SCIENCE CHINA
Information Sciences

• LETTER •

April 2019, Vol. 62 049202:1–049202:3
https://doi.org/10.1007/s11432-018-9673-2

An ensemble model based on weighted support vector regression and its application in annealing heating process

Yongyue ZHANG[1], Weihua CAO[2,3*], Yali JIN[2,3] & Min WU[2,3]

[1]*School of Information Science and Engineering, Central South University, Changsha 410083, China;*
[2]*School of Automation, China University of Geosciences, Wuhan 430074, China;*
[3]*Hubei key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, China University of Geosciences, Wuhan 430074, China*

Received 7 September 2018/Accepted 23 November 2018/Published online 22 February 2019

Citation Zhang Y Y, Cao W H, Jin Y L, et al. An ensemble model based on weighted support vector regression and its application in annealing heating process. Sci China Inf Sci, 2019, 62(4): 049202, https://doi.org/10.1007/s11432-018-9673-2

Dear editor,

In the steel industry, the annealing heating production directly impacts the economic benefits of enterprises; hence, accurately predicting the required gas flow is very important [1]. On account of high equipment complexity, it costs lots of time for maintaining a traditional mechanism model at a good working condition; data-driven model, such as support vector regression (SVR), has been used in industry processes [2].

However, owing to diversity of operating conditions, real data collected from annealing heating process for modeling contains many singular points and noise. Therefore, achieving high predicting precision with an SVR model implemented using real factory data is difficult. Moreover, because of the structural risk minimization principle in the SVR algorithm, the predicted data is smoother, and the influence of singular points which creates from some rare operation conditions on the model is eliminated, which leads to greater error output by the model at some operation conditions and negatively affects the industrial production [3].

To increase the influence of singular points on the prediction model, Diao et al. [4] proposed the weighted support vector regression (WSVR) method to change the weight values of sample points based on the time difference, such that the model exhibits stronger timeliness. However, this method of setting the sample weight value cannot adapt to samples without time characteristics; sample weight assignment is difficult for untrained data.

Recently, machine learning [5] has gained popularity; moreover, ensemble learning can overcome the challenge of sample weight assignment in classification problems [6]. AdaBoost [7] is an ensemble learning algorithm that can adjust the sample weights to adapt the prediction accuracy of a classifier and transforming those from weak classifiers into a strong classifier. However, predicting the temperature in the annealing heating process is a regression problem. Defining a weight adjustment strategy in the regression model is the primary task in applications of the AdaBoost algorithm.

To solve the problem of low prediction accuracy achieved by the ordinary SVR model in the annealing heating process, we first define the standard of the accuracy of the regression model prediction results. Then, we use WSVR to train the prediction models using the original data set; the prediction accuracy of the WSVR models determines the sample and model weights. Finally, the AdaBoost algorithm is used to integrate multiple

* Corresponding author (email: weihuacao@cug.edu.cn)

weak regression models into an ensemble model to improve the accuracy of the process model.

WSVR algorithm. Let $S = \{(x_i, y_i), i = 1, \ldots, m\}$ be a training set comprising sample data collected from the annealing heating process, where $x_i \in \mathbb{R}^N$ and $y_i \in \mathbb{R}$, N represents the input characteristics used to describe the process. The selection of this value determines the number of input parameters for the regression model.

First, the variable x is nonlinearly mapped to the high-order feature space Ω by $\phi(x)$, and the linear regression function $f(x) = W^{\mathrm{T}}\phi(x) + b$ is constructed in Ω. According to the standard SVR algorithm, W and b can be obtained by solving the convex quadratic programming problem as follows:

$$\min_{W,b} P = \frac{1}{2}W^{\mathrm{T}}W + C\sum_{i=1}^{m}(\xi_i + \xi_i^*),$$

where C is the penalty coefficient, ξ_i and ξ_i^* are slack variables. In WSVR, the penalty coefficient of each sample changes according to the weight of the sample. Simultaneously, WSVR has the same constraints as the standard SVR:

$$\begin{aligned}&\min_{W,b} P = \frac{1}{2}W^{\mathrm{T}}W + C\sum_{i=1}^{m}q_i(\xi_i + \xi_i^*),\\ &\text{s.t.}\begin{cases} y_i - (W^{\mathrm{T}}\phi(x_i) + b) \leqslant \varepsilon + \xi_i,\\ (W^{\mathrm{T}}\phi(x_i) + b) - y_i \leqslant \varepsilon + \xi_i^*,\\ \xi_i, \xi_i^* \geqslant 0,\ i = 1, 2, \ldots, m.\end{cases}\end{aligned} \tag{1}$$

The Lagrangian multiplier method can be used to convert (1) to its dual problem. And α and α^* are solved using

$$\begin{aligned}&\min_{\alpha,\alpha^*} D = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}Q_{ij}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) +\\ &\varepsilon\sum_{i=1}^{m}(\alpha_i + \alpha_i^*) - \sum_{i=1}^{m}y_i(\alpha_i - \alpha_i^*),\\ &\text{s.t.}\begin{cases} 0 \leqslant \alpha_i, \alpha_i^* \leqslant Cq_i,\\ \displaystyle\sum_{i=1}^{m}(\alpha_i - \alpha_i^*) = 0,\ i = 1, 2, \ldots, m,\end{cases}\end{aligned} \tag{2}$$

where $Q_{ij} = \phi(x_i)^{\mathrm{T}}\phi(x_j) = K(x_i, x_j)$, $K(x_i, x_j)$ is a kernel function. The final regression function can be expressed as

$$f(x) = \sum_{i=1}^{m}(\alpha_i - \alpha_i^*)K(x_i, x) + b. \tag{3}$$

AdaBoost algorithm. Let WSVR be a basic regression model and the AdaBoost maximum training times be G. q_i^g represents the weight of the i-th sample at the g-th training, and δ is the maximum allowable error between the sample value and the predicted value. Generally, in the industrial field, the value of δ is directly related to the assessment standard. The steps of this algorithm are as follows.

Step 1. Initially assume that all samples have equal weights, as $q_i^1 = 1/m$.

Step 2. Set $g = 1, 2, \ldots, G$, the regression model $f^g(x)$ is obtained according to WSVR, and the error rate is calculated using

$$r^g = \frac{\sum_{i=1}^{m} l_i^g}{m} \times 100\%, \tag{4}$$

where $l_i^g = \begin{cases}1, & |y_i - f^g(x_i)| \geqslant \delta\\ 0, & |y_i - f^g(x_i)| < \delta.\end{cases}$ If r^g is greater than 50%, the loop is exited.

Step 3. Set

$$\beta^g = \frac{1}{2}\ln\left(\frac{1 - r^g}{r^g}\right) \tag{5}$$

as the model weight, and adjust the sample weights using the following formula:

$$q_i^{g+1} = \frac{u_i^{g+1}}{\sum_{i=1}^{m} u_i^{g+1}}, \tag{6}$$

where $u_i^{g+1} = q_i^g \times \begin{cases}\exp(-\beta^g), & |y_i - f^g(x_i)| < \delta,\\ \exp(\beta^g), & |y_i - f^g(x_i)| \geqslant \delta.\end{cases}$ In the training process, this algorithm tries to provide different weights for those singular points to improve their performance. Then, return to Step 2.

Step 4. At the end of the cycle, the final ensemble model ABWSVR is shown as

$$F(x) = \frac{\sum_{g=1}^{G}\beta^g f^g(x)}{\sum_{g=1}^{G}\beta^g}. \tag{7}$$

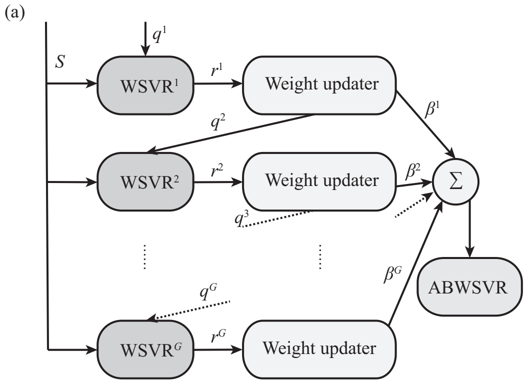
Finally, the structure of ABWSVR is as shown in Figure 1(a).

Verification. We collected 150 sets of annealing furnace heating data that were stably produced under different operating conditions on an industrial site. We used the coil width, coil height, coil temperature, process speed, gas-air ratio and the calorific value as input together with gas flow rate as output for training and prediction. The first 100 samples were set as the training set, and the remaining 50 samples were included in the testing set.

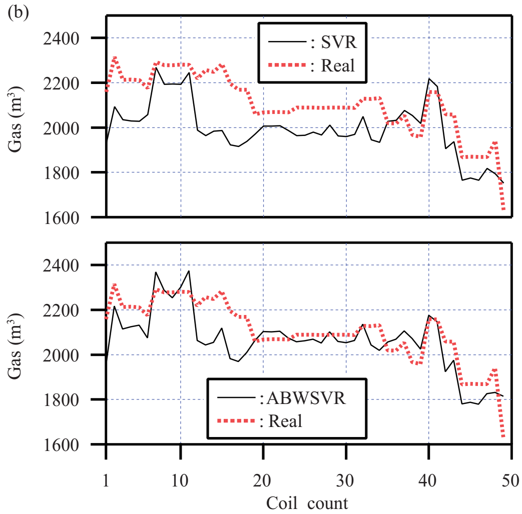
First, we used SVR to build a prediction model. Figure 1(b) shows the values predicted by the model. While the SVR results can be fitted to the trend of the data, the overall error is too large and the tracking effect on some operation conditions is not ideal.

Next, we built an ABWSVR model by applying the algorithm proposed in this study. we used the AdaBoost maximum training time $G = 8$, and $\delta = 100\ \mathrm{m}^3$ according to the field accuracy requirements. Figure 1(b) shows the predicted results

(a)



(b)



(c)

Model	r (%)	MAE (m^3)
SVR	58	126.2
ABWSVR	28	79.2

Figure 1 (Color online) (a) Structure of ABWSVR; (b) predicted results of SVR and ABWSVR; (c) comparison of the models.

based on the test samples. The fitting curve in the figure is significantly more realistic than that of the SVR model, and the tracking effect for singular points is more obvious.

According to (4), $r_{\rm SVR}$ and $r_{\rm ABWSVR}$ can be calculated. Figure 1(c) shows the results. Compared to the SVR model, the prediction error rate of the ABWSVR model is reduced by 30%, base on this result, using chi-square test to calculate $\chi^2 = 16.34$ which proves that the difference between the two models has highly statistically significance and the ABWSVR model is better than the SVR model. In addition, the mean absolute error (MAE) is reduced by 47 m^3, which also proves the effectiveness and accuracy of the ABWSVR model.

Conclusion. In this study, new model based on WSVR is proposed to achieve higher prediction accuracy for the annealing heating process than an SVR based model can provide. First, the WSVR algorithm is realized to improve the identification of sample specific values; the model weights are given in combination with the on-site assessment criteria. The training sample weights of each WSVR model are adaptively adjusted by the last WSVR model accuracy to fit the different operation conditions. Finally, the ensemble model is obtained using the AdaBoost algorithm, which effectively improves the prediction accuracy and creates significant benefits for enterprise production.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61773354), Hubei Provincial Natural Science Foundation of China (Grant No. 2015CFA010), and 111 Project (Grant No. B17040).

References

1 Wu H, Benschop B, Ben Driss O, et al. Furnace combustion and control renovation to improve the productivity of a continuous annealing line. Energy Procedia, 2017, 120: 454–461
2 Zhu B, Xie L, Han D, et al. A survey on recent progress in control of swarm systems. Sci China Inf Sci, 2017, 60: 070201
3 Sanz-Garcia A, Fernandez-Ceniceros J, Antonanzas-Torres F, et al. GA-PARSIMONY: a GA-SVR approach with feature selection and parameter optimization to obtain parsimonious solutions for predicting temperature settings in a continuous annealing furnace. Appl Soft Comput, 2015, 35: 13–28
4 Diao X, Li Q. On-line training algorithm and its application based on weighted SVR. J Syst Simul, 2007, 19: 3970–3973
5 Jordan M I, Mitchell T M. Machine learning: trends, perspectives, and prospects. Science, 2015, 349: 255–260
6 Xiao L, Dong Y X, Dong Y. An improved combination approach based on Adaboost algorithm for wind speed time series forecasting. Energy Conv Manag, 2018, 160: 273–288
7 Kim E Y. Vision-based wheelchair navigation using geometric AdaBoost learning. Electron Lett, 2017, 53: 534–536