Prediction study of temperature deviations on left and right sides of power plant boiler reheaters based on regression algorithm learning machine and expert experience

Yu-jia Cui1, Ying Huang, Xiao-chun Ma, Guo-hang Ma
Harbin Boiler Factory Company Limited, Xiangfang District Haerbin 150040, China;
1 Email: 2419724044@qq.com

Abstract. Based on regression algorithm learning machine and expert experience, the temperature deviation prediction of the left and right sides of the power plant boiler reheater is based on real-time or off-line data collected by the power plant boiler. The data is divided by expert experience and algorithm processing data. Based on the machine learning regression algorithm, a regression algorithm learning machine is established. Through this learning machine and expert experience, the temperature deviation of the left and right sides of the power plant boiler reheater is predicted. The results show that the model can accurately predict the temperature deviation of the left and right sides of the power plant boiler reheater, thus providing reference guidance for the operation of the power plant boiler.

1. Introduction
Driven by economic benefits and environmental protection requirements, power plants need to continuously improve unit efficiency and reduce pollutant emissions to improve competitiveness. [1] The boiler is one of the three major components of the thermal power unit, and ensuring the stability of its operation is of vital importance for improving the efficiency of the thermal power unit. Due to the diversity of boiler equipment composition, changes in the external environment, aging of boiler equipment and other factors, the possibility of temperature deviations on the left and right sides of the reheater is increasing.

In order to ensure the reliability and safety of power station boilers and improve service quality, it is urgent to predict the temperature of the left and right sides of the reheater. [2] Use power generation enterprise control system, plant-level monitoring system (SIS) and various information systems are used to accumulate massive historical data of equipment operation, and use artificial intelligence methods such as big data calculation models, machine learning theory and solution reference to provide remote diagnosis of temperature deviations on the left and right sides of the reheater. [3]

Based on this, first a large amount of discrete data is collected, and the data is processed through mathematical methods and expert experience. A regression algorithm learning machine is established to predict the temperature deviation of the left and right sides of the power plant boiler reheater, compare the evaluation indicators, and find the best prediction method. When the actual temperature exceeds expectations, the system alarms and measures are taken in advance, which can effectively reduce the risk of power plant boiler failure, significantly improve the economics and stability of power plant boiler operation, and provide guidance for the operation of power plant boilers.
2. Data preprocessing
The data collected at the power plant is large in number, high in dimensionality, and redundant. Before forecasting, the data must be integrated, and the required data must be extracted for subsequent model establishment and analysis.

2.1. Data filtering
One-year operation data of a power plant is selected. First delete all non-numerical variable features, and integrate all features at the same time according to time. Through observation, it is found that the values of some features are missing, and the average value of the features is used for filling. Then observe the time series graphs of some of the variables for the entire time period, experts filter, and manually delete abnormal data.

Take the mixed wind temperature at the A mill entrance of the power plant as an example, as shown in Figure 1. It can be seen from the figure that most of the mixed air temperature at the inlet of the A mill is distributed between 250-350. Near 61233, this data is partially missing. At some points, the data is abnormal. Now use the average value to fill in the missing values, and manually delete individual outliers. The distribution of the mixed wind temperature data at the entrance of the A mill after processing is shown in Figure 2. In Figure 1 and Figure 2, the horizontal axis is the serial number of the data, and the vertical axis is the value of the mixed air temperature at the entrance of the A mill.

![Figure 1. Raw data of mixed wind temperature at A mill entrance.](image1)

![Figure 2. Data after mixed air temperature treatment at A mill inlet.](image2)

Determine the influencing factors related to the temperature deviation of the left and right sides of the reheater through expert experience. Determined including main feed water flow, desuperheating water flow, reheater outlet pressure, total coal volume, total air volume, mixed air temperature at each mill inlet, desuperheater inlet and outlet temperature, furnace pressure, reheater outlet flue gas temperature, superheater 121 influencing factors such as outlet flue gas temperature.

2.2. Data processing
The principal component analysis method is used to reduce the dimension of 121 features and eliminate part of the system noise. PCA is a multivariate statistical technique for data compression and data mining. Under the principle of ensuring the minimum loss of data information, the dimensionality reduction process of high-dimensional variable space is performed to eliminate the overlapping of information caused by multiple correlations of variables. The principle is to convert multiple variables into a few comprehensive variables (ie principal components), and the principal components are not related to each other, so that these principal components can reflect most of the information of the
original variables. After passing the principal component analysis method, the feature results are used as the input of the training set of the prediction algorithm.

Due to the different dimensions and units of the data vector output by the simulation, the data standardization analysis is used to preprocess the required data. The formula is as follows:

\[
S_n^k = \frac{X_n^k - \frac{1}{K} \sum_{k=1}^{K} X_n^k}{\sqrt{\frac{1}{K} \sum_{k=1}^{K} (X_n^k - \frac{1}{K} \sum_{k=1}^{K} X_n^k)^2}}
\]  

(1)

In the formula: \(S_n^k\) represents the production parameter of the kth sample under the standardization of n-dimensional data, \(X_n^k\) represents that the kth sample arranges n-dimensional metadata in time series, and K is the number of data sets.[4]

Divide all data into training data and verification data. The input parameters of the model are main feed water flow, desuperheating water flow, reheater outlet pressure, total coal volume, total air volume, mixed air temperature at each mill inlet, desuperheater inlet and outlet temperature, furnace pressure, reheater outlet flue gas temperature, superheat flue gas temperature etc. The temperature deviation of the left and right sides of the reheater is used as the output parameter to establish the model.[5]

3. Establishment of regression algorithm learning machine

The four algorithms of stochastic gradient descent, ridge regression, lasso regression, and support vector regression machine SVR are combined into a regression algorithm learning machine, which predicts the temperature deviation of the left and right sides of the reheater.

Assume that the loss function of the regression model is:

\[
J(\beta) = \sum_{i=1}^{n} (y_i - \sum_{j=1}^{d} x_{ij} \beta_j)^2
\]  

(2)

In the formula, \(i\) is the \(i\) training sample, \(n\) represents the number of instances in the training set, \(y_i\) represents the output of the \(i\) instance, \(x_{ij}\) represents the input of the \(i\) instance, and \(\beta\) is the regression coefficient.[2]

The goal of the regression algorithm is to minimize it. Now we use the following four methods to find the optimal loss function.

3.1. Stochastic gradient descent

Gradient descent uses an iterative form to find the minimum value of the loss function. The update rules are as follows:

\[
\beta_j = \beta_j - \alpha \frac{\partial}{\partial \beta_j} J(\beta)
\]  

(3)

Where: \(\alpha\) is the learning rate.[2]

Bring formula (2) into formula (3) and get:

\[
\beta_j = \beta_j - \alpha \frac{\partial}{\partial \beta_j} (\sum_{i=1}^{n} (y_i - \sum_{j=1}^{d} x_{ij} \beta_j)^2)
\]  

(4)

Iterate and change continuously so that the loss function decreases according to the direction of gradient descent until it reaches the minimum value.

Stochastic gradient descent is to randomly select one of the training data to update the parameters to reduce the complexity of the calculation process and converge faster.
3.2. Ridge regression

Ridge regression is to increase the sum of squares of all parameters on the basis of the loss function formula (2), that is, the L2 norm.

\[ J(\beta) = \sum_{i=1}^{n} (y_i - \sum_{j=1}^{p} x_{ij} \beta_j)^2 + \lambda \sum_{j=0}^{p} \beta_j^2 \]  

(5)

Where: \( \lambda \) is the regular term coefficient. By choosing the appropriate regular term coefficients \( \lambda \), the optimal loss function is obtained. In this prediction, \( \lambda \) take 0.1.

3.3. Lasso regression

Lasso regression adds the sum of the absolute values of all parameters on the basis of the loss function formula 1, that is, the L1 norm, and the loss function becomes:

\[ J(\beta) = \sum_{i=1}^{n} (y_i - \sum_{j=1}^{p} x_{ij} \beta_j)^2 + \lambda \sum_{j=0}^{p} | \beta_j | \]  

(6)

Where: \( \lambda \) is the regular term coefficient. In this prediction, \( \lambda \) takes 0.005. The selection of \( \beta_j \) uses the sub-gradient method to derive the partial derivative of the RSS in the first part of formula (6), select one dimension at a time for optimization and iterate continuously to obtain the optimal regression coefficient.

3.4. Support vector regression machine

The loss function of support vector regression machine SVR is as follows:

\[
\begin{align*}
& y_i - w^T x_i - b \leq \varepsilon \\
& w^T x_i + b - y_i \leq \varepsilon
\end{align*}
\]

(7)

Where: \( \varepsilon \) is the accuracy of the fitting error. On this basis, the slack variables \( \xi_i \) and \( \xi_i^* \) are introduced to control the punishment of the samples beyond the error range[5]. For non-linear SVR, it is necessary to introduce a kernel function, through non-linear mapping to a high-dimensional feature space, to perform linear regression operations.

In this prediction algorithm, the fitting error precision is \( \varepsilon \) and the relaxation variable C is 1. The kernel function is radial basis rbf.

4. Result analysis

The training data is trained by a regression algorithm learning machine to form four different regression algorithm networks.

The input data of the verification set are respectively passed through these four networks to obtain the output result. By comparing the evaluation indexes such as fitting degree and root mean square error, the algorithm with high fitting degree and small error is selected as the final prediction algorithm.

After the data set passes the four regression algorithms of stochastic gradient descent, ridge regression, lasso regression, and support vector regression machine SVR, the results obtained are shown in Figure 3 and Figure 4, respectively.

Figure 3 shows the actual data and prediction data of the validation set after the data passes through the regression algorithm learning machine. It can be seen from the figure that these four algorithms have a good prediction effect on the temperature deviation of the left and right sides of the reheater. The prediction effect of stochastic gradient descent, ridge regression, and lasso regression is not much different. It can accurately predict most data, but it cannot accurately predict the very special values in the data. For SVR, the prediction degree is better than the other three algorithms.

From these data, select 100 consecutive time periods, you can clearly see the prediction effect, the results are shown in Figure 4.
Figure 3-1. Stochastic gradient descent prediction results.

Figure 3-2. Ridge regression prediction results.

Figure 3-3. Lasso regression prediction results.

Figure 3-4. Support vector regression prediction results.

The prediction of the four regression algorithms can be seen more clearly from Figure 4. The specific evaluation indicators are shown in Table 1.

Table 1 shows the fitting degree, MAE, MSE, and RMSE errors of the four algorithms of stochastic gradient descent, ridge regression, lasso regression, and support vector regression.

| Algorithm                     | $R^2$  | MAE    | MSE    | RMSE   |
|-------------------------------|--------|--------|--------|--------|
| Stochastic gradient descent   | 0.9009 | 1.5922 | 4.9379 | 2.2221 |
| Ridge regression              | 0.9054 | 1.5497 | 4.7181 | 2.1721 |
| lasso                         | 0.8991 | 1.6110 | 5.0318 | 2.2432 |
| SVR                           | 0.9314 | 1.2050 | 3.5232 | 1.8770 |
It can be seen from the graph and table that the SVR has the best fitting degree and the MAE, MSE, and RMES errors are the smallest, so that the SVR prediction result can be used as the final output.

According to the experience of experts, the threshold of the temperature deviation of the left and right sides of the reheater is set to $5 \degree C$. When the predicted reheater deviation exceeds $\pm 5 \degree C$, the system alarms. The effect is shown in Figure 5. The temperature deviation of the two sides of the predicted reheater is black within the normal range. When the threshold is exceeded, the system display becomes red.
5. Conclusion

Based on the regression algorithm learning machine and expert experience method, it has a high predictive ability for the temperature deviation of the left and right sides of the power plant boiler reheater. According to the data used in this article, stochastic gradient descent, ridge regression, lasso, and support vector regression machine SVR are the four regression algorithms. The SVR evaluation index is the best and the best solution for prediction.

Through this regression algorithm, the learning machine can clearly and quickly find a more accurate method for predicting the temperature deviation of the left and right sides of the reheater. Compared with using a single algorithm to predict one by one, the prediction time can be shortened, and the prediction efficiency and quality can be significantly improved.

By predicting the temperature deviation of the left and right sides of the reheater, the change can be predicted on the power plant side. If over temperature occurs, it can be alerted in advance to reduce losses and provide guidance for the safe operation of the power plant.

References

[1] Riccardi, A, Fernandez-Navarro, F, Carloni, S 1898-1909 Cost-Sensitive AdaBoost Algorithm for Ordinal Regression Based on Extreme Learning Machine[J] IEEE Transactions on Cybernetics 44(10)

[2] 2016 Application of Extreme Learning Machine Algorithm in the Regression Fitting[C]// International Conference on Information System & Artificial Intelligence. IEEE

[3] Wang J S, Lin C W, Yang Y T C, et al. 2011 A PACE Sensor System with Machine Learning-Based Energy Expenditure Regression Algorithm[C]// Bio-Inspired Computing and Applications-7th International Conference on Intelligent Computing, ICIC 2011, Zhengzhou,China, August 11-14. 2011, Revised Selected Papers. Springer-Verlag

[4] Xu J, Song X . Ant Colony Optimization for Nonlinear Inversion of Rayleigh Waves[C]// Bio-Inspired Computing and Applications - 7th International Conference on Intelligent Computing. ICIC 2011, Zhengzhou,China, August 11-14. 2011, Revised Selected Papers. Springer-Verlag, 2011.

[5] Xu Y, Wang J. 2012 Industrial Process Modeling Based on Online Learning Algorithm for Regression Least Squares Support Vector Machine[J] Advanced Materials Research 472-475 505-509