TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.5, May 2014, pp. 4056 ~ 4062
DOI: http://dx.doi.org/10.11591/telkomnika.v12i5.4364

Application of Support Vector Machine Model in Mine Gas Safety Level Prediction

Huaping Zhou[1], Ruixin Zhang[2]
[1]School of Computer Science and Engineering Anhui University of Science and Technology, HuaiNan, China
[2]Faculty of Resources & Safety Engineering, China University of Mining & Technology, Beijing, China

Abstract

For the limitation of traditional information fusion technology in the mine gas safety class predicition, an intelligent algorithm is proposed in which Genetic Algorithms is adopted to optimize the parameters of the least squares support vector machine and establishes a multi-sensor information fusion model GA-LSSVM which overcomes the subjectivity and blindness on parameters selection, and thus improves its classification accuracy and convergence speed. Experimental results show that compared to the least squares support vector machine model not been optimized and the least squares support vector machine model optimized by the grid searching algorithm, GA-LSSVM model can be a good solution on the issue of the high-dimensional, nonlinear and uncertainty of the small sample in coal mine underground environment level evaluation.

Keywords: *information fusion, genetic algorithms, least squares support vector machine, parameter optimization cross validation*

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

The information fusion methods in mine safety are mainly involves Bayes estimation theory [1], fuzzy information fusion [2, 3], Vague sets information fusion [4], adaptive estimation methods in batches [5], DS evidence theory[6], rough set [7], neural networks or a combination of both methods [8], which complete respectively one level fusion and decision level fusion. Decision-level information fusion method has advantages and disadvantages. DS evidence theory is difficult to find the more reasonable basic probability assignment for specific circumstances. The membership value of Fuzzy Information Fusion is a single value which may not also indicate the evidence of Supporting and opposing and is not the best theory for dealing with the uncertainty. Vague sets that consider both membership and non-membership information, but target selection method is more difficult to determine. Nerve network information fusion algorithm has the shortcomings such as training slower, more difficult parameter selection, easy to over-fitting. So, it is difficult to adapt to the mine. In this paper, the above analysis, genetic algorithm optimizing least squares support vector machine (GA-LSSVM) information fusion optimization mode is proposed. Support vector machine solves the question by quadratic optimization, so the solution is global optimal solution, avoiding local minima. Least squares support vector machine is a form of the model which can improve the training speed and classification speed of model. The parameters of support vector machine have a greater impact on the model, so this article adopts genetic algorithms to optimize least squares support vector machine model parameters.

2. Least Squares Support Vector Machine

Support Vector Machine (SVM) [9] constructs the optimal separating hyperplane, and makes the points of the training set away from it far as possible. The nonlinear question is solved through the introduction of nonlinear mapping mapped into a high dimensional feature space, thus transformed to a linear problem. The construction of the optimal separating hyperplane is divided into linearly separable and linearly inseparable. Support vector machine is initially presented in the case of linear separable.

Received October 17, 2013; Revised December 28, 2013; Accepted January 15, 2014

Assuming the training sample set $T=\{(x_1,y_1),...,(x_l,y_l)\}\in(x\times \mathrm{y})^l$, among it, $x_i\in x=R^n, y_i\in y=\{-1,+1\}, i=1,...,l$; constructing and solving the optimization problem for the variables ω and b , the objective function is:

$$\min_{w,b}\frac{1}{2}\|w\|^2 \qquad \text{s.t. } y_i[(w\cdot x_i)+b]\geq 1, i=1,...,l \tag{1}$$

The $w*$ and $b*$ is the optimum solution. The optimal separating hyperplane constructed is: $(w*\cdot x)+b*=0$, Get the following decision function:

$$f(x)=\mathrm{sgn}[(w*\cdot x)+b*] \tag{2}$$

When the training set are linearly inseparable, the error must exist which is referred to as a ξ.According to the structural risk minimization principle, the introduction of slack variables, denoted by ξ_i .s.t. $\xi_i \geq 0$ .The constraints is relaxed: $y_i[(w\cdot x_i)+b]+\xi_i\geq 1$. $\sum_{i=1}^{l}\xi_i$ is adopted as a measure which describes the degree of misclassification of the training set.

While also ensure $2/\|w\|^2$ maximum. Therefore a penalty parameter C is introduced as the combination of these two target weights. The objective function becomes the following form:

$$\frac{1}{2}\|w\|^2+C\sum_{i=1}^{l}\xi_i \tag{3}$$

Introducing Lagrange multipliers α_i, α_i^* , η_i , and η_i^*, Establishing Lagrange function:

$$\begin{aligned}L=&\frac{1}{2}\|w\|^2+C\sum_{i=1}^{l}(\xi_i+\xi_i^*)-\sum_{i=1}^{l}\alpha_i(\varepsilon+\xi_i-y_i+<w\cdot x_i>+b)-\\&\sum_{i=1}^{l}\alpha_i^*(\varepsilon+\xi_i^*+y_i-<w\cdot x_i>-b)-\sum_{i=1}^{l}(\eta_i\xi_i+\eta_i^*\xi_i^*)\end{aligned} \tag{4}$$

Take partial derivative with respect to w,b, ξ_i and ξ_i^* and set to zero. Obtain the following form:

$$\begin{cases}w=\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)x_i\\ \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)=0\\ C-\alpha_i-\eta_i=0\\ C-\alpha_i^*-\eta_i^*=0\end{cases} \tag{5}$$

Original problem is transformed into its dual form:

$$\begin{cases}\min\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)<x_i\cdot x_j>+\varepsilon\sum_{i=1}^{l}(\alpha_i^*+\alpha_i)-\sum_{i=1}^{l}y_i(\alpha_i^*-\alpha_i)\\ \qquad s.t. \quad \sum_{i=1}^{l}(\alpha_i^*-\alpha_i)=0\\ \qquad 0\leq\alpha_i,\alpha_i^*\leq C\end{cases} \tag{6}$$

Solving the above convex quadratic programming problems, getting the optimal classification equation:

$$f(x) = \operatorname{sgn}[\sum_{i=1}^{l} y_i \alpha * < x_i \cdot x > + b*] \tag{7}$$

For nonlinear separable case, through the introduction of kernel functions meeting Mercer condition, transform into linear problem in high-dimensional space through nonlinear action. In high-dimensional space seek the optimal separating hyperplane. Kernel functions to meet the conditions mainly are as Table 1:

Table 1. Kernel Function

Polynomial kernel function	$k(x,z)=(<x,z>+c)^d, d\in Z^+, c\geq 0$
Gaussian kernel function	$k(x,z) = \exp(-\frac{\lVert x-z \rVert^2}{2\sigma^2}), \sigma > 0$
Exponential radial basis kernel	$k(x,z) = \exp(-\frac{\lVert x-z \rVert}{2\sigma^2}), \sigma > 0$
B-spline kernel	$k_1(x,z) = k_1(x,z;t_1,...,t_m) = \sum_{i=1}^{m} (x-t_i)_+^p (z-t_i)_+^p, \forall x,z \in R$
Fourier kernel	$k_1(x,z) = \frac{1-q^2}{2(1-2q\cos(x-z)+q^2}, \forall x,z \in R$
RBF kernel	$k(x,z)=\exp(-\frac{\lVert x-z \rVert}{2\sigma^2}), \sigma>0$

σ is the width of the radial basis function. The parameters of radial basis kernel function are less with a simple calculation, and the performance is better and it has more common applications. So this paper uses RBF kernel function as a support vector machine kernel function. After kernel function is introduced, its decision function is:

$$f(x) = \operatorname{sgn}[\sum_{i=1}^{l} y_i \alpha_i^* K(x_i, x) + b*] \tag{8}$$

Suykens [10] announces least squares support vector machine model to the public for the first time in the last century. Standard support vector machine model is the inequality constraints $y_i[(w \cdot x_i) + b] \geq 1, i = 1,...,l$ .But the least squares support vector machine is equality constraints which is s.t. $y_i[(w \cdot x_i) + b] = 1, i = 1,...,l$. Thus, solving linear equations instead of solving quadratic programming problems, thereby reducing the support vector machine model computational complexity, speeding up the solving speed [11, 12].

For SVM multi-class classification problems, this paper uses paired classification algorithms, namely one-against-one algorithm (abbreviated 1-a-1 SVM). Training a classifier each two types, for a problem of n-type, there are n (n-1) / 2 category function. Each classifier is to take any data of two categories to train the [13, 14]. For the training between class i and class j, you need to solve the following two types of classification:

$$\min_{w^{ij}, bi^j, \xi^{ij}} \frac{1}{2}(w^{ij})^T w^{ij} + C\sum_t \xi_t^{ij}$$

$$\begin{cases} (w^{ij})^T \varphi(x_t) + b^{ij} \geq 1 - \xi_t^{ij}, y_t = i, \xi_t^{ij} \geq 0 \\ (w^{ij})^T \varphi(x_t) + b^{ij} \leq -1 + \xi_t^{ij}, y_t = j, \xi_t^{ij} \geq 0 \end{cases} \tag{9}$$

In this paper, "the biggest referendum act" is adopted determine which category the sample is, that is, each two classifier both judges the categories of the sample, the class of the most votes is the belonging class of the unknown sample. Classify the unknown sample x, the decision function is:

$$f(x) = \mathrm{sgn}[(w^{ij})^T \phi(x_t) + b^{ij}] \quad (10)$$

3. GA-LSSVM Prediction Model

Genetic algorithms (GA) is first proposed by John Holland in the 1860s. The intelligent search of genetic algorithm is adopted in the process of parameter selection of support vector machine algorithm in this paper and find the optimal parameter. Looking for the optimal support vector machine model for the sample of the coal mine. By comparing with grid search algorithm, genetic search algorithm can quickly obtain the satisfactory parameters.

SVM model includes the qualitative options and quantitative options. The former includes how to identify specific support vector machine algorithm and Kernel. The latter is support vector parameter selection. LSSVM parameter choice includes: kernel function parameters and the error penalty parameter . The error penalty parameter of different SVM is named different. The name of different kernel function parameters is not the same. However, the role and significance is both the same. For convenience of description herein, penalty parameter and kernel parameters are expressed using γ and σ.

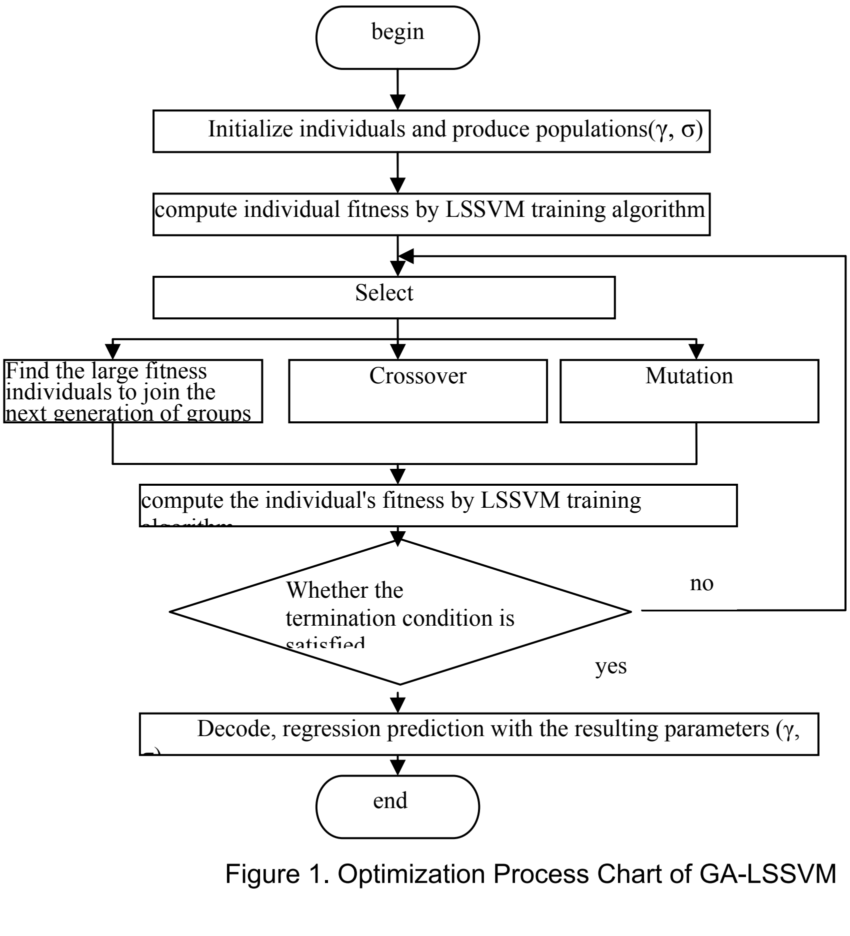

Figure 1. Optimization Process Chart of GA-LSSVM

Kernel parameter selection of least squares support vector machine is directly related to the learning performance and generalization ability of least squares support vector machine. The parameter selection methods commonly used are mainly cross-validation method and nuclear calibration. Cross-validation method requires a lot of computing, to determine the optimum parameters. Especially when the number of parameters is large, it will take a lot of time to strike the optimal solution. Nuclear calibration method is related to much knowledge and research of nuclear matrix, so it is more difficult to achieve it. To compensate for the insufficient of existing parameter selection algorithm, the SVM model predicts the gas level with the genetic algorithm. The algorithm is not only able to achieve a global search, and search speed can be guaranteed. The practical application shows: the improved support vector machine parameters selection algorithm based on genetic algorithm can get optimal operating parameters of non-stationary time series and nonlinear prediction model. It is a proven method of selecting SVM kernel parameters. Optimizing process shown in Figure 1.

4. Simulation Results

4.1. Mine Gas Level Evaluation Model Data

Selecting 200 samples from more than 20 coal mines in China in this paper, of which 125 samples as SVM training samples, the establishment of model training, 75 samples as test samples. Comprehensive literature [1] and the literature [2], extracting four characteristic parameters, which are most relevant to gas accidents, dust, temperature, wind speed, gas content concentration as input dimension of support vector machine. According to the mine safety rules gas state safety class were divided into safer, more secure, general safety, more dangerous and hazardous, respectively values 0,1,2,3 and 4. This will adopt GA-LSSVM classification algorithm and training samples to establish predictive model. In order to test the correctness and the generalization performance of the model, training samples and test samples selected are disjoint. And ensure that the test sample contains all grades of coal mine safety. Then for the given characteristic parameters the model can make an intelligent decision for environmental conditions underline the coal mine. Since the data sample values of 4 feature vectors vary greatly, in order to make the different dimension and magnitude of feature vectors minimize the impact on the prediction model, thus ensure the accuracy of SVM prediction model, the data samples should be normalized and be converted to a value between 0 and 1, the mean method is adopted in this paper.

For a data series, its average value is divided by all of this data series, the mean value of the data series is a new sequence after treatment. Suppose the original series denoted as x0=(x0(l),x0(2), …, x0(n)). The average denoted as $\overline{x_0}$. The original data sequence x0 which is averaged is the data sequence y0. Calculated as follows:

$$y_0 = \{y_0(1), y_0(2), \cdots, y_0(n)\} = \left\{ \frac{x_0(1)}{\overline{x_0}}, \frac{x_0(2)}{\overline{x_0}}, \cdots, \frac{x_0(n)}{\overline{x_0}} \right\} \tag{11}$$

4.2. Analysis of Experimental Results

Table 1. Comparative Analysis

Training time	Prediction Model	Parameter γ	σ	Training time	Optimal number of iterations	Classification accuracy（%）
1	LSSVM	0.125	0.5	18.78		75.33
	GS-LSSVM	1	0.5	15.47	50	82.67
	GA-LSSVM	16.125	12.625	10.343	20	85.33
2	LSSVM	0.5	0.25	18.87		75.33
	GS-LSSVM	8.37	6.725	19.59	100	74.67
	GA-LSSVM	42.25	1.37	11.75	20	89.33
3	LSSVM	25.25	10.45	17.47		78.67
	GS-LSSVM	62.35	12.345	16.34	140	89.33
	GA-LSSVM	60.26	60.675	9.56	25	91.33
4	LSSVM	60.75	0.125	16.48		75.33
	GS-LSSVM	80.67	25.38	18.27	155	83.33
	GA-LSSVM	66.57	0.0093	10.56	20	89.33

LIBSVM is adopted as training and testing tools of support vector classification model in Matlab software platform in the paper. In order to better verify the validity of the predictive model GA-LSSVM and remove the chance of prediction results, 125 training samples were selected and to obtain the optimal parameters γ and σ ,not parameters optimized least squares support vector machine (LSSVM), Meshing algorithm optimization Least squares support vector machine (GS-LSSVM) and genetic algorithm least squares support vector machine (GA-LSSVM) adopted by the paper, training three times each prediction model, GS-LSSVM model and GA-LSSVM model are adopted respectively and each prediction model is trained three times. GS-LSSVM model and GA-LSSVM model, respectively, obtain the optimal parameters γ and σ, training time, the optimal number of iterations. LSSVM model parameters are for the artificial random assignment, and then each model is validated and the classification accuracy is obtained with the 75 test samples. Results are shown in Table 1.

It can be seen from Table 1 that the average training time of LSSVM, GS-LSSVM and GA-SVM is 16.48s, 18.27s and 10.56s respectively, the average number of iterations of GS-LSSVM, GA-SVM is 155 and 20 times respectively, the he average classification accuracy is 75.33%, 83.33% and 89.33%. Classification accuracy rate of without the parameter optimization LSSVM model is the lowest. The parameters γ and σ have a greater impact on the classification performance of support vector machine. The training time of GS-LSSVM is the longest, so it is bound to affect the classification efficiency in the case of large amount of training samples. All particles of genetic algorithm converge quickly to the optimal solution, and the classification accuracy rate is the highest, reaching 89.33%, 12.666% higher than the LSSVM model, 6.532% higher than GS-LSSVM model. Therefore, the genetic algorithm optimization least squares support vector machine classifier prediction model proposed in the paper has better generalization ability and higher classification capability.

5. Conclusion

Extracting four characteristic parameters, which are most relevant to gas accidents, dust, temperature, wind speed, gas content concentration as the factors of coal mine environment grade evaluation. Dividing environment grade into 5 grade that are safer, more secure, general safety, more dangerous and hazardous. Adopting genetic algorithms to optimize the parameters of least squares support vector machine model to establish the GA-LSSVM model of coal mine environment classification. Compared to without the parameter optimization model (LSSVM), grid search algorithm to optimize the least squares support vector machine model (GS-LSSVM), this model has a higher processing speed and higher classification accuracy.

Acknowledgements

This work is supported by Anhui Provincial Natural Science Foundation Of Universities Key Projects (Grant No. KJ2010A083), the Natural Science Foundation of the Anhui Higher Education Institutions of China (KJ2012A099),and is aslo supported by the National Natural Science Foundation of China (Grant No. 51174257).

References

[1] Fu Hua, Zhao Dan, Zhou Fang. Research on application of RS-RBF information fusion in gasmonitorin. *Transducer and Microsystem Technologie*. 2009; 28(12): 30-32.
[2] LIU Bing,LI Hui,XING Gang. Fuzzy Information Fusion Target Recognition Based on Weighted Evidence Theory. *Computer Engineering.* 2012; 38(15): 172-174.
[3] JIN Hai. Application of Multisensor Fuzzy Information Fusion Algorithm in Coal Mine Gas Monitoring. *Coal Technology.* 2012; 31(8):82-84.
[4] FU Hua, GAO Ting, YANG Xin. Application of vague set information fusion theory in mine safety monitoring. Application Research of Computers. 2009; 26(6): 2282-2284.
[5] SUN Ke-lei, Qin Ru-xiang. Study of multi-sensor data fusion based on adaptive batch estimation algorithm for gas monitoring. *Transducer and Microsystem Technologies*. 2011; 30(10): 47-49.
[6] FU Hua, LI Bo, XUE Yong-cun. Analysis of underground monitor method based on D-S decision-making data fusion. *Transducer and Microsystem Technologie*. 2007; 26(1): 27-29.
[7] DU Xiao-kun, CHEN Feng. Application of information fusion in coal mine monitoring system. *Transducer and Microsystem Technologie*. 2010; 29(7): 124-126.
[8] Vapnik VN. The nature of Statistical learning theory. Berlin:Spring-Verlag. 1995.
[9] Suykens JAK, Vandew alle J. Least squares support vector machine classifiers. *Neural Processing Letters*. 1999; 9(3): 293-300.

[10] Sana M Vieira, Luís F Mendonça, Gonçalo J. Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. *Applied Soft Computing.* 2013; 13(8): 3494-3504.
[11] Nghe Wang, Xinyi Zhao, Baotian Wang. LS-SVM and Monte Carlo methods based reliability analysis for settlement of soft clayey foundation. *Journal of Rock Mechanics and Geotechnical Engineering*. 2013; 5(4): 312-317.
[12] En yongqi. LS_SVM Parameters Selection Based on Hybrid Complex Particle Swarm Optimization. *Energy Procedia*. 2012; 17: Part A: 706-710.C
[13] Heng-Lung Huang, Jian-Fan Dun. A multiple kernel framework for inductive semi-supervised SVM learning. *Neurocomputing*. 2012; 90(1): 46-58.
[14] Ehua Liu, Hui Qian, Guang Dai, Zhihua Zhang. An iterative SVM approach to feature selection and classification in high-dimensional datasets. *Pattern Recognition*. 2012; 46(9): 2531-2537.