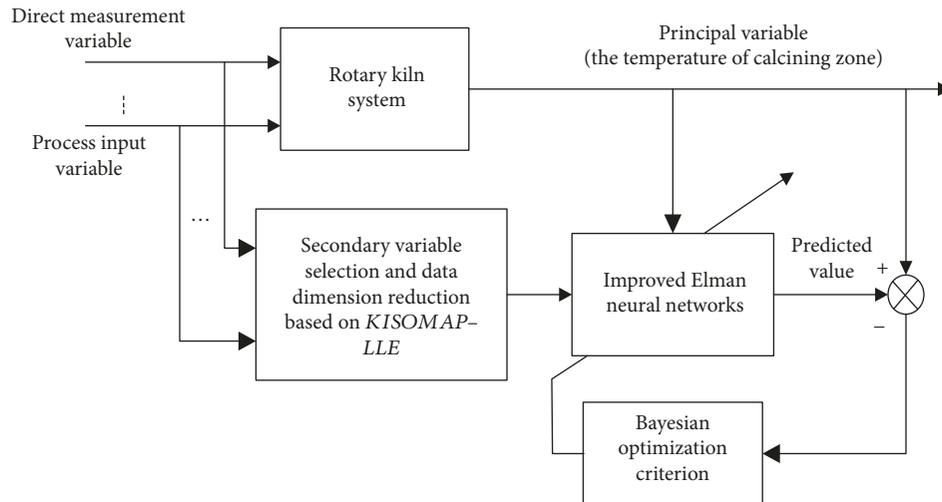


FIGURE 4: The diagram of temperature prediction in the rotary kiln calcining process based on improved Elman NN with variable data preprocessing.

Step 2.6. Obtain the input variable data without redundancy and noises disturbing.

Step 3. Evaluate the parameter optimization by equation (19) and obtain the new parameters by equations (5) and (20).

Step 4. Update the weight parameters by equation (17) and obtain novel neural network by equation (18).

4. Case Study: Application to Temperature Prediction in Rotary Kiln Calcining Process

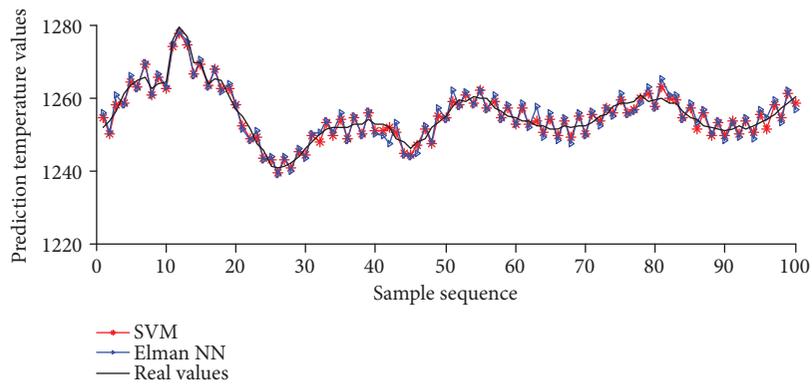
The rotary kiln is an important part of pellet sintering. Green pellets with a diameter of 9~16 mm are generated in the pellet making system. Then the pellets pass through the kiln tail into the rotary kiln after preheating at the chain grate. In the rotary kiln, the oxidizing roast and consolidation process of pellets is completed, which makes pellets have the physical and chemical properties of the blast furnace burden. To ensure and improve the product quality of pellets, the real-time control of temperature in the rotary kiln calcining process is very important. And the temperature in the rotary kiln calcining process is directly related to production efficiency, energy consumption, and emissions of harmful gas.

By the analysis of the heat balance mechanism and technological characteristics of the rotary kiln calcining process, the temperature in the rotary kiln calcining process is coupled with multiple technological index variables in the grate system and ring cooler system, which makes it very difficult to use the instrument directly to measure the temperature in the rotary kiln calcining process because of poor working conditions or high maintenance costs of hardware sensors. Thus, the soft sensor model based on improved Elman NN with variable data selected and Bayesian optimization is applied to estimate the temperature in the rotary kiln calcining process. For the prediction of the temperature, we use the technological index variables which are correlated with prediction temperature and are easy to measure as the

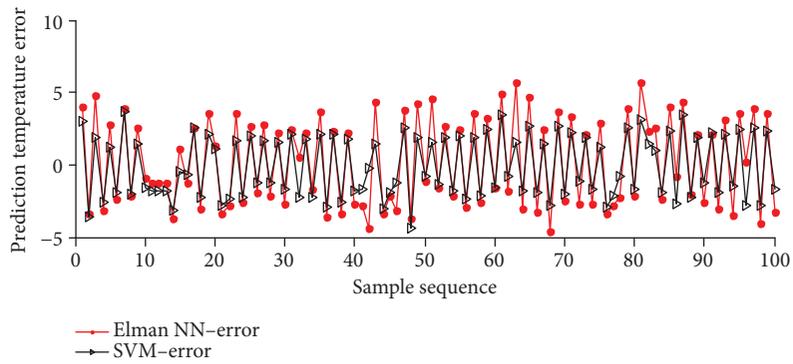
input variables of the soft sensor model in this process, which are listed in Table 1. Besides, a kind of physical and chemical reactions in the rotary kiln calcining process also affect the calcining temperature. A more detailed physico-chemical mechanism of the sintering process is described in literature [31, 32].

4.1. The Soft Sensor Model of Temperature in Rotary Kiln Calcining Process Based on Improved Elman NN with Variable Data Preprocessing. According to the technological characteristics of the rotary kiln and technology analysis, the secondary variables are selected based on the combining of ISOMAP with LLE through kernel function, which can simplify the redundant and abnormal data of the selected variables. The secondary variables are chosen as input of the soft sensor model based on improved Elman NN to predict the temperature in the rotary kiln calcining process. The Bayesian optimization criterion is applied to estimate the accuracy of prediction model and adjust the weights in order to amend the error. The structure diagram of temperature prediction in the rotary kiln calcining process based on improved Elman NN with variable data preprocessing is shown in Figure 4.

4.2. Results and Discussion. For the improved Elman NN training, data samples have been collected and are partitioned into two parts, in which 3000 process datasets of the secondary variables and corresponding temperature in the rotary kiln calcining process are selected for training, and the 1500 process datasets are applied as the testing dataset. At the same time, the data preprocessing method based on the combination of ISOMAP with LLE is used to remove the noise and data redundancy that exist in input data in order to improve prediction performance. Considering that SVM is a state-of-the-art soft sensor model with good generalization ability, the improved Elman NN with variable dataset preprocessing based on ISOMAP and LLE is compared with the soft sensor model based on SVM. The kernel

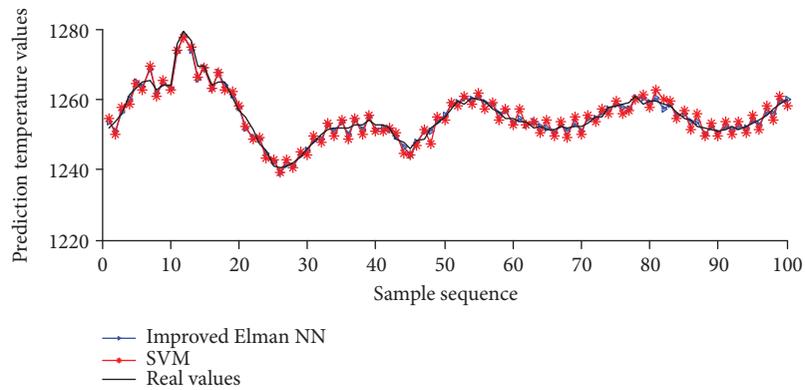


(a) Prediction curve of the kiln burning zone temperature

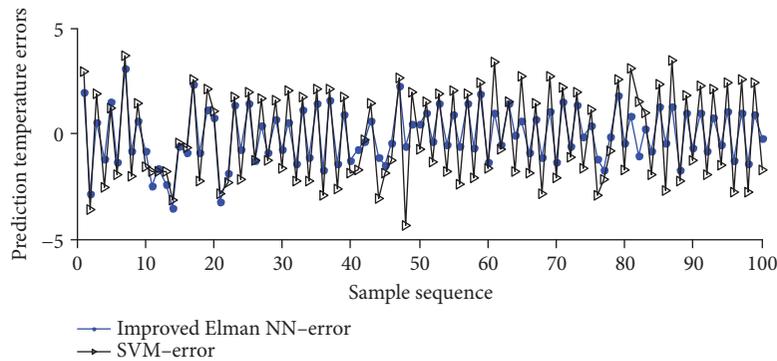


(b) Error curve of the kiln burning zone prediction temperature

FIGURE 5: Comparison of temperature generalization test in rotary kiln between Elman NN and SVM.



(a) Prediction curve of the kiln burning zone temperature



(b) Error curve of the kiln burning zone prediction temperature

FIGURE 6: Comparison of temperature generalization test in rotary kiln between SVM and improved Elman NN.

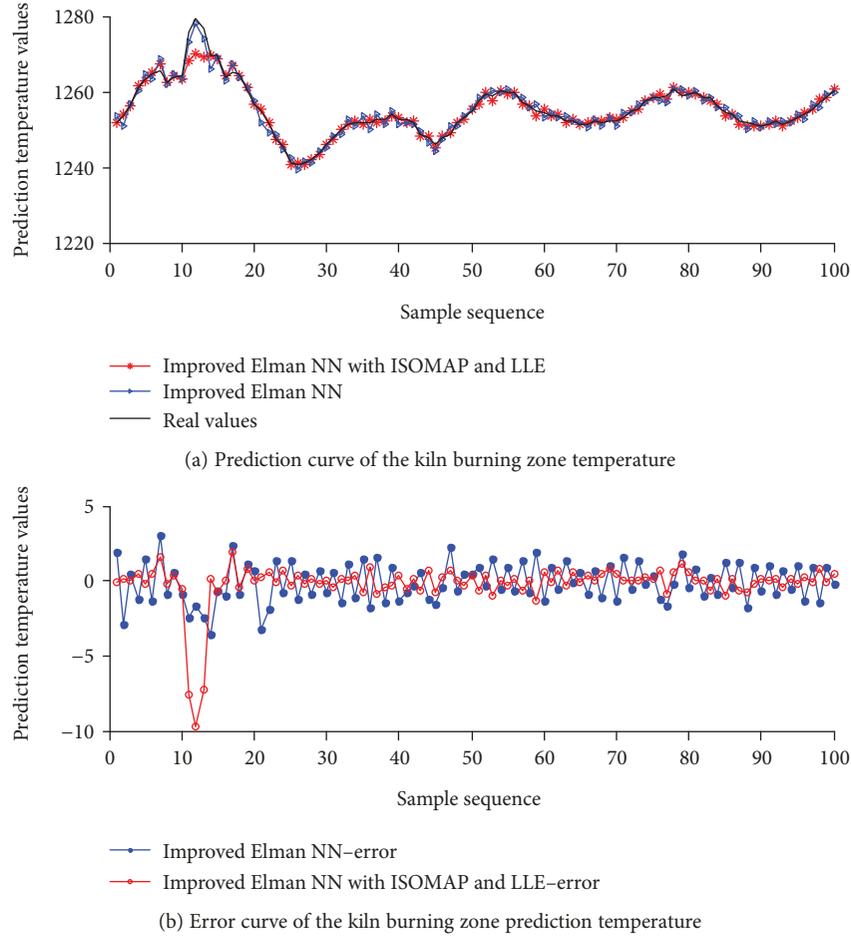


FIGURE 7: Comparison of temperature generalization test in rotary kiln between improved Elman NN and improved Elman NN with ISOMAP and LLE.

function of SVM is the radial basis kernel function and the number of support vectors in the SVM is specified as 6. In addition, for a fair comparison, the number of hidden nodes and context nodes is the same in the Elman NN and improved Elman NN. And the number of hidden nodes is set to 36.

In the improved Elman NN with data preprocessing experiments, the gradient descent is employed to train the improved Elman NN. The datasets are chosen by the ISOMAP and LLE method. In addition, the parameters are optimized and decided by comparing validation error curves according to Bayesian criterion. Once the weight parameters are fine-tuned, all training data are employed to train the model again. And the learning rate according to the loss's change is adopted to obtain an excellent performance. To test the prediction performance of the soft sensor model based on different model methods, 100 testing samples are shown in Figures 5–7 because of that the testing datasets are too large to show. To evaluate the performance of the soft sensor model, the root mean square error (*RMSE*), the maximum negative error (*MNE*), and the maximum positive error (*MPE*) are typically used. The values of the testing dataset calculated through the different soft sensor models are compared in Table 2.

TABLE 2: Performance comparison results of prediction temperature under different soft sensor models.

Soft sensor model	RMSE	MNE	MPE
Elman NN	3.0496	-4.582	5.697
SVM	2.1471	-4.332	3.719
Improved Elman NN	1.3003	-3.535	3.061
Improved Elman NN with variable data preprocessing	0.5450	-1.275	1.941

Table 2 shows that improved Elman NN with variable data preprocessing has more excellent learning performance than other soft sensor models. The learning ability of improved Elman NN is improved to some extent and the proposed variable data preprocessing method is effective to remove the redundant data and noises.

Figures 5 and 6 demonstrate that the improved Elman NN has good performance in predicting the temperature in the rotary kiln calcining process. The soft sensor model based on improved Elman NN has much better generalization and learning performance than the soft sensor model based on SVM and conventional Elman NN. The error curves of the SVM prediction, traditional Elman NN prediction, and improved Elman NN prediction which are shown in

Figures 5(b) and 6(b) clearly indicate the difference between the different methods. It can be seen that the prediction error from the improved Elman NN is smaller than that from the SVM and conventional Elman NN.

However, the three methods are limited to handle the noises and data redundancy which exist in input variable data to improve the prediction performance. The soft sensor model based on improved Elman NN with ISOMAP and LLE is able to remove abnormal working condition noise and data redundancy, as shown in Figure 7. It is seen that when several abnormal noise data are added in the early stage, the improved Elman NN with ISOMAP and LLE can effectively avoid disturbance which is from noise. The improved Elman NN prediction is smooth without large fluctuations and shows better robustness than SVM prediction and improved Elman NN prediction.

5. Conclusion

This paper develops a soft sensor model based on improved Elman NN and proposes a data preprocessing method based on ISOMAP and LLE for input variable data selection. The improved Elman NN introduces local feedback and feedforward networks into the Elman NN, which is trained to obtain better generalization and avoid overfitting to a certain extent. The data preprocessing method based on ISOMAP and LLE effectively keeps the neighborhood structure and the global mutual distance of the datasets and better reflects the real character of the original datasets, which is used to remove the abnormal working condition data, noises, and data redundancy in the input variables from improved Elman NN. The soft sensor model based on improved Elman NN with data preprocessing by ISOMAP and LLE is applied to estimate the temperature in the rotary kiln calcining process. The prediction values of the soft sensor model based on improved Elman NN with data preprocessing through ISOMAP and LLE follow the varying trend of the temperature very effectively and have excellent robustness and generalization performance. The great performance illustrates that the soft sensor model based on improved Elman NN with data preprocessing through ISOMAP and LLE provides a powerful and promising method for complex industrial process applications.

Data Availability

The prediction temperature of the rotary kiln calcining process in Sintering and Pelletizing data used to support the findings of this study have not been made available because the data which is used is the proprietary data from a Sintering and Pelletizing company that will share the data only with regulatory agencies but not with researchers. So the jurisdiction of the data is limited.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

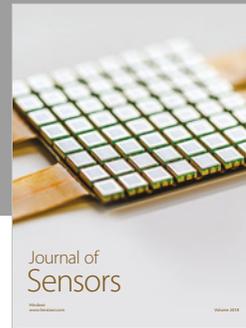
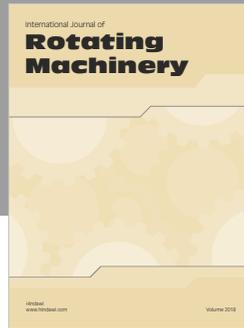
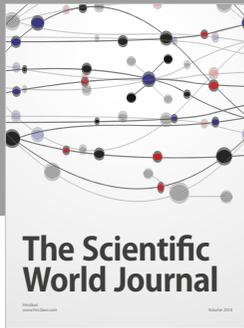
Acknowledgments

This work was supported in part by the Project by National Natural Science Foundation of China under Grant (61473054).

References

- [1] W. Yan, D. Tang, and Y. Lin, "A data-driven soft sensor modeling method based on deep learning and its application," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 4237–4245, 2017.
- [2] D. W. Clarke and C. Mohtadi, "Properties of generalized predictive control," *Automatica*, vol. 25, no. 6, pp. 859–875, 1989.
- [3] P. Albertos and R. Ortega, "On generalized predictive control: two alternative formulations," *Automatica*, vol. 25, no. 5, pp. 753–755, 1989.
- [4] C. R. Cutler and B. L. Ramaker, "Dynamic matrix control - computer control algorithm," *Joint Automatic Control Conference*, vol. 17, p. 72, 1980.
- [5] J. Richalet, A. Rault, J. L. Testud, and J. Papon, "Model predictive heuristic control: applications to industrial processes," *Automatica*, vol. 14, no. 5, pp. 413–428, 1978.
- [6] D. I. Adelani and M. K. Traoré, "Enhancing the reusability and interoperability of artificial neural networks with DEVS modeling and simulation," *International Journal of Modeling, Simulation, and Scientific Computing*, vol. 7, no. 2, article 1650005, 2016.
- [7] A. Afram, F. Janabi-Sharifi, A. S. Fung, and K. Raahemifar, "Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: a state of the art review and case study of a residential HVAC system," *Energy and Buildings*, vol. 141, pp. 96–113, 2017.
- [8] P. Chopra, R. K. Sharma, and M. Kumar, "Prediction of compressive strength of concrete using artificial neural network and genetic programming," *Advances in Materials Science and Engineering*, vol. 2016, Article ID 7648467, 10 pages, 2016.
- [9] Q. Cheng, Z. Qi, G. Zhang, Y. Zhao, B. Sun, and P. Gu, "Robust modelling and prediction of thermally induced positional error based on grey rough set theory and neural networks," *International Journal of Advanced Manufacturing Technology*, vol. 83, no. 5–8, pp. 753–764, 2016.
- [10] J. Zheng and Z. Song, "Semisupervised learning for probabilistic partial least squares regression model and soft sensor application," *Journal of Process Control*, vol. 64, pp. 123–131, 2018.
- [11] K. Y. Bae, H. S. Jang, and D. K. Sung, "Hourly solar irradiance prediction based on support vector machine and its error analysis," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 935–945, 2017.
- [12] Z. Ju and H. Gu, "Predicting pupylation sites in prokaryotic proteins using semi-supervised self-training support vector machine algorithm," *Analytical Biochemistry*, vol. 507, pp. 1–6, 2016.
- [13] L. Zago, P. Y. Hervé, R. Genuey et al., "Predicting hemispheric dominance for language production in healthy individuals using support vector machine," *Human Brain Mapping*, vol. 38, no. 12, pp. 5871–5889, 2017.
- [14] R. Grbić, D. Slišković, and P. Kadlec, "Adaptive soft sensor for online prediction and process monitoring based on a mixture of Gaussian process models," *Computers & Chemical Engineering*, vol. 58, no. 22, pp. 84–97, 2013.

- [15] M. Zhang and X. Liu, "A soft sensor based on adaptive fuzzy neural network and support vector regression for industrial melt index prediction," *Chemometrics and Intelligent Laboratory Systems*, vol. 126, no. 126, pp. 83–90, 2013.
- [16] M. Ławryńczuk, "Computationally efficient model predictive control algorithms," *Studies in Systems, Decision and Control*, vol. 3, no. 1, pp. 137–149, 2014.
- [17] J. Sun, Z. Cheng, R. Yang, and S. Shan, "Soft-sensor modeling for paper mill effluent COD based on PCA-PSO-LSSVM," *Computers & Applied Chemistry*, vol. 34, no. 9, pp. 706–710, 2017.
- [18] A. Najafi, A. Joudaki, and E. Fatemizadeh, "Nonlinear dimensionality reduction via path-based isometric mapping," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 7, pp. 1452–1464, 2016.
- [19] J. Wang, "Laplacian Eigenmaps," in *Geometric Structure of High-Dimensional Data and Dimensionality Reduction*, pp. 235–247, Springer, 2012.
- [20] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2014.
- [21] Y. Huang, L. Jin, H. S. Zhao, and X. Y. Huang, "Fuzzy neural network and LLE algorithm for forecasting precipitation in tropical cyclones: comparisons with interpolation method by ECMWF and stepwise regression method," *Natural Hazards*, vol. 91, no. 1, pp. 201–220, 2018.
- [22] X. Wang, S. Chen, and M. Yao, "Data dimensionality reduction method of semi-supervised isometric mapping based on regularization," *Journal of Electronics & Information Technology*, vol. 33, no. 1, pp. 241–245, 2016.
- [23] G. Kanishka and T. I. Eldho, "Watershed classification using Iso-map technique and hydrometeorological attributes," *Journal of Hydrologic Engineering*, vol. 22, no. 10, article 04017040, 2017.
- [24] G. Ren, Y. Cao, S. Wen, T. Huang, and Z. Zeng, "A modified Elman neural network with a new learning rate scheme," *Neurocomputing*, vol. 286, pp. 11–18, 2018.
- [25] K. Kolanowski, A. Świetlicka, R. Kapela, J. Pochmara, and A. Rybarczyk, "Multisensor data fusion using Elman neural networks," *Applied Mathematics and Computation*, vol. 319, pp. 236–244, 2018.
- [26] Y. Ge, Y. Huang, D. Hao, G. Li, and H. Li, "An indicated torque estimation method based on the Elman neural network for a turbocharged diesel engine," *Proceedings of the Institution of Mechanical Engineers Part D Journal of Automobile Engineering*, vol. 230, no. 10, pp. 1299–1313, 2016.
- [27] A. Wysocki and M. Ławryńczuk, "Elman neural network for modeling and predictive control of delayed dynamic systems," *Archives of Control Sciences*, vol. 26, no. 1, pp. 117–142, 2016.
- [28] H. Choi and S. Choi, "Robust kernel Isomap," *Pattern Recognition*, vol. 40, no. 3, pp. 853–862, 2007.
- [29] Y. Zhang, Y. Fu, Z. Wang, and L. Feng, "Fault detection based on modified kernel semi-supervised locally linear embedding," *IEEE Access*, vol. 6, pp. 479–487, 2018.
- [30] B. Jiang, C. Ding, and B. Luo, "Robust data representation using locally linear embedding guided PCA," *Neurocomputing*, vol. 275, pp. 523–532, 2018.
- [31] Y. Qianli and Q. Liyun, "A diagnosis method of rotary kiln lining based on extreme learning machine (ELM) and its application," *Sintering and Pelletizing*, vol. 40, no. 6, pp. 18–22, 2015.
- [32] G. Zhang and J. Zhang, "A novel RS-SVM dynamic prediction approach and its application," *Process Automation Instrumentation*, vol. 26, no. 2, pp. 17–20, 2005.



Hindawi

Submit your manuscripts at
www.hindawi.com

