

# Research of the Machinery Fault Diagnosis and Prediction Based on Support Vector Machine

Xiujuan Qie<sup>a</sup>, Jing Zhang<sup>b</sup>, Jiangya Zhang<sup>c</sup>

Department of Electrical and Mechanical Engineering, HeBei College Of Science And Technology, China

<sup>a</sup>942711710@qq.com, <sup>b</sup>42588122@qq.com, <sup>c</sup>153293237@qq.com

**Keywords:** Machinery fault diagnosis; Fault trend prediction; Support vector machine

**Abstract.** This paper analyzes the theory of support vector machines-SVM and discusses the algorithms of SVM classification and regression. After overviewed the SVM application research on machinery fault diagnosis and prediction recently, it discusses the merits and deficiencies of SVM and the points out the bright application research on machinery fault diagnosis and prediction. It presents the SVM model for machine condition trend prediction. It is proved that SVM model has good predict ability for long time period by comparing the AR model and SVM model for a test system vibration signal.

## Introduction

In order to achieve predictable on the basis of online condition monitoring and fault analysis based on electromechanical equipment maintenance, the use of intelligent diagnosis and indicate the method based on neural network. This method has the ability to approximate complex nonlinear systems and classification capabilities, but requires a lot of typical faults data samples and a priori knowledge. Due to various reasons, in practice it is difficult to obtain a large number of samples. Statistical learning theory is a specialized machine learning theory of small samples under the law, and the core concept is the VC dimension. On the basis of VC dimension theory and structural risk minimization principle, V. Vapnik et established and developed the support vector machines (SVM). It is based on a limited sample of information to find the best compromise between complexity and learning ability of the model to get the best generalization ability[1].

## Support vector machine

SVM is a new learning method based on structural risk minimization principle of the machine, which can make full use of the limited sample learning acquisition decision function with high generalization ability. Considering a two classification models, a training sample is defined as  $\{(x_i, y_i)\}$ , where  $x_i$  represents the input vector and  $y_i$  represents the classification sign. The two classes of linear discriminant function of separable cases are as follows:

$$f(x) = w \cdot \phi(x) + b. \quad (1)$$

Where  $x$  is the sample vector,  $w$  is the weight vector,  $b$  represents the classification threshold. Supposing there is a classification decision plane:

$$f(x) = w \cdot \phi(x) + b = 0. \quad (2)$$

Which satisfies:

$$f(x) = \begin{cases} w \cdot \phi(x) + b > 0 & y_i = +1 \\ w \cdot \phi(x) + b < 0 & y_i = -1 \end{cases}. \quad (3)$$

Thus, Eq.(2) is defined as the classification and ultra flat surface of support vector.

A two types of linear separable linear sample set  $\{(x_i, y_i)\}^N$ ,  $i = 1, 2, \dots, n$  is given. The weight vector  $w$  in hyperplane and classification threshold  $b$  are both obtained. As to any support vector  $x_{sv}$  and its classification label  $y_{sv}$ , it satisfies the following condition.

$$f(x_{sv}) = W^T x_{sv} + b = \begin{cases} +1 & y_{sv} = +1 \\ -1 & y_{sv} = -1 \end{cases}. \quad (4)$$

The maximum interval of the classification plane can be calculated as  $\frac{2}{\|w\|}$ , Therefore, the maximum class interval is equivalent to the minimum. If the surplus plane need to classify all samples correctly, it must satisfy:

$$y_i(W^T x_i + b) - 1 \geq 0, i = 0, 1, 2, \dots, N. \quad (5)$$

Thus, the SVM hyperplane by solving the following constrained optimization problem was the solution:

$$\min_w \frac{1}{2} W^T W. \quad (6)$$

Which is a A typical quadratic programming problem. The optimal solution is a saddle point of the following Lagrangian:

$$L(w, b, a) = \frac{1}{2} (w, w)^T - \sum_{i=1}^l \alpha_i [(x_i, w) - b] - 1. \quad (7)$$

Where  $a_i$  is the Lagrange multipliers, The above optimization problem inverts into a dual form to get the solution:

$$\max W(\alpha) = \min \left( -\frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j k(x_i, x_j) \right) \sum_{i=1}^n \alpha_i. \quad (8)$$

According KKT conditions, those points on the boundary of two types of sample points fall in support vector machine SVM required solution. Classification decision function is obtained as follows:

$$f(x) = \text{sign} \left( \sum_{i,j=1}^n \alpha_i y_i k(x_i, x_j) + b \right). \quad (9)$$

Where  $k(x_i, x_j)$  is the kernel function.

The Gaussian kernel function is selected as the support vector machine kernel function in this paper. The form of Gaussian kernel function is shown as follows:

$$K(x, y) = \exp \left( -\frac{\|x - y\|^2}{\sigma^2} \right). \quad (10)$$

Thus, the optimal separating hyperplane is required:

$$f(x) = \text{sign} \left( \sum_{i,j=1}^n \alpha_i y_i \exp \left( -\frac{\|x_i - x_j\|^2}{\sigma^2} \right) + b \right). \quad (11)$$

For SVM with Gaussian kernel function, the parameters conclude penalty parameter  $C$  and core width  $\sigma$ . The penalty parameter  $C$  is made between the structure and the risk of sample

error compromise. The value of  $C$  is related to the tolerable error. A larger value allows small error and the smaller value allows larger errors. Core width  $\sigma$  is related to the input space of learning sample or the width. If the extent of the input sample is large, the value is large. On the contrary, if the extent of the input sample is small, the value is also small.

### The prediction model based on SVM

For a certain time series  $\{x_1, x_2, \dots, x_n\}$ ,  $i = 1, 2, \dots, n$ ,  $\{x_n\}$  represent the prediction target value. Establish the mapping relationship between input and output  $y_n = \{x_n\}: R^m \rightarrow R$ , where  $m$  is defined as the embedding dimension. Then, the learning samples of Support vector machine can be obtained, which are shown as follows[2]:

$$X = \begin{bmatrix} x_1 & x_1 & \cdots & x_m \\ x_2 & x_3 & \cdots & x_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-m} & x_{n-m-1} & \cdots & x_{n-1} \end{bmatrix}, \quad Y = \begin{bmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{bmatrix} \quad (12)$$

The regression function for training support vector machine is defined as:

$$y_i = \sum_{i=1}^{n-m} (\alpha_i - a_i^*) K(x_i^* - x_t^*) + b, \quad t = m+1, m+2, \dots, n \quad (13)$$

Then, the prediction value of No.  $m+1$  can be calculated:

$$y_{n+1} = \sum_{i=1}^{n-m} (\alpha_i - a_i^*) K(x_i^* - x_{n-m+1}^*) + b \quad (14)$$

The final prediction error (FPE) criterion is applied to evaluate the applicability of the proposed model and the embedding dimension of training model. The prediction process of SVM is shown as follows:

- (1) Enter the history sample data and make pretreatment; Calculate the embedding dimension according to FPE criteria; Obtain the training and testing samples.
- (2) Model initialization. Assign a random initial value for  $\alpha_i$ ,  $\alpha_i^*$  and  $b$ ; Establish the objective function through training samples and calculate the value of  $\alpha_i$ ,  $\alpha_i^*$  and  $b$ .
- (3) Put the obtained parameter values into estimated function, and calculate the predicted value at a future time with testing samples.
- (4) Compute the error function, when the absolute value of the error is less than a preset value, the end of the learning process, otherwise continue to learn.

### The specific example analysis

SVM theory is mainly applied in face detection and other pattern recognition, function approximation and nonlinear system control. The fundamental work of machinery and equipment operating status monitoring and fault is how to monitor and external signs, and get the correct information characteristic parameters for analysis and identification. The development bottleneck of troubleshooting is the lack of fault samples. Since SVM can achieve a good classification promotional purposes in small samples, more and more scholars applied SVM in this area research. Reference [3] discussed the classification algorithm of support vector machine in the field of fault diagnosis. Reference [4] used SVM model to distinguish the fault spectrum and the correct spectrum, and its result showed that the proposed model can correctly classify a variety of electronic equipment failures. Reference [5] based on the multi-class SVM algorithm, DAGSVM, proposed a variable increment iterative algorithm using misclassification of samples. Reference [6] combined the wavelet packet decomposition energy structure with the SVM classifier to fault diagnosis.

In this paper, the vibration intensity value is used as the basis for the overall state of the

mechanical system failures forecasts. Take the mechanical system test bed vibration intensity daily average as a sample, a total of 94 sample data, the data is divided into two groups, the first 60 sets of samples are defined as training data, the other samples are defined as the testing data. The parameters settings of SVM is shown in Table 1.

Table 1 The parameters settings of SVM

Parameters	Value
$\varepsilon$	0.0001
$m$	17
$C$	1000
Kernel function	Radial basis function

The one step ahead prediction and twenty steps ahead prediction of SVM model and Auto regression (AR) model are respectively implemented. The prediction results are shown in Fig.1 and Fig.2.

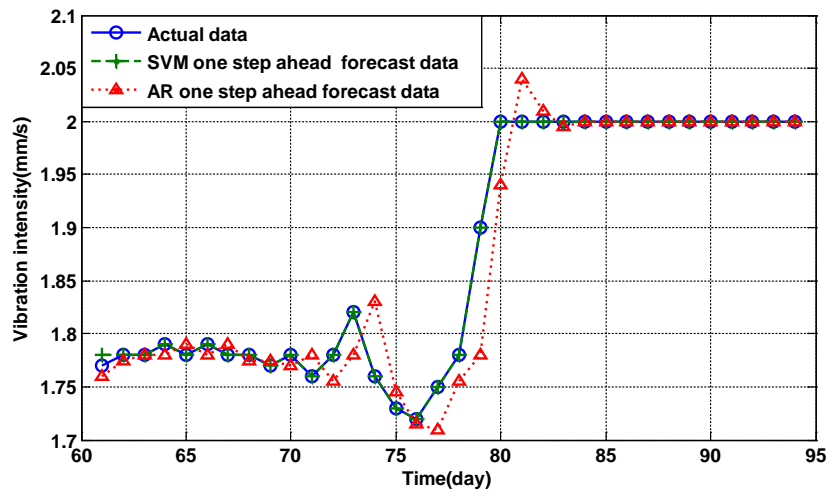


Fig.1 The prediction results of one step ahead

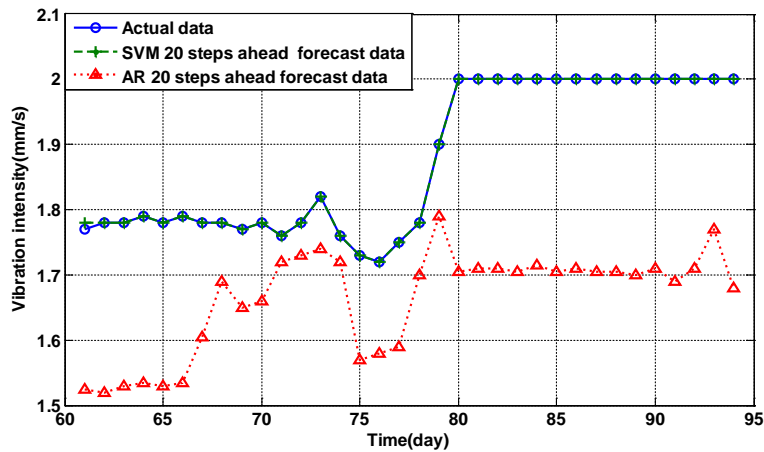


Fig.2 The prediction results of twenty steps ahead

As can be seen from the above, when multi-step prediction, the prediction accuracy of SVM model is significantly better than the AR model, we can see that SVM model has good generalization ability.

## Conclusions

The main advantage of support vector machine approach are: (1) Specifically for small sample. The real problems through non-linear transformation to convert to high-dimensional feature space, to ensure that the SVM has good generalization ability, while addressing the curse of dimensionality; (2) The AR model and SVM model experiments table vibration intensity forecast show that SVM

model has a high long-range prediction accuracy.

Although the application research of SVM has yielded good results, there are some issues that need further study: (1) The performance of SVM depends on the choice of kernel function, how to choose the right kernel function to meet the non-linear non-mechanical equipment stable condition is worthy of discussion; (2) Further research is needed to improve the training of SVM algorithm slow to adapt to the requirements of real-time fault diagnosis. It will be achieved that SVM in the fault diagnosis and trend forecasting machinery and equipment will have good prospects.

## References

- [1] Vapnik V. The natural of statistical theory. New York: Springer-Verlag, 1995.
- [2] Nello C, John S T. An introduction to support vector machines and other kernel-based learning methods. Cambridge: Cambridge University Press, 2000.
- [3] Hsu Ch W, Lin Ch J. A comparison of methods for multi2 class support vector machines. IEEE Transactions on Neural Networks. Vol. 13 (2002), p. 415-425.
- [4] Leng Chongfu, Jiang Yaquun, Huang Chun. Implementation of Dual Support Vector Machine Based Adaptive Single-Phase Reclosure of Transmission Line. Vol. 38 (2014), p. 2822-2827.
- [5] Hu Haigang, Pang Honglei, Fu Lei. Fault diagnosis methods about torsional vibration of shaft based on DAGSVM. Vol. 37 (2011), p. 141-147.
- [6] Li Ziguo, Hao Wei, Li Lingjun. Method of fault diagnosis based on wavelet packet decomposition and support vector data description. Vol. 29 (2007), p. 365-369.