Research on Pattern Recognition of Gas Sensor Array with Weighted Support Vector Regression Machine

Shudi Peng [a], Lingyun Wan [b], Wei Song [c] and Gaolin Wu [d]

Chongqing Electric Power Research Institute, Chongqing 401123, China

[a]pengshudi@163.com, [b]wanlingyun@163.com, [c]songwei@163.com, [d]wugaolin@163.com

Abstract

In this study, an weighted support vector regression machine (W-SVR) is proposed and applied for pattern regression of gas sensor array. Simultaneously, the back propagation neural network (BPNN) and the standard support vector regression machine (SVR) are also employed for comparison. All testing results show that this proposed W-SVR method is more powerful than traditional neural BPNN and the standard SVR in the aspect of forecasting accuracy and generalization, and effectively solve the cross-sensitivity of the multi-component gas sensor array.

Keywords

Gas sensor array, pattern recognition, support vector regression machine, back propagation neural network.

1. Introduction

Pattern recognition is one of the key technologies in the gas sensor array detection. At present, there are a large number of pattern recognition methods with different advantages and disadvantages [1-2]. Support Vector Machine SVM is a new machine learning method based on statistical learning theory. It has a strict mathematical foundation, and overcome the traditional neural network shortcomings based on the structural risk minimization criteria [3-4]. It also has good statistical regularities and good generalization ability in the case of small statistical sample size, and is mainly used in pattern classification and regression analysis. SVM in many fields such as image processing, data analysis, classification, and fault diagnosis has been successfully applied [5].

With the introduction of insensitive loss function ε, SVM has been extended to nonlinear regression function estimation, and shows good learning performance. In standard SVR learning algorithm for regression, the choice of design parameters C and ε is crucial to construct the regression function. However, in practical applications, we often find that some important sample data require a small training error [6]. In other words, the accuracy requirements for different sample data will be different. Therefore, in the description of the optimization, each sample data should have a different error requirements and penalty coefficient to obtain much more accurate regression estimation.

To this end, an improved support vector regression machine model is presented for pattern recognition of the monitoring sensor array for multi-component gases dissemination in transformer oil. On the basis of standard regression model, C and ε are weighted with different weight coefficients according to the importance of the sample data. And a comparative study is also investigated with traditional BP neural network and standard SVR method.

2. Standard Support Vector Regression Machine (S-SVR)

$\left\{\left(x_i, y_i\right), i=1,2,\cdots,l\right\}$ is the the given training data, where $x_i \in R^d$ is the input value of sample i, $x_i=\left[x_i^1, x_i^2,\cdots,x_i^d\right]^T$ is a d dimension column vector, and $y_i \in R$ is the corresponding target value. The linear ε insensitive loss function is defined as follow:

$$\left|y-f(x)\right|_{\varepsilon}=\begin{cases}0 & \left|y-f(x)\right|\le\varepsilon\\ y-f(x)-\varepsilon & \left|y-f(x)\right|>\varepsilon\end{cases} \tag{1}$$

If the difference between the target value y and the regression value $f(x)$ is less than ε, the linear ε insensitive loss value equal to 0. By defining the appropriate kernel function $K(x_i, x_j)$, support vector machine nonlinearly maps the input sample space to another feature space, and then builds the regression estimation function in this feature space[7-8]. The nonlinear regression function is defined as $f(x)=w^T\phi(x)+b$ and the goal of SVR is to find w and b.

Taking the constraints into account, two slack variables ξ_i 、 $\xi_i^{'}$ are introduced, and the optimization problem is:

$$\begin{cases}\min\limits_{w} \ \dfrac{1}{2}w^T w+C\sum\limits_{i=1}^{l}\left(\xi_i+\xi_i^{'}\right)\\ s.t.\, y_i-w\cdot\phi(x_i)-b\le\varepsilon+\xi_i\\ \quad w\cdot\phi(x_i)+b-y_i\le\varepsilon+\xi_i^{'}\\ \qquad \xi_i\ge 0\\ \quad \xi_i^{'}\ge 0 \quad i=1,2,\cdots,l\end{cases} \tag{2}$$

And the dual optimization problem is turned into:

$$\begin{cases}\min\limits_{\alpha,\alpha'}\left[-\dfrac{1}{2}\sum\limits_{i=1}^{l}\sum\limits_{j=1}^{l}\left(\alpha_i-\alpha_i^{'}\right)\left(\alpha_j-\alpha_j^{'}\right)K\left(x_i,x_j\right)-\right.\\ \qquad \left.\varepsilon\sum\limits_{i=1}^{l}\left(\alpha_i+\alpha_i^{'}\right)+\sum\limits_{i=1}^{l}y_i\left(\alpha_i+\alpha_i^{'}\right)\ \right]\\ \qquad s.t. \quad \sum\limits_{i=1}^{l}\left(\alpha_i-\alpha_i^{'}\right)=0\\ \qquad\qquad 0\le\alpha_i\le C\\ \qquad\qquad 0\le\alpha_i^{'}\le C\end{cases} \tag{3}$$

The obtained SVR function is:

$$\begin{cases}\varepsilon-y_i+f(x_i)=0, & 0<a_i<C\\ \varepsilon+y_i-f(x_i)=0, & 0<a_i'<C\end{cases} \tag{4}$$

$$b=\frac{1}{N_{NSV}}\left\{\sum_{0<\alpha_i<C}\left[y_i-\sum_{x_i\in sv}\left(\alpha_j-\alpha_j^{'}\right)K\left(x_j,x_i\right)-\varepsilon\right]+\sum_{0<\alpha_i^{'}<C}\left[y_i-\sum_{x_j\in sv}\left(\alpha_j-\alpha_j^{'}\right)K\left(x_j,x_i\right)+\varepsilon\right]\right. \tag{5}$$

Where N_{NSV} is the number of standard support vector.

3. Weighted Support Vector Regression Machine (W-SVR)

For the output signal of the sensor array, the pattern recognition effect of the standard SVR is not conclusive, especially with different square deviation σ_i^2 of the random error. To solve this problem, a proper weighting λ_i is introduced to adjust the weight of the items in the regression model. That is to say the weight coefficient is used to adjust the impact of square deviation in the standard SVR. And the optimization problem is converted into as follow form.

$$\begin{cases} \min\limits_{w} \ \frac{1}{2} w^T w + C \sum\limits_{i=1}^{l} \lambda_i \left(\xi_i + \xi_i^{'} \right) \\ s.t.\, y_i - w \cdot \phi(x_i) - b \le \varepsilon + \xi_i \\ w \cdot \phi(x_i) + b - y_i \le \varepsilon + \xi_i^{'} \\ \xi_i \ge 0 \\ \xi_i^{'} \ge 0 \quad i = 1, 2, \cdots, l \end{cases} \tag{6}$$

And the dual optimization problem is:

$$\begin{cases} \min\limits_{\alpha,\alpha'} \Big[-\frac{1}{2} \sum\limits_{i=1}^{l} \sum\limits_{j=1}^{l} (\alpha_i - \alpha_i^{'})(\alpha_j - \alpha_j^{'}) K(x_i, x_j) - \\ \varepsilon \sum\limits_{i=1}^{l} (\alpha_i + \alpha_i^{'}) + \sum\limits_{i=1}^{l} y_i (\alpha_i + \alpha_i^{'}) \ \Big] \\ s.t. \quad \sum\limits_{i=1}^{l} (\alpha_i - \alpha_i^{'}) = 0 \\ 0 \le \alpha_i \le \lambda_i C \\ 0 \le \alpha_i^{'} \le \lambda_i C \end{cases} \tag{7}$$

The expression of $f(x)$ is:

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^{'}) K(x_i, x) + b \tag{8}$$

According to the Karush-Kuhn-Tucker (KKT) optimization condition, the follow formation can be obtained.

$$\begin{cases} \varepsilon - y_i + f(x_i) = 0, & 0 < a_i < \lambda_i C \\ \varepsilon + y_i - f(x_i) = 0, & 0 < a_i' < \lambda_i C \end{cases} \tag{9}$$

Based on equation 9, the parameter b can be calculated.

4. Pattern Recognition of Gas Sensor Array with W-SVR

The gas sensor array detection system used in this article mainly consists of three components, the gas sensor array, signal conditioning and data acquisition, and the pattern recognition unit with W-SVR. The gas sensor array is composed of six gas sensors MQ_1~MQ_6 and installed in the power transformer drain valve to sensing dissolved H_2, CO, CH_4, C_2H_2, C_2H_4 and C_2H_6, respectively.

In order to obtain satisfactory prediction, we firstly normalizes the data to eliminate the size difference between the data on the calculated result. The maximum difference dormalization method is selected to normalize the training set and test set in this study.

$$x_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{10}$$

Where x_i is the data after normalization, X_i is the gas sensor array output value, $X_{\max}$ and $X_{\min}$ are respectively the corresponding maximum and minimum output values of the gas sensor array .

The selection of the kernel function has a great influence on the accuracy of the prediction W-SVR model. Different kernel functions can constructe different types of learning machines. Because the radial basis kernel function (RBF) have only one variable, we select the RBF function in this W-SVR machine：

$$K\left(x_i, x\right) = \exp\left(-\left\|x_i - x\right\| / 2\sigma^2\right) \tag{11}$$

Some previous literatures have reported that C, ε, and parameters σ in RBF Kernel function affect greatly to the prediction effect of W-SVR. In this paper, we use the leave-one cross-validation method to obtain the optimum combination of parameters C, ε, and σ. According to this method, the best parameters for C, ε, and σ are selected as 200, 3.62, and 0.08, respectively.

The maximum error and the average relative percentage error are used as the indicators to evaluation the prediction accuracy.

$$\begin{cases} E_1 = Max\left|y_k - \hat{y}_k\right| \\ E_2 = \dfrac{1}{N}\sum_{k=1}^{N}\dfrac{\left|y_k - \hat{y}_k\right|}{y_k} \end{cases} \tag{12}$$

Where E_1 and E_2 are the maximum error and the average relative percentage error, y_k 、 $\hat{y}_k$ are separately the actual and predicted values, and N is the number of forecast points.

Table 1 Identification results with BPNN(μL/L)

	H_2	CO	CH_4	C_2H_4	C_2H_2	C_2H_6
1	43.25	105.47	18.32	24.87	0.34	22.38
2	85.76	265.27	74.01	42.97	0.45	32.48
3	105.66	289.78	137.26	71.32	0.08	55.18
4	211.39	312.25	240.04	207.36	0.32	42.97

Table 2 Identification results with standard SVR(μL/L)

	H_2	CO	CH_4	C_2H_4	C_2H_2	C_2H_6
1	42.07	103.78	20.96	18.94	0.12	26.87
2	82.63	258.21	72.68	43.67	0.28	33.07
3	103.76	285.43	141.78	75.27	0.05	52.75
4	206.48	308.36	245.69	204.39	0.26	38.36

Table 3 Identification results with weighted SVR(μL/L)

	H_2	CO	CH_4	C_2H_4	C_2H_2	C_2H_6
1	41.23	101.66	19.57	20.85	0.06	26.25
2	81.85	255.36	68.64	43.97	0.20	33.89
3	102.37	283.86	141.21	77.16	0.02	48.34
4	204.31	305.23	252.38	202.75	0.19	41.02

Table 1, Table 2, and Table 3 represent the recognition results of the gas sensor array using the BP neural network, the standard support vector regression machine, and the weighted support vector regression machine, respectively. And the corresponding maximum error E_1 and average relative percentage error E_2 are separately shown in Table 4, Table 5, and Table 6.

Table 4 Identification error with BPNN

	H_2	CO	CH_4	C_2H_4	C_2H_2	C_2H_6
E_1(μL/L)	11.39	15.27	9.96	8.68	0.45	5.18
E_2(%)	6.67	4.79	5.02	10.9	0.45	8.87

Table 5 Identification error with standard SVR

	H_2	CO	CH_4	C_2H_4	C_2H_2	C_2H_6
E_1(μL/L)	6.48	8.36	4.31	4.73	0.28	2.75
E_2(%)	3.87	2.95	2.91	4.09	0.28	5.65

Table 6 Identification error with weighted SVR

	H_2	CO	CH_4	C_2H_4	C_2H_2	C_2H_6
E_1(μL/L)	4.31	5.36	2.38	2.84	0.20	1.66
E_2(%)	2.48	1.73	1.48	2.87	0.20	3.51

One can clearly see from Table 1 to Table 6 that, in the case of small amount of samples, the weighted support vector machine for regression model can adjust the parameters to make the error as small as possible and hand the regression function as smooth as possible.Compared with traditional BP neural network and the standard SVR, the proposed W-SVR is more powerful in the aspect of forecasting accuracy and generalization, and effectively solving the cross-sensitivity of the multi-component gas sensor for monitoring.

5. Conclusion

Aiming at the problems such as difficult determination of net structure, over fitting and low convergence rate local minimization of traditional neural networks, and the problem of no considering the importance of each sample in the standard SVR, combined with multi-component gas monitoring sensor array for power transformer oil, an weighted support vector machine for regression is applied in gas sensor array signal pattern recognition. The testing results show that this pattern recognition method is more powerful than traditional neural network and the standard SVR in the aspect of forecasting accuracy and generalization, and effectively solve the cross-sensitivity of the multi-component gas sensor for monitoring.

References

[1] V. Vapnik. The nature of staistical learning theory [M]. New York: Springer-Verlag, 1995.
[2] J. Lin, J. Zhang Jing, and Z. T. She. Neural network based on-line detection of dissolved gas in transformer oil [J]. Automation of Electric Power Systems, Vol. 25 (2001) No.8, p.56-58.

[3] C. Burges. A tutorial on support vector machines for pattern recognition [J]. Data Mining & Knowle. Discov, Vol. 2 (1998) No.2, p.121-127.
[4] G. L. Su, and F. P. Deng. Introduction to model selection of SVM regression [J]. Bulletin of Science and Technology, Vol. 22 (2006) No.2, p.154-158.
[5] W. J. Peng, and X. Q. Luo. Research on vibrant fault diagnosis of hydro-turbine generating unit based on wavelet packet analysis and support vector machine [J]. Proceedings of the CSEE, Vol. 26 (2006) No.24, p.164-168.
[6] G. A. Bakken, G. W. Kauffman, P. C. Jurs, and K. J. Albert. Pattern recognition analysis of optical sensor array data to detect nitroaromatic compound vapors [J]. Sensors and Actuators B: Chemical, Vol. 79 (2001) No.1, p.1-10.
[7] X. Shi, L. Wang, N. Kariuki, J. Luo, C. J. Zhong, and S. Lu. A multi-module artificial neural network approach to pattern recognition with optimized nanostructured sensor array [J]. Sensors and Actuators B: Chemical, Vol. 117 (2006) No.1, p.65-73.
[8] Q. Zhou, W. G. Chen, X. P. Su, and S. D. Peng. Quantitative recognizing dissolved hydrocarbons with genetic algorithm-support vector regression [J]. TELKOMNIKA Indonesian Journal of Electrical Engineering, Vol. 11 (2013) No.9, p.5509-5516.