International Journal of u- and e- Service, Science and Technology
Vol.9, No. 12 (2016), pp.442-450
http://dx.doi.org/10.14257/ijunesst.2016.9.12.39

Study of Boiler NOx Emission Model Based on Improved Deep Learning and Genetic Algorithm

Mingzhu LU[1,2*], Jianhua GANG[1], Haiyi SUN[3] and Wei ZHENG[2,4]
[1]*Department of mechanical and electrical engineering, Cangzhou Normal University, Hebei 061001, China;*
[2]*School of electrical engineering and automation, Tianjin University, Tianjin 300072, China;*
[3]*Cangzhou bureau of traffic and transportation, Hebei 061001, China;*
[4]*School of Mechatronical Engineering and Automation Tianjin Vocational Institute, Tianjin 300410, China;*

Abstract

Boiler efficiency and emission load of NO_X are the key evaluation indicators of operation performance of the coal-fired boiler. It has become a popular academic research topic concerning the reduction of NO_X emissions while maintaining the same boiler efficiency, as well as how to build a model for boiler emission. Based on a prediction model that is constructed for boiler efficiency and emission load of NO_X with the application of deep belief algorithm, genetic algorithm is used to optimize the tilting angel of boiler burners and the flow velocity of pulverized coal, thereby effectively reducing the emission load of NO_X. Simulation results indicate that this method effectively optimizes the parameters of the boiler, and provides a new way to optimize the parameters of the boiler.

Keywords: *DBN network; information entropy; prediction; Genetic Algorithm (GA); NO_X*

1. Introduction

Boiler efficiency and emission load of NO_X are important evaluation indicators of the coal-fired boiler. To reduce the emission load of NO_X, it is of great significance to build an accurate model for boiler emissions in the case of maintaining the same boiler efficiency. However, the emission characteristics and efficiency characteristics of the coal-fired boiler are relatively complex. There exist a number of influencing factors, such as the type of coal, load, flow velocity of coal and boiler wind distribution, and the coupling effect occurs among these factors. Therefore, it is difficult to build a relatively accurate model for boiler emissions. Usually, the impact of each influencing factor on the indicators is evaluated by the method of practical test. However, the impact mechanism of each factor on the indicators varies, and there exists the effect of mutual overlapping and amplification among different factors. As a result, it is difficult for the common prediction methods, such as partial least squares, neural network and support vector machine, to obtain favorably accurate prediction on above-mentioned indicators.

The data is collected from the SG-1025/17.5-M897 subcritical natural circulation boiler produced by a power plant in Shanghai Boiler Factory from February 1, 2015 to February 9, 2015. A total of 12,960 sets of sample data in terms of influencing factors and evaluation indicators have been collected at the interval of one minute. The coefficients acquired after the partial least squares regression are employed to evaluate the impact of 16 influencing factors on boiler combustion efficiency.

[*] Corresponding Author

ISSN: 2005-4246 IJUNESST

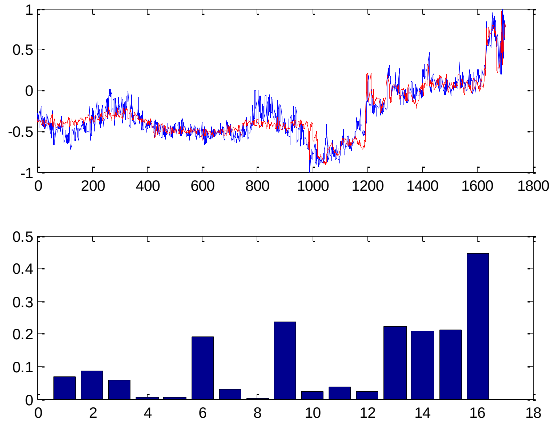

Figure 1. Test Chart of Partial Least Squares

Compared with the common prediction algorithms, by virtue of its characteristics of successive extraction and step-by-step learning, deep belief is able to accurately predict the unknown objects in accord with object features. On the basis of the study on the DBN network, this paper applies information entropy and T-type RBM network to improve the DBN network. A prediction model is constructed for boiler efficiency and emission load of NO_X with the application of deep belief algorithm, and genetic algorithm is used to optimize the tilting angel of boiler burners and the flow velocity of pulverized coal, thereby reducing the emission load of NO_X while maintaining the same boiler efficiency. Consequently, a new train of thought is provided to optimize the parameters of the boiler.

2. DBN Network

2.1. RBM Network

Restricted Boltzmann Machine (RBM) is the basic component unit of the DBN network. A typical RBM network is made up with a visible layer at the lower, a visible layer at the upper, and a hidden layer at the lower, as shown in Figure 2 [1-2].

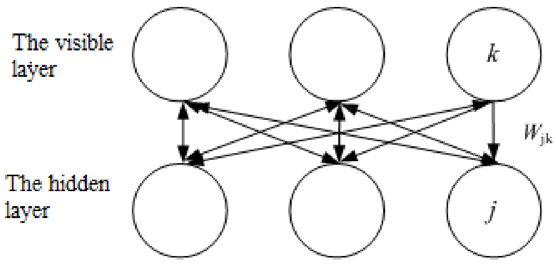

Figure 2. RBM Network

According to Figure 2, the visible layer and the hidden layer of the RBM network are interconnected. Weight W is used to represent the strength of connection. Units in the visible layer and in the hidden layer contain paranoid units, which are demonstrated as the state of activation and the state of close in the form of probability.

The basic principle of RBM is from thermodynamics in physics. The classic thermodynamics believes that the energy state of the particle is inversely proportional to the probability. The particles with lower energy hold a higher probability of occurrence, while the particles with higher energy hold a lower probability of occurrence. The energy in the RBM network is defined as the following Formula (1).

$$E(v,h|\theta) = -\sum_i \sum_j W_{i,j} v_i h_j \quad (1)$$

According to the principle of thermodynamics, the formula of energy conversion probability is illustrated in Formula (2).

$$P_i = 1/Z * e^{-E_i} \quad (2)$$

In terms of the RMB network, each pair of units in the visible layer and in the hidden layer is required to be calculated twice. Thus, for the RBM network that includes m units in the visible layer and n units in the hidden layer, the total number of the probability state is 2^{n+m}.

Since the RBM network includes units in the hidden layer and in the visible layer, the units within the layer are not connected. In terms of probability, the units maintain a state of relative independence. As for unit k in the hidden layer, its probability of activation depends on the activation state and connection weight W of all the units in the visible layer that are linked with units in this hidden layer. Its probability of activation is calculated in Formula (3) as below.

$$P(H_k = 1|V,W) = sigmoid(\sum V_i W_{ij}) \quad (3)$$

Wherein, sigmoid function is a step function, whose characteristics is that it approaches 0 in the direction of the large negative, and it is close to 1 in the direction of the large positive.

2.2. DBN Network

The so-called deep belief refers to a type of machine learning classification algorithm that conducts feature extraction for data through the process of multi-layer feature extraction [3]. Deep belief holds its advantages such as multilayer network depth and feature learning, and enjoys better feature expression capabilities than the conventional shallow networks. Deep Belief Network (DBN) is the most widely-used network in the research and in the application of in-depth learning, which consists of multiple RBM networks and BP networks. During training, the DBN network first employs the unsupervised learning method to train the RBM network step by step, thereby preserving feature information through layer-by-layer mapping of the original input data. Then, with the application of the supervised method, the output of the last layer of the RBM network is used to train the entire network. The structure of DBN network including a 2-layer RBM network and a 1-layer BP network is illustrated in Figure 3 [4].

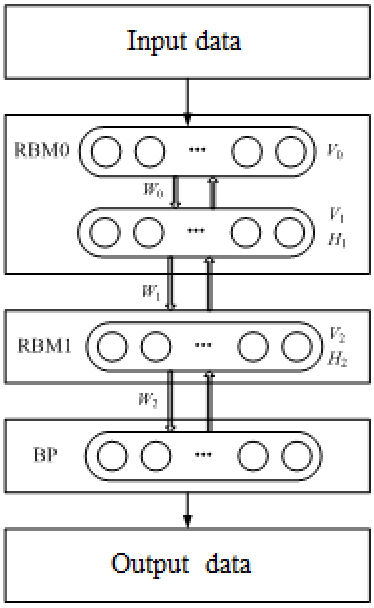

Figure 3. The structure of DBN network

In Figure 3, *V* and *H* respectively represent the node values of the visible layer and the hidden layer of the RBM network, and *W* stands for the value of connection weight. Since the RBM network is symmetrical, the node values of the visible layer is calculated by the node values of the hidden layer, as shown in Formula（4）.

$$p(v_i = 1) = \frac{1}{1 + e^{-c_i - \sum_j H_j W_{ji}}} \tag{4}$$

To be specific, c_i is the parameter, H_j is the node value, and w_{ji} is the weight.

The joint probability distribution of the feature vector V in the visible layer and the feature vector H in the hidden layer is shown in Formula（5）.

$$p(V,H) \infty \exp(-E(V,H)) = e^{H^T WV + b^T V + c^T H} \tag{5}$$

Wherein, W is the weighted value between the visible layer and the hidden layer. $E(V,H)$ is the mathematical expectation of feature vector V and H. The purpose of network training is to solve $\theta = (W,b,c)$ and to maximize the joint probability distribution $p(V,H)$ in Formula (5). The conventional method is the Markov Chain Monte Carlo Method (MCMC). In fact, the $p(V,H)$ and the terminal joint probability distribution $p(V_i^\infty, H_j^\infty)$ obtained by the Markov chain method are difficult to guarantee the convergence. In the experiment, the CD guideline is used to enhance the computing speed and to guarantee the computing accuracy.

2.3. Improvements of the DBN Network

For DBN network, favorable feature extraction capabilities are the key to construct a network with preferable prediction capabilities. The DBN network extracts the characteristics of input data layer by layer, so the feature extraction capabilities for each layer of the RBM network exert relatively great impact on the final results [5-6].

The classic DBN network has been able to carry out favorable feature extraction layer by layer; however, in this paper, it is believed that this network still has room for improvement. First, the continuous process of extracting data feature tends to be inevitably more and more simplified, which enables a more simplified and more efficient expression of data nature. Furthermore, the size of data feature can be predicted by a number of approaches. If some algorithms are able to automatically predict the size of data feature, the number of units in the hidden layer could be advised, thereby achieving the purpose of automatic adjustment of parameters.

Based on the meaning of naming deep learning, the process of feature extraction should be carried out layer by layer. At the same time, deep learning is able to further abstract the information at each layer. Therefore, the data in the final results should be featured with a high level of abstraction and simplification. To achieve this, this paper designs the T-RBM network, and the network structure is shown in Figure 4.

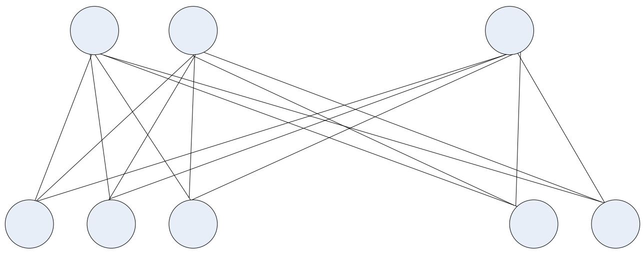

Figure 4 T-RBM Network Structure

According to Figure 4, The T-RBM network applies trapezoidal node structure, and there are fewer nodes in the hidden layer. Thus, the conversion of data information from the bottom network to the top network is able to be simplified. In this way, it not only favorably narrows down the data problem, but also effectively avoids the over-fitting phenomenon.

The number of nodes in the hidden layer of the T-RBM network is relatively small. But if the node value of the hidden layer is fixed, and feature extraction capabilities are still favorable on the basis of reduced feature dimension, it is a key issue to enhance the performance of the DBN network [7].

Information entropy is a commonly used method in the theory of information, which is applied to present the proportion of useful information to the data. In terms of the coal-fired boiler, its useful information is the degree of how much each factor deviates from the boiler at the stable state. Thus, the entropy function of the T-RBM Network is calculated as Formula（6）.

$$H(V) = E(\log_2(1/p(V_i)) = -\sum P(V_i)\log_2(p(V_i)) \quad (i = 1, 2, ..., n)$$

(6)

To be specific, $p(V_i)$ represents the probability that the i-th factor deviates from the stable value. $H(V)$ stands for the value of the entropy.

Concerning the factor of the boiler, the greater the current operation deviates from the stable value, the larger the information is included in the input elements. The number of nodes in the hidden layer is proportional with the entropy value, as calculated in Formula (7).

$$C(V) = Alpha * H(V)$$

(7)

$C(V)$ represents the number of the units in the hidden layer, and $Alpha$ is the fixed parameter, which equals 100 in most cases.

3. Output Forecast of Coal-Fired Boiler

3.1. Coal-Fired Boiler and its Influencing Factors

Data in this paper is based on the SG-1025/17.5-M897 subcritical natural circulation boiler produced by Shanghai Boiler Factory, with features like single furnace and tangential combustion. The coal pulverizing system is positive pressure direct firing, with five roller coal pulverizers. Boilers of this kind hold advantages like small size, high thermal cycle efficiency and long service life, which have been widely used in various types of factories. For this type of boiler, its evaluation indexes—boiler efficiency and emission load of NO_X are mainly impacted by factors including primary air, secondary air, over fire air, SOFA burners, the flow velocity of coal, the total amount of coal, the tilting angel of burners, load, and the coal quality coefficient. From February 1, 2015 to February 9, 2015, influencing factors and evaluation indicators of this boiler have been collected at the interval of one minute. A total of 12,960 sets of sample data have been collected, and 16-dimensional input data and 2-dimensional output data are constructed in the model based on sample data.

3.2. Index Prediction and Analysis

First, based on the influencing factors of normalization, the method discussed in 1.2 is applied to estimate the number of nodes in the hidden layer. The information entropy formula is used to calculate the maximum value of information entropy among the input data—1.6, resulting in 160 nodes in the hidden layer. On this basis, the DBN network consisting of a two-layer T-RBM network and a one-layer BP network is constructed. The first 11460 sets of sample data are employed to train the DBN network, and then the last 500 sets of data are used to test the prediction capabilities of the trained DBN network. The forecasting results are shown in Figure 4.

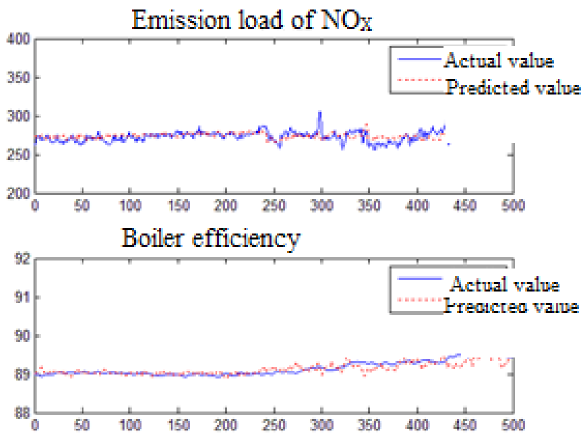

Figure 4. Prediction Results

Results of the prediction accuracy of the improved DBN network and other prediction methods are compared, and the comparative results are shown in Table 1.

Table 1. Comparison of Different Prediction Methods

Prediction Methods	Emission load of NO_X		Boiler efficiency	
	Relative error	Error variance	Relative error	Error variance
Partial least squares	7.41%	10.52	7.88%	15.78
Neural network	6.32%	9.71	7.12%	16.12

Support vector machine	6.01%	8.91	5.73%	14.31
Deep Belief	4.93%	8.81	5.16%	13.78
Improved Deep Belief	4.19%	8.15	4.88%	11.21

As seen from Figure 4 and Table 1, the improved deep belief algorithm proposed in this paper is able to relatively accurately predict the emission load of NO_X and boiler efficiency. Compared with the several commonly-used prediction methods, the prediction results of this method indicate relatively smaller relative errors and error variances, which could be used to optimize input parameters in the genetic algorithm.

4. Optimization Parameters of Genetic Algorithm

4.1. Overview of the Algorithm

Genetic Algorithm (GA) originates from the human observation of biological evolution and genetic phenomenon in the nature. This algorithm is a global optimization algorithm with parallel computing capabilities and high search efficiency, which is able to locate the optimal solution of the problem in a short time. The idea of this algorithm was first proposed by Professor John Holland from the University of Michigan. Goldberg generalized and summarized genetics through the combination of computer technology and mathematical theory, and in the 1980s, he proposed the integrated genetic algorithm theory and implementation method. GA adopts the technique of probabilistic search. On the basis of population initialization, new individuals are constantly generated through selection, crossover, mutation and other operations, and the entire population is directed and evolved towards the direction of the optimal solution through fitness function. The advantage of this method is the high efficiency of algorithm search, the strong capability of searching the globally optimal solution, the favorable universality, and a certain degree of parallelism and robustness.

Genetic algorithm is an optimal search method that simulates biological evolution. Based on the setting of fitness function, this method initializes a potential solution set of the problem. Each individual in the solution set represents the potential solution of the problem. Then by selecting between crossover operation and mutation operation, the population moves in the direction of the optimal solution until the optimal location of the problem is identified. Genetic algorithm is used to optimize the input evaluation index data of the boiler, so as to reduce the emission load of NO_X in case of maintaining the same boiler efficiency. In terms of the practical use of boilers, the flow velocity of coal and the tilting angel of burners are two commonly-used parameters of adjusting the boiler. Therefore, genetic algorithm is employed to optimize these two parameters.

4.2. Algorithm Settings

(1) Individual Coding

Optimized parameters are continuously adjustable, so the individual coding applies real number encoding.

(2) Fitness Function

According to the goal of algorithm optimization, the fitness function is set as Formula (8) below.

$$fit = efficiency_{DBN} \text{-} efficiency_{real} + NO_{real} - NO_{DBN} \tag{8}$$

Specifically, $efficiency_{DBN}$ and NO_{DBN} represent the prediction values obtained by the DBN network when the current boiler inputs the parameters. $efficiency_{real}$ and NO_{real} indicate the actual values. The larger the value of fitness, the better.

(3) Genetic Manipulation

Genetic manipulation mainly includes selection operation, crossover operation and mutation operation.

The selection operation applies the roulette wheel selection algorithm. In other words, the greater the value of individual fitness, the greater the probability of the individual being selected. As a single point crossover operation, the crossover operation first randomly selects crossover individuals and crossover positions from the population, and then the crossover is realized by means of swapping genes in the corresponding positions. The mutation operation employs the method of real number variation. Mutation individuals and mutation positions are first randomly selected from the population, and new individuals are obtained by random variation.

4.3. Results of Optimization

Genetic algorithm is used to optimize the parameters at a given state of the coal-fired boiler. In the case that the boiler efficiency remains unchanged, the emission load of NO_X is reduced. The state parameters under the circumstances that the boiler is in stable operation are demonstrated in Table 2.

Table 2. Operating Parameters of the Boiler

Primary air	22.49 9.68 -0.24 -0.24 -0.21
Secondary air	29.26 7.32 -0.30 31.04
Overfire air	99.28 98.41 98.26 97.80
Total amount of coal	136.71
Flow velocity of coal	27.25
Tilting angel of burners	49.93
Boiler efficiency	90.09
Emission load of NO_X	197.97

The parameter settings of genetic algorithm are as follows: the number of the population is 20, the number of iteration is 100, the crossover probability is 0.1, and the mutation probability is 0.1. The optimization process of genetic algorithm is illustrated in Figure 5.

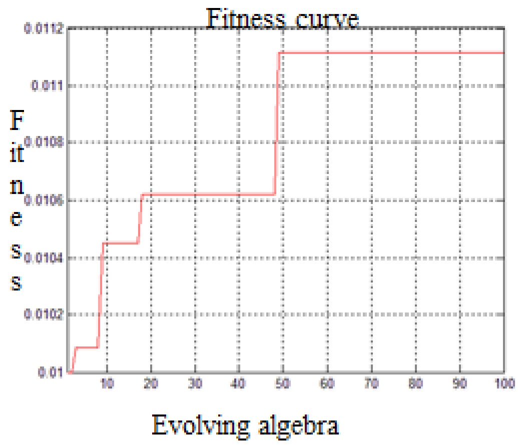

Figure 5. Genetic Algorithm Optimization

Under genetic algorithm, the optimized flow velocity of coal is 28.43, and the tilting angel of burner is 51.51. When the boiler efficiency is not reduced and other parameters remain unchanged, the emission load of NO_X is reduced from 197.97 to 196.91.

5. Concluding Remarks

The process of coal-fired boiler combustion is a complex process, with problems like strong nonlinearity and multiple relevant factors. The current commonly used methods are difficult to predict the emission load of NO_X and the boiler efficiency on the basis of the values of the current influencing factors of the coal-fired boiler. In this paper, the DBN network is adopted to construct a index prediction model for the coal-fired boiler. The existing DBN network is improved through T-RBM network and information entropy, so that the DBN network has stronger prediction capabilities. On the basis of the network construction, genetic algorithm is used to optimize the tilting angel of burner and the flow velocity of coal, thereby effectively reducing the emission load of NO_X on the premise of not reducing boiler efficiency. In this way, a new approach is offered to optimize the parameters of the boiler.

Acknowledgements

Fund Project: Application of variable selection in boiler combustion optimization (project number: 15210336)

References

[1] G. Hinton, “A practical guide to training restricted Boltzmann machines”, Momentum, vol. 9, no. 1, **(2010)**, pp. 926.
[2] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks”, Science, vol. 313, no. 5786, **(2006)**, pp. 504-507.
[3] Z. Lu, “A Study on DBN Classification Network Based on T-RBM Algorithm”, Jilin University, vol. 5, no. 13, **(2014)**.
[4] Y. Bengio and D. Olivier, “On the expressive power of deep architectures”, Algorithmic Learning Theory. Berlin: Springer /Heidelberg, **(2011)**, pp. 18-36.
[5] R. Salakhutdinov and G. Hinton, “An efficient learning procedure for deep Boltzmann machines”, Neural Computation, vol. 24, no. 8, **(2012)**, pp. 1967–2006.
[6] J. W. Liu, Y. Liu and X. L. Luo, “Research Progress of Boltzmann Machine”, Computer Research and Development, vol. 51, no. 1, **(2014)**, pp.1-16.
[7] T. Tieleman, “Training Restricted Boltzmann Machines Using Approximations to the Likelihood gradient”, Proceedings of the 25th international conference on Machine learning. Canada: ICML, , **(2008)**, pp. 1064-1071.
[8] T. T. Yang, “A Data-based Study on Energy Optimization Control of Power Plant”, North China Electric Power University, vol. 8, **(2010)**, pp. 47-55.
[9] F. Gao, “A Study on the Modeling and the Optimization of Large-capability Boiler Combustion System Based on Intelligent Computing”, North China Electric Power University, vol. 6 **(2013)** , pp. 52-56
[10] J. L. Wang and E. J. Li, “Boiler Efficiency of 600MW Coal-fired Power Station and Multi-objective Optimization of NO_X Emission Concentration”, Power System Engineering, vol. 5, no. 3, **(2015)** pp. 47-48.

