Detection method of coal low calorific value based on machine learning

Zelin Yan[a], Dong Xiao*[a], Jian Li[b], Hongfei Xie[a], Zhang Li[a]
[a]College of Information Science and Engineering, Northeastern University, Shenyang 110819, China; [b]Unit 31434 of the Chinese People's Liberation Army, Shenyang 110000, China

ABSTRACT

Coal is a major part of the world's energy. The low calorific value of coal is the key index to measure the calorific value of coal. Therefore, real-time detection of low calorific value of coal mine plays an indispensable guiding role in the formulation of resource mining plan. To solve the above problems, this paper proposes a DP_WTELM (D: direct orthogonal signal correction; P: principal component analysis; W: whale optimization algorithm) model based on machine learning to detect the low calorific value of coal. The model is improved based on the classical forward propagation network that is extreme learning machine (ELM). The prediction accuracy of the network is improved by using preprocessing, whale optimization algorithm, and increasing the network depth. The experiments show that in contrast to ELM models, the DP_WTELM model has better performance in predicting the low calorific value of coal and can meet the industrial requirements.

Keywords: Machine learning, extreme learning machine, coal, low calorific value

1. INTRODUCTION

With the increasing maturity of spectral technology, more and more scholars apply it to the fields of ore analysis, grade identification and food qualification. Wang et al.[1] applied principal component analysis and support vector machine to analyze the spectrum of 199 different types of coal samples, proposed a high accuracy and speed coal sample classification method and obtained accurate prediction results. Zeng et al.[2,3] designed an extended Kalman filter to obtain the moisture content of coal and then evaluate the coal quality. Hu et al.[4] proposed a near-infrared spectral analysis method using the property information of organic matter, which can effectively improve the prediction accuracy of coal composition. Wang et al.[5] introduced the confidence learning machine into the near-infrared spectrum and combined it with the support vector machine to build the online and offline coal classification models. The results obtained have great application significance for the on-site coal classification. Ren et al.[6] proposed an all-optical spectroscopy method for identifying bulk commercial coal. This method can quickly and effectively obtain the classification information of coal. Qin et al.[7] proposed a method that combines laser-induced breakdown spectroscopy with Fourier transform infrared spectroscopy. This method can extract both element and molecular information at the same time. The analysis results of caloric value of coal have also been improved. Therefore, the wide application of spectral technology provides an efficient, fast, accurate and economical method for identifying and analyzing coal.

The problem solved in this paper is the prediction of the medium and low calorific value of coal. Firstly, the direct orthogonal signal correction method and principal component analysis method are used to preprocess the infrared spectrum data, which can denoise and simplify the data while retaining the key information. Compared with the ordinary extreme learning machine[8], the two hidden layers extreme learning machine (TELM)[9] can obtain higher accuracy when the number of nodes in the same. Therefore, we use the whale optimization algorithm to optimize the weight and bias of the TELM model and then obtain a coal low calorific value detection and analysis model based on the DP_WTELM algorithm, and we prove the accuracy of the model through simulation experiments.

2. COLLECTION AND PROCESSING OF COAL SPECTRAL DATA

Usually, the spectrum directly collected by the spectrometer is not a smooth curve with a clear peak, but it will contain a lot of "burrs". The reason for these "burrs" is that some information or noise irrelevant to the sample is included in the spectral information. These noises mainly come from the experimental environment, such as temperature, natural light, and the instrument itself. If the noise contained in the spectral data is not removed, it will have a certain impact on the

*xiaodong@ise.neu.edu.cn

International Conference on Computer Application and Information Security (ICCAIS 2021),
edited by Yingfa Lu, Changbo Cheng, Proc. of SPIE Vol. 12260, 122600S · doi: 10.1117/12.2637768

accuracy of the model. Therefore, before using spectral data to establish a model, it is necessary to filter out the information and noise irrelevant to the sample spectral data as much as possible. The spectral curve of the original coal data is shown in Figure 1.

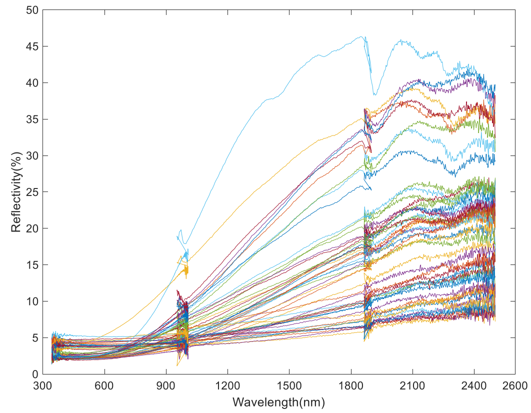


Figure 1. Spectral curve of coal data.

3. ESTABLISH DP-WTELM MODEL

3.1 Direct orthogonal signal correction

The principle of the direct orthogonal signal correction[10] method is to orthogonalize the spectral array with the concentration array before establishing the quantitative correction model, filter out the signals irrelevant to the concentration array, and then carry out a multivariate correction, to simplify the model and enhance the prediction ability of the model. The orthogonal mathematical method filters out the spectral signals independent of the concentration array, which can reduce the number of main factors used in establishing the model and improve the prediction ability and robustness of the correction model. The spectral curve of coal data processed by the direct orthogonal signal correction method is shown in Figure 2.

3.2 Principal component analysis

Principal component analysis (PCA) is not only a basic mathematical analysis method but also a method widely used to reduce dimension. The principle is to establish a few new variables as possible according to many variables with certain correlations, and these new variables are irrelevant. At the same time, these new variables can contain as much information as possible. In the DP_WTELM model, we mainly use PCA to reduce the dimension of data. Each spectral data has 1024 features. We set the cumulative contribution rate to 99% and extracted 9 principal components from the original data, as shown in Figure 3.

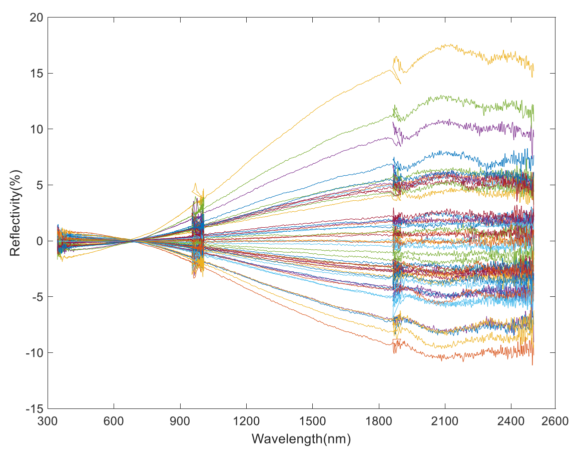


Figure 2. Spectral curve of coal data after DOSC processing.

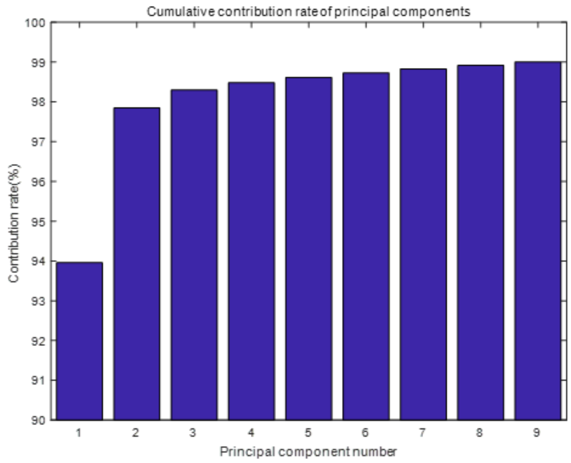


Figure 3. Cumulative contribution rate of 9 principal components.

3.3 Whale optimization algorithm

Whale optimization algorithm (WOA) is an algorithm proposed by Mirjalili et al.[11] to simulate the humpback whale hunting strategy. In contrast to classical algorithms, WOA is superior to classical algorithms in both accuracy and speed.

3.4 Extreme learning machine

In contrast to the traditional neural network structure, extreme learning machine (ELM) has the advantages of less iterations and fast training speed. We choose the TELM network with two hidden layers as the infrastructure. TELM not only takes into account these advantages of ELM, but also has higher accuracy. In this paper, the WOA is used to find the optimal weight and bias parameters of TELM, and the DP-WTELM network model for detecting coal quality is obtained. Figure 4 is a schematic diagram of the TELM network structure.

For any different N samples (x_i, t_i), here $\mathbf{x}_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T \in \mathbf{R}^n$, $\mathbf{t}_i = [t_{i1}, t_{i2}, \cdots, t_{im}]^T \in \mathbf{R}^m$. According to the ELM algorithm, the output matrix $\mathbf{H}_1$ of the first layer can be calculated and the output weight $\beta_1 = \mathbf{H}_1^+ \mathbf{T}$. Defining the expected output matrix of a second hidden layer as $\mathbf{H}_{2E}$, this matrix is obtained by equation (1).

$$\mathbf{H}_{2E} = \mathbf{T}\beta_1^+ \tag{1}$$

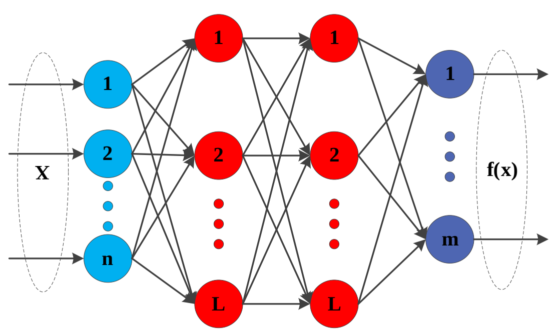


Figure 4. Network structure of TELM.

In the equation, β_1^+ is the Moore-Penrose generalized inverse matrix of β_1. Define a matrix $\mathbf{W}_{H2} = [\mathbf{B}_2 \ \mathbf{W}_2]$, where $\mathbf{B}_2$, $\mathbf{W}_2$ is the input deviation and input weight of the second hidden layer, then $\mathbf{W}_{H2}$ is obtained from equation (2).

$$\mathbf{W}_{H2} = g^{-1}(\mathbf{H}_{2E})\mathbf{M}_2^+ \tag{2}$$

In the equation, $\mathbf{M}_2 = [\mathbf{1}\ \mathbf{H}_1]^T$, vector $\mathbf{1}$ is a column vector composed of N scalars 1; $\mathbf{M}_2^+$ is the Moore Penrose generalized inverse matrix of $\mathbf{M}_2$, and the calculation method is the same as that of β_1^+; $g^{-1}(x)$ is the inverse of the activation function $g(x)$.

The output of the second hidden layer is defined as $\mathbf{H}_{2A}$. It can be obtained by equation (3).

$$\mathbf{H}_{2A} = g(\mathbf{W}_{H2}, \mathbf{M}_2) \tag{3}$$

Then, the output weight of the second hidden layer can be calculated β_2.

$$\beta_2 = \mathbf{H}_{2A}^+ \mathbf{T} \tag{4}$$

In the equation, $\mathbf{H}_{2A}^+$ is the Moore Penrose generalized inverse matrix of $\mathbf{H}_{2A}$, and the calculation method is the same as that of β_1^+.

Finally, the actual output of the whole TELM network is represented by equation (5).

$$f(x) = \mathbf{H}_2 \beta_2 \tag{5}$$

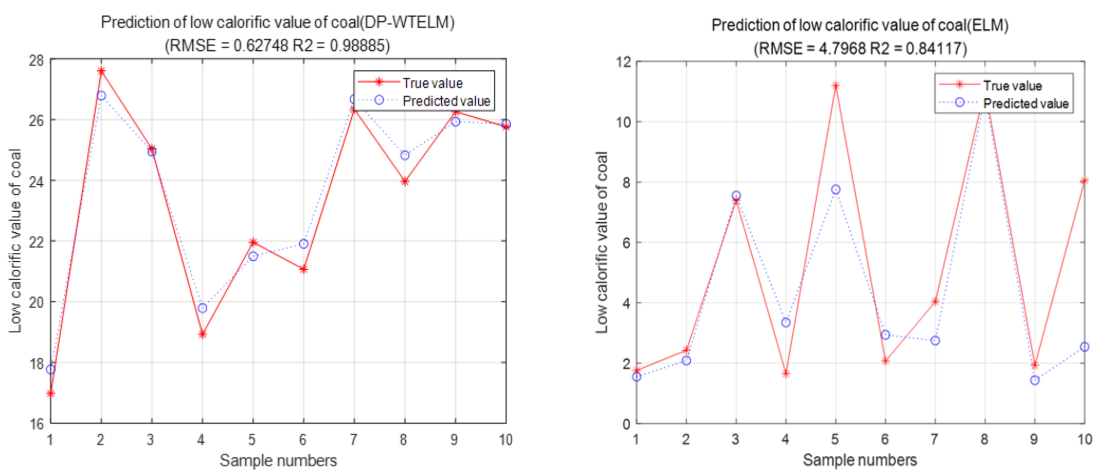
4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Preprocessing of coal spectral data

The spectral data of coal directly measured by spectrometer contains a lot of noise and the dimension is 1024 dimensions, if these data are directly input into the model, it will lead to problems such as complex model structure and long training time. Therefore, the DOSC and PCA are used to preprocess the data. After preprocessing, the data dimension is reduced from 1024 to 9, which greatly enhances the training speed of the model.

4.2 Coal quality test results

In the experiment, we collected 50 coal samples, of which 40 were used as training sets and 10 as test sets. Based on the coal spectral data, the coal quality detection model is established by using the DP_WTELM neural network algorithm. The test results are shown in Figure 5a.



(a) Based on DP_WTELM (b) Based on ELM

Figure 5. Prediction results of low calorific value based on different models.

Figure 5 shows the prediction results of the DP_WTELM model and ELM model on the low calorific value of coal respectively. It can be seen that the overall prediction effect of low calorific value of coal samples in figure 5a is marvelous, while the prediction results of low calorific value of multiple coal samples in figure 5b have large errors. It can be seen that the DP_WTELM model can obtain more accurate results in predicting the low calorific value of coal than the ELM model. From the experimental results, the R2 value is 0.98885 and the RMSE value is 0.62748, which verifies that the DP_WTELM model has good prediction performance.

5. CONCLUSION

In this paper, a coal quality prediction model based on machine learning is proposed, and the practicability and accuracy of the model are verified by field coal samples. Firstly, 50 coal samples and their spectral data are collected, preprocessed by DOSC and PCA, and then the low calorific value in coal is predicted by WTELM network. The results show that the R^2 value and RMSE value of this method are 0.98885 and 0.62748, which can meet the industrial requirements. In contrast to the traditional coal quality detection methods, the method based on the DP-WTELM model has many advantages, which provides a new idea for coal quality detection.

REFERENCES

[1] Wang, Y. S., Yang, M., Wei, G., Hu, R. F., Luo, Z. Y. and Li, G., "Improved PLS regression based on SVM classification for rapid analysis of coal properties by near-infrared reflectance spectroscopy," Sensors and Actuators B: Chemical, 193, 723-729 (2014).

[2] Zeng, D. L., Hu, Y., Liu, J. Z., Zhao, Z. and Gao, S., "Soft sensing of coal moisture," Measurement, 60, 231-239 (2015).

[3] Zeng, D. L., Wang, Y. H., Gao, S. and Liu, J. Z., "Modeling and control of ball mill system considering coal moisture," Measurement, 90, 43-51 (2016).

[4] Hu, R. F., Wang, Y. S., Yang, M., Li, X. M., Luo, Z. Y. and Li, G., "Improved analysis of inorganic coal properties based on near-infrared reflectance spectroscopy," Analytical Methods, 7(12), 5282-5288 (2015).

[5] Wang, Y. S., Yang, M., Luo, Z. Y., Wang, Y., Li, G. and Hu, R. F., "Rapid coal classification based on confidence machine and near infrared spectroscopy," Spectroscopy and Spectral Analysis, 36(6), 1685-1689 (2016).

[6] Ren, Y., Sun, X. J., Dai, X. A., Cen, Y., Tian, Y. M., Wang, N. and Zhang, L. F., "Variety identification of bulk commercial coal based on full-spectrum spectroscopy analytical technique," Spectroscopy and Spectral Analysis, 38(2), 352-357 (2018).

[7] Qin, H. Q., Lu, Z. M., Yao, S. C., Li, Z. H. and Lu, J. D., "Combining laser-induced breakdown spectroscopy and fourier-transform infrared spectroscopy for the analysis of coal properties," Journal of Analytical Atomic Spectrometry, 34, 347-355 (2019).

[8] Huang, G. B., Zhu, Q. Y. and Siew, C. K., "Extreme learning machine: Theory and applications Neurocomputing, 70(1-3), 489-501 (2006).

[9] Qu, B. Y., Lang, B. F., Liang, J. J., Qin, A. K. and Crisalle, O. D., "Two-hidden-layer extreme learning machine for regression and classification," Neurocomputing, 175, 826-834 (2016).

[10] Westerhuis, J. A., de Jong, S. and Smilde, A. K., "Direct orthogonal signal correction," Chemom. Intell. Lab Syts., 56(1), 13-25 (2011).

[11] Mirjalili, S. and Lewis, A., "The whale optimization algorithm," Advances in Engineering Software, 95, 51-67 (2016).