Applied Thermal Engineering 130 (2018) 745–753

Contents lists available at ScienceDirect

Applied Thermal Engineering

journal homepage: www.elsevier.com/locate/apthermeng

Research Paper

An improved extreme learning machine integrated with nonlinear principal components and its application to modeling complex chemical processes

Qun-Xiong Zhu, Xiao Wang, Yan-Lin He *, Yuan Xu *

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China
Engineering Research Center of Intelligent PSE, Ministry of Education of China, Beijing 100029, China

HIGHLIGHTS

- An improved ELM with a special structure is presented.
- ITNN is used to extract the nonlinear principal components (NPCs) from inputs.
- Two independent input subnets based on the correlation coefficient are built.
- A improved ELM integrated with the extracted NPCs (NPCs-IELM) is proposed.
- The effectiveness of NPCs-IELM is validated by modeling a complex chemical process.

ARTICLE INFO

Article history:
Received 6 October 2017
Accepted 12 November 2017

Keywords:
Extreme learning machine
Input training neural network
Correlation coefficient analysis
Modeling
Chemical processes

ABSTRACT

In order to enhance the performance of extreme learning machine (ELM) in modeling complex chemical processes, an improved ELM integrated with nonlinear principal components is proposed. Firstly, an improved ELM (IELM) model is presented. The IELM has a special structure with two independent input subnets: a positive correlation subnet and a negative correlation subnet. The two independent input subnets are developed based on the correlation coefficient between input attributes and output attributes. The nonlinear principal components of original input attributes are extracted using input training neural network (ITNN). The extracted nonlinear principal components are connected to output layer nodes. Thus, the output nodes not only connect with the positive correlation subnet and the negative correlation subnet, but also with the extracted nonlinear principal components. Thus, an IELM integrated with nonlinear principal components (NPCs-IELM) model can be built. The effectiveness of the proposed NPCs-IELM is verified by modeling a high density polyethylene process. Simulation results indicate that the proposed NPCs-IELM can achieve higher accuracy and better stability.

1. Introduction

Modern chemical plants are complicated and composed of a large number of integrated and interdependent process units [1,2]. There are two challenges that should be handled to ensure high efficiency of operation and high-qualified product. Firstly, the increasing complexity of processes requires an increasing accuracy of system description by using modeling technologies. Secondly, integrated and interdependent chemical units lead to high dimensional input patterns. For the first challenge, artificial neural networks (ANNs) are promising methods. ANNs have been widely used in many fields, such as process modeling [3], prediction [4] and optimization [5]. Conventionally, the gradient-based learning algorithms are relatively slow in the learning phase and may easily get stunk in local minimum. In order to solve this problem, training feedforward neural networks with random weights is an alternative method [6,7]. Some randomized neural networks like random vector functional-link net (RVFL), have been discussed [8,9]. Recently, random weights based neural networks have attracted more and more attention. Among the random weights based neural networks, a single-hidden layer feed-forward neural network (SLFN) named extreme learning machine (ELM) was proposed by Huang et al. [10]. In ELM, the input weights and the biases of the hidden layer nodes are randomly assigned, and the output weights

* Corresponding author.
E-mail addresses: heyl@mail.buct.edu.cn (Y.-L. He), xuyuan@mail.buct.edu.cn (Y. Xu).

https://doi.org/10.1016/j.applthermaleng.2017.11.061
1359-4311/

are analytically determined by using generalized inverse. ELM has also been widely applied in many fields, such as regression [11], classification [12], approximation [13], and so on.

Some enhanced ELM models have been proposed by Huang et al. [14,15], Deng et al. [16], He et al. [17], Sun [18] and Samet et al. [19]. For most of the above improved algorithms, the improved networks ignore the effect of the input attributes on the outputs. Some of the input attributes may have a positive effect on the outputs and some other may have a negative effect. In our previous work, an improved extreme learning machine (IELM) has been proposed to handle this problem [20]. The IELM has a special structure with two separated input subnets. The two subnets are built according to the correlation coefficient between input attributes and output attributes. However, the original inputs are directly connected to the output nodes, which may limit the performance.

The second challenge is that the integrated and interdependent of plants require more variables to be monitored due to the high nonlinearity problem. Many researchers have been devoted to solve this problem [21–23]. In the paper of He et al. [24], the original inputs are connected to the output nodes. However, there is highly nonlinear relationship between original input patterns and output patterns. So the performance of IELM model is still limited due to the absence of consideration about the nonlinear relationship between original inputs and outputs. In this work, we further improve the generalization performance by extracting the nonlinear principal components information and then directly connect the extracted nonlinear information to output nodes. Thus, an improved ELM (IELM) integrated with nonlinear principal components extraction method of ITNN is proposed for improving the performance of ELM when dealing with the actual highly nonlinear chemical processes in this paper. There are many methods could be used to extract components. The best-known method to reduce data dimensionality is PCA depicted in Qi et al. [25]. However, the linear PCA literally cannot efficiently deal with nonlinearly correlated variables. Discussed in Zhu et al. [26], Geng et al. [27] and Reddy et al. [28], the nonlinear PCA methods, like principal curves or principal surface, auto-associative neural network (AANN) and ITNN, can be adopted to overcome such a bottleneck. However, the principal curves cannot commonly be applied to all nonlinear systems and the AANN has a relatively complicated net structure compared with ITNN. ITNN is composed of an input layer, a de-mapping layer, and an output layer, which is easier to train than AANN. Based on these facts, ITNN is selected to extract the nonlinear principal components.

The extracted nonlinear principal components are utilized as the direct input attributes linked with the output layer nodes. Finally, an IELM integrated with nonlinear principal components (NPCs-IELM) is proposed. Different from the traditional ELM model, the proposed model has two salient features. Firstly, the model has two independent input subnets: a positive correlation subnet and a negative correlation subnet. The two input subnets are established according to the correlation coefficient between input attributes and output attributes. Secondly, output nodes not only receive the information from the hidden nodes but also directly receive the information of the extracted nonlinear principal components. In order to verify the effectiveness of the proposed model, a complex chemical processes named high density polyethylene (HDPE) is selected.

The remaining parts of this paper is organized as follows: Section 2 provides some preliminaries about methods of traditional extreme learning machine, input training neural network and correlation coefficient analysis; details of the learning algorithm and the construction steps of the proposed NPCs-IELM are described in Section 3; Section 4 presents the case study using the high density polyethylene process; Section 5 contains conclusions.

2. Preliminaries

A brief review of traditional ELM, ITNN and the correlation coefficient analysis is presented in this section.

2.1. Extreme learning machine

For p arbitrary training samples $U = (x_i, t_i) \in [R^m \times R^s](i = 1, 2, \cdots, p)$, the output of the ELM with k hidden layer nodes is shown as follows:

$$\Theta(x_i) = \sum_{k=1}^{K} \beta_k g(A_k \cdot x_i + b_k) \tag{1}$$

where $A_k = [a_{k1}, a_{k2}, \cdots, a_{km}]^T$ is the weight vector connecting the input nodes to the kth hidden node; b_k is the threshold of the k_{th} hidden node; $\beta_k = [\beta_{k1}, \beta_{k2}, \cdots, \beta_{ks}]^T$ is the weight vector connecting the k_{th} hidden node and output nodes; $A_k \cdot x_i$ denotes the inner product of A_k and x_i; $g(\cdot)$ is the activation function of hidden layer nodes.

For all P samples, an equivalent compact form of Eq. (1) can be written as follows:

$$H\beta = t \tag{2}$$

$$H = \begin{bmatrix} g(A_1 \cdot x_1 + b_1) & \cdots & g(A_k \cdot x_1 + b_k) \\ \vdots & \vdots & \vdots \\ g(A_1 \cdot x_p + b_1) & \cdots & g(A_k \cdot x_p + b_k) \end{bmatrix}_{p \times k}$$

$$\beta = [\beta_1^T, \ldots, \beta_k^T]_{k \times s}^T, t = [t_1^T, \ldots, t_s^T]_{p \times s}^T \tag{3}$$

where H is called as hidden layer output matrix; β is the weight vector connecting the hidden nodes and output nodes; t is the desired output value.

In ELM, the learning parameters A_k and $b_k, k = 1, \cdots, K$ are randomly generated. The output nodes are assigned with a linear sum function. Thus, the output weights can be analytically determined by finding a least-square solution as follows:

$$\hat{\beta} = H^{\dagger} T \tag{4}$$

where $H^{\dagger}$ is the Moore Penrose generalized inverse of the hidden layer output matrix H. The topological structure of ELM is shown in Fig. 1.

2.2. Input training neural network

Unlike other feed-forward neural networks, the inputs of the ITNN subnet are not given. Both the weights of ITNN and the input values reproducing the data are adjusted as accurately as possible. For each input vector, it is adjusted to minimize the error between the corresponding output of ITNN and the expected value. After the subnet and the inputs are properly adjusted, we can gain a matrix and a de-mapping neural network model. The matrix X can be viewed as nonlinear principal components. Thus the reduction of high-dimensional data can be fulfilled through training a de-mapping network and simultaneously the input layer. The structure of ITNN is presented in Fig. 2.

The algorithm of ITNN is illustrated as follows. Supposed that N arbitrary training samples $\{(x_i, t_i)\}_{i=1}^N$ are available, where x_i consisting of m elements is the m -dimensional vector of the ith sample, and t_i consisting of n elements is the n -dimension vector of the ith output. Each input vector $[\xi_{N1}, \xi_{N2}, \cdots, \xi_{Np}]$, where the number of the samples is N, and p is the number of the extracted nonlinear principal components, is adjusted to minimize the error between