Proceedings of the 6th World Congress on Intelligent Control
and Automation, June 21 - 23, 2006, Dalian, China

Fault Diagnosis of Blast Furnace Based on SVMs*

Anna Wang, Lina Zhang, Nan Gao and Mingshun Wang

College of Information Science and Engineering
Northeastern University
Shenyang, Liaoning Province , China
wanganna@mail.neu.edu.cn

Abstract **–For the complexity and confusion in fault diagnosis of blast furnace, a new artificial intelligent algorithm is presented to solve this problem. The Support Vector Machines (SVMs) are one type of large margin classifier based on statistical methods. With the property of dealing with high dimension data, studying small quality of samples and training large data sets, it is feasible to use it to make classifier. By using one-against-one strategy, construct it with different kernel functions, and select the best one for the classifier. The simulation results show that using polynomial function is superior to the others. Besides, compared with expert system and neural networks, SVMs have a better performance on recognizing patterns, capabilities of fault-tolerance and generalization.**

Index Terms –**blast furnace, fault diagnosis, SVMs, kernel function**

I. INTRODUCTION

At present, the applications of artificial intelligence on fault diagnosis of blast furnace are mainly based on expert system and neural networks [1]. However, there are many difficulties in acquiring and expressing knowledge [2], and summing up the experience perplexes the system; the neural networks also consists of the disadvantage of generalization, determination numbers of concealed level and the problem of studying large-scale samples [3].The most important problem is the veracity of these two methods always unsatisfied. Hence, the diagnosis system of blast furnace hasn't made a great step.

The support vector machines (SVMs) was originated as a new technique of data mining by Vapnik in 90's of twenty centuries [4]. In recent years, c-SVC, v-SVC, one-class SVM, c-SVR and v-SVR appeared successively. This method has the capability of dealing with high dimension data, optimal policy of computing and fault-tolerance. It leads the decrease of training samples and nicer adaptation of new data. For many practical problems, including pattern matching and classification function approximation, data clustering and forecasting, SVMs have drawn much attention [5]. Now it has been applied into many important fields such as spaceflight, automation industry, petroleum industry, etc. Aimed at the disorder and abnormality in the process of pudding, the method based on SVMs is used for real-time diagnosis. By simulated large real data, we acquired better results.

II. SUPPORT VECTOR MACHINES

Support Vector Machines use optimal method to solve the machine study problem [6, 7]. It is based on the Statistics Learning Theory (SLT). Its principle is the minimum of structure risk that is, minimizing a bound on the generation error of a model, which is different with the minimum of mean square error over the data set in neural networks. Hence, it settles the problems of large scale of training samples and intrinsically training difficulty.

SVMs could switch the linear problem from non-linear issue, and map it into a higher dimension space (decision surface). Then the hyperplane is constructed in this feature space that bisets the two categories and maximizes the margin of separation between itself and those points lying nearest to it (called the support vectors), as shown in Fig.1. This decision surface can then be used as a basis for classifying vectors of unknown classification. In the end, it could find the optimal solution to the problem in the condition of finite samples, and realize perplex classify.

Given a training set of instance-label pairs $(x,y), i=1,...n$, where $x \in R$ and $y \in \{+1,-1\}$, defining the hyperplane equation as $\omega \cdot x + b = 0$, for the sake of maximizing the classify margin, $\min \|\omega\|^2 / 2$ needs to be calculated.

Subject to:

$$y_i(\omega \cdot x_i + b) - 1 \geq 0 \quad i=1,...,n \tag{1}$$

The dual optimization question of above is:

$$\left.\begin{aligned} &Q(\alpha)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1,j=1}^{n}\alpha_i\alpha_j y_i y_j (x_i - x_j) \\ &\sum_{i=1}^{n} y_i\alpha_i = 0 \quad \alpha_i \geq 0, i=1,...,n \end{aligned}\right\} \tag{2}$$

* This work is partially supported by the open fund , the key laboratory of process industry, automation of MOE, China

1-4244-0332-4/06/$20.00 ©2006 IEEE

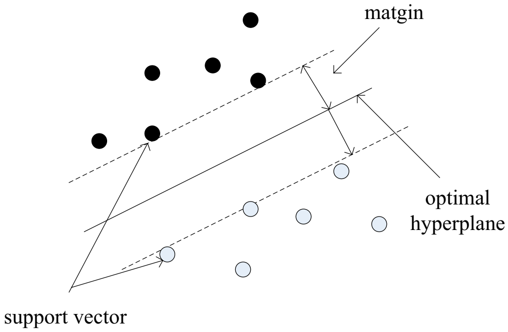

Fig.1 optimal classification hyperplane

Following the Kuhn-Tucker, the optimal solution should subject to:
$$\alpha_i y_i(\omega \cdot x_i + b) - 1 = 0 \quad i = 1,\ldots,n \tag{3}$$

For the problem is always linear inseparably, here add a loose parameter ξ_i, $i=1,\ldots,n$, which should not be negative and a penalty parameter C. Then the question above turns to:

$$\left.\begin{aligned} &\min \|\omega\|^2 / 2 + C\sum_{i=1}^{n}\xi_i \\ &(\omega \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, i=1,\ldots,n \end{aligned}\right\} \tag{4}$$

The corresponding optimal decision-making function is:

$$f(x) = \mathrm{sgn}[\sum_{i=1}^{n} y_i \alpha_i (x \cdot x_i) + b] \tag{5}$$

Among all the solutions to this equation, the samples making α nonzero are called support vector. According to these vectors, classify could be computed easily and correctly.

When the problem is non-linear, the algorithm allows mapping the samples x into a higher dimension space H. Following the Mercer, linear classifier is constructed by using different kernel function. Then the optimal decision-making function turns to:

$$f(x) = \mathrm{sgn}[\sum_{i=1}^{n} y_i \alpha_i K(x_i, x_j) + b] \tag{6}$$

There are four popular kernel functions:

1) Linear: $K(x_i, y_j) = x_i \cdot y_j$
2) Polynomial: $K(x_i, y_j) = [(x_i \cdot y_j) + 1]^d$
3) Radial basis function: $K(x_i, y_j) = \exp(-\gamma |x - y|^2)$
4) Sigmoid: $K(x_i, y_j) = \tanh(v(x_i^T \cdot x_j) - c)$

where d, γ, v and c are kernel parameters.

III. METHOD OF SOLUTION

During the process of puddling, a lot of process parameters will generate, such as top temperature, pressure, ventilation property, material velocity, etc. These parameters could be detected and collected through the sensors and then transferred to computers. Any a little deviant data can result in different fault, the operators used to have to observe these data continuously. Because the data are large-scale and the machines are normal in most occasions, it wastes lots of manpower.

Nowadays, we use intelligent system to let machine predict the tendency of disorder and abnormality by integrating and analyzing these data sets and we have gained delightful results. To make the decision more accuracy, we propose a new approach to solve this problem.

The former models are mainly based on expert system or neural networks. However, these methods demand lots of experience or samples to study, and they could only predict several classical sorts. Besides, the accuracy of results is also unsatisfactory. As to the new sort, their adaptability is not so well in common. For the same samples, using SVMs algorithm, its superiority represents not only the raised veracity and capability of identifying more sorts and more complex tendency, but also reducing the quantity of training data. All these advantage shows making good use of SVMs may be a better way to solve the diagnosis problem of blast furnace.

Apparently, the diagnosis of blast furnace is a multi-class identification matter. To solve it effectively, it is necessary to reconstruct the SVMs model. There are three common approaches: one-against-all, one-against-one and DAGSVM algorithm [7]. According to a great many of experience, Chih-Jen Lin concludes that one-against-one method is superior to the others in actual application although it might be a little complex.

The basic thought of this algorithm is constructing $k(k-1)/2$ binary classifiers. Between every two categories, system will train a SVM to distinguish themselves. For training data from the i th and j th classes, the following binary classification problem is to be solved:

$$\min_{\omega^{ij}, b^{ij}, \zeta^{ij}} \frac{1}{2}(\omega^{ij})^T \omega^{ij} + C(\sum_t (\zeta^{ij})_t) \tag{7}$$

subject to:

$((\omega^{ij})^T \phi(x_t)) + b^{ij} \geq 1 - \zeta_t^{ij}$, if x_t in the i th class.

$((\omega^{ij})^T \phi(x_t)) + b^{ij} \leq 1 + \zeta_t^{ij}$, if x_t in the j th class.

$\zeta_t^{ij} \geq 0$.

where ϕ is kernel function, C is penalty parameter of error term, ζ is loose parameter.

While predicting the testing samples, it used a voting strategy that each binary classification is considered to be a voting where votes can be cast for all data pointers x; in the end, point is designated to be in a class with maximum number of votes. In case that two classes have identical votes, though it may not be a good strategy, it would simply select the one with the smallest index.

In the course of designing model, it is very important to select a proper kernel function. For different kernel has its own characters, they would brings on a various performance on the same problem. In this paper, SVMs classifier is constructed by the four kernels respectively. According to the test results, select the best one to solve the question.

TABLE I

KERNEL COMPARISON

	Training samples	Support vector	Training Accuracy (%)	Testing data	Testing Accuracy (%)
Linear	49	39	97.96	578	96.02
RBF	49	49	100	578	90.83
Sigmoid	49	49	20.41	578	16.96
Polynomial	49	38	100	578	96.54

IV. SIMULATION

This paper refers the data from 5# blast furnace of a steel enterprise. We collected 578 set of real-time data and select several important control parameters as judgments criterions, such as furnace temperature and press, material velocity, venting quality, etc. Through simulation, we got the comparisons among kernel functions, methods, and the final results of classifier.

A. kernel function study

According to the same training samples and testing data, different kernel functions are used to train them. The results are shown in Table1.

By comparing different kernel function, it can be seen that every kernel has its own effect on the same problem [9]. From the table, it is clear that sigmoid kernel can hardly recognize any fault, but polynomial function has a satisfied performance on this problem. The difference between them is very significantly. Hence, selecting a proper kernel function is vital to the research. The following work is based on the polynomial function.

B. fault diagnosis study

To examine the feasibility of SVMs, construct the classifier using the selected function above. Through testing the real data, the results are listed below.

TABLE2

DIAGNOSIS RESULTS

	Training samples	Testing samples	Distortion rate (%)	Omission rate (%)	Accuracy (%)
Pipe fracture	49	578	0.3460	0.1730	
Cooler	49	578	0.5190	0.1730	
Warmer	49	578	0.1730	0.1730	
Hanging	49	578	0.3460	0	
Slip	49	578	0.1730	0.3460	96.54
Moving handicap	49	578	0.1730	0	
Centre gas flow	49	578	0.5190	0.1730	
Brim gas flow	49	578	0.1730	0	
Low stockline	49	578	0	0	
Normal	49	578	0	0	

By observing the results in table2, we could find that use SVMs to identify disorder and abnormality, the accuracy of identification is not dropped though the types of fault rise. The distortion rate and omission rate are all within 0.33%, and each type of signal diagnosis can attain 98%. Thus, it is feasible to construct SVMs classifier with polynomial function.

C. algorithm comparison

To illuminate the validity of SVMs algorithm, we have compared SVMs with expert system and neural networks method [10, 11]. We also used the same data but different quantity of studying samples for the demands of practice. The simulation results are listed in table3.

TABLE3

METHOD COMPARISON

	Studying samples	Testing samples	Accuracy (%)
Expert system	198	578	91.46
Neural networks	120	578	93.23
Support vector machines	49	578	96.54

From the table, we could find the studying samples of SVMs are the smallest, but the diagnosis accuracy is the topmost. Thus, the algorithm we present is valid.

According to the analysis above, we conclude that using SVMs method is feasible, suitable and effective. On the basis of it, we constructed the classifier with polynomial kernel function, and the accuracy of diagnosis is raised. We can also use this method to judge more fault and predict unhealthy tendency.

V. CONCLUSION

According to the simulation, it is easy to find that using SVMs, system could gain better results on accuracy, in the condition of smaller training samples and more parameters. Besides, it could identify more sorts of fault and tendency. For the using of basic SVMs algorithm, there are still many ways to improve and complement it, for instance, selection of SVMs and the disposition before training. All these work would be studied continually.

REFERENCES

[1] J.C. Song, *Theory and Operation in Blast Furnace Puddling*, Beijing: Metallurgical Industry Press, 2005.

[2] C.Z. Cao, N. Wang, *Intelligent Technology*, Beijing: Tsinghua University Press, 2004.

[3] B. Samanta, "Gear fault detection using artificial neural networks and support vector machines with genetic algorithms", *Mechanical Systems and Signal Processing*, vol.18, no.3, pp.625-644, 2004.

[4] A. Shilton, M. Palaniswami, D. Ralph, et al, "Incremental training of support vector machines", *IEEE Trans. Neural Networks*, vol.16, no.1, pp.114-131, January 2005.

[5] G. X. Zhang, Z.X. Gao and Y.J. Gu, "A hybrid classifier based on rough set theory and support vector machines", *LNCS*, vol.3613, pp.1287-1387, 2005.

[6] N.Y. Deng, Y.J. Tian, *Support Vector Machines: A New Method in Data Mining*, Beijing: Science Press, 2004.

[7] C.W. Hsu, C.J. Lin, "A comparison of methods for multi-class support vector machines", *http://citeseer.nj.nec.com/hsu01comparison.html*, 2001.

[8] V.N. Vapnik, *Statistical Learning Theory*, New York: Wiely, 1998.

[9] K.R. Muller, S. Mika, G. Ratsch, et al, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Networks*, vol. 12, no. 2, pp. 181-201, Mar. 2001.

[10] J.G. Li, H.F. Guo and W. Zhou, "The design and realization of the decision system for the blast furnace abnormal state", *Journal of Northeastern University*, vol.18, no.2, pp.178-182, Apr.1997.

[11] Y.T. Wang, J.B. Xia, J.C. Zhou, et al, "Neuro-fuzzy network expert system for parameter learning of blast furnace", *Journal of Iron and Steel Research*, vol.11, no.5, pp.72-76, Oct.1999.