Application of SVM in Analyzing the Headstream of Gushing Water in Coal Mine

YAN Zhi-gang, ZHANG Hai-rong, DU Pei-jun

School of Environment and Spatial Informatics, China University of Mining & Technology, Xuzhou, Jiangsu 221008*, China*

Abstract: To recognize the presence of the headstream of gushing water in coal mines, the SVM (Support Vector Machine) was proposed to analyze the gushing water based on hydrogeochemical methods. First, the SVM model for headstream analysis was trained on the water sample of available headstreams, and then we used this to predict the unknown samples, which were validated in practice by comparing the predicted results with the actual results. The experimental results show that the SVM is a feasible method to differentiate between two headstreams and the H-SVMs (Hierachical SVMs) is a preferable way to deal with the problem of multi-headstreams. Compared with other methods, the SVM is based on a strict mathematical theory with a simple structure and good generalization properties. As well, the support vector W in the decision function can describe the weights of the recognition factors of water samples, which is very important for the analysis of headstreams of gushing water in coal mines.

Key words: support vector machine; gushing water; headstream recognition; H-SVMs

CLC number：P6411.42

1 Introduction

Water gushing is one of the serious natural disasters in the coal mine industry. The recognition of headstreams of gushing waters is the foundation of the prevention and control of gushing water, which includes dewatering, a declining hydraulic pressure and injection of slurry. Hydrogeochemistry technology is effective in distinguishing headstreams of gushing water[1–7], on the basis of which the methods for recognizing headstreams can be reduced to five types: (1) expert's experience[4], (2) multivariate statistics[1,6–7], (3) degree of grey incidence[2], (4) fuzzy comprehensive evaluation[3] and (5) ANN(artificial neural network)[5]. Among these, the expert's experience method depends on experts who are acquainted with the hydrogeological conditions in coal mines, with no uniform decision rules. Multivariate statistical methods can derive decision rules by clustering and regression analyses based on samples of gushing water whose headstreams are known. The rules are expressed as regression equations and clustering distances. Compared with the first type, this method has a rigorous mathematical theory, is more scientific and less dependent on experts. But its accuracy critically depends on the choice of the model. If an incorrect regression equation is selected, it is hard to arrive at a right conclusion. The degree of grey incidence mainly evaluates the correlation between predicted results and available samples and then determines the most likely pattern to which the predicted samples belong, based on correlation coefficients. The superiority of this method lies in the fact that it is brief and easily used and can deal with some complicated instances, such as multi-headstreams and mixed headstreams, but this method has a theoretical shortcoming when calculating and deducing the degree of correlation[8], which very much affects its application. Fuzzy comprehensive evaluation is a kind of analytical method based on fuzzy mathematics and a mixture between headstreams will lead to uncertainties. Fuzzy mathematics is a useful tool to dispose of uncertainty, but the method still depends on expert's experience in the choice of membership functions and fuzzy inference rules[10]. ANN can effectively deal with nonlinear, incomplete and fuzzy decision-making problems. It is a novel method to recognize the headstream of gushing water, but there are some inconclusive theoretical problems and the parameters of ANN are still determined by experience.

Except for the expert experience method, the above methods are all based on traditional statistical theories. They describe the ultimate characteristics of the sample set when the set capacity approaches infinity, but the maximum of the training set in references [1–7] is 12, which is far from the theoretical precondition. This is an essential limitation of the above methods[10].

Since the 1990's, the theoretical research on machine learning on limited samples has come of age and a more perfect theoretical system — SLT (statistical learning theory) has been developed. Based on

Received 20 May 2006; accepted 15 July 2006

Project 40401038 supported by the National Natural Science Foundation of China and 2003047 by the Top 100 Outstanding Doctoral Dissertation Foundation of China

Corresponding author. Tel:+86-13776786832; E-mail address: zhg-yan@163.com

SLT, the SVM (Support Vector Machine) has become a new popular research area in the field of machine learning and obtained a series of satisfactory results in classification and regression problems.

Taking the data provided in references [5–6] as training samples, SVM was used to analyze the headstream of gushing water in coal mines. The results show that the application of SVM to predict headstreams is feasible and preferable.

2 Introduction to SVM

SVM was proposed by Vapnick and his cooperator at the AT&T Bell Laboratory[9]. With a view of analyzing the headstream of gushing water in coal mines, the SVM is introduced as follows[9–10].

According to pattern recognition problems, one must find a calculable decision function $y = f(x)$, $\boldsymbol{x} \in R^n, y \in \{-1,1\}$. For a given training set of k samples $(x_1, y_1), (x_2, y_2), \cdots, (x_k, y_k), x \in R^n, y \in \{-1, 1\}$, we need to find a hyperplane which can separate the training set into two classes, that is $\boldsymbol{Wx} + b = 0$, $\boldsymbol{W} \in R^n, b \in R$. The corresponding decision function is:

$$f(x) = \sin(\boldsymbol{Wx} + b) \tag{1}$$

The hyperplane is subject to the following condition:

$$y_i[\boldsymbol{W}x_i + b] > 0,\ i = 1,\ 2,\ \cdots,\ k \tag{2}$$

There should be a distance between each class and the hyperplane; a slack variable $\xi_i \geq 0$ is adopted to allow misclassification. Then the hyperplane becomes subject to:

$$y_i[\boldsymbol{W}x_i + b] \geq 1 - \xi_i,\ i = 1,\ 2,\ \cdots,\ k \tag{3}$$

The optimal hyperplane is defined to maximize the minimum distance from the hyperplane to the nearest training sample of each class; thus, the classification problem is converted into a minimization problem which must meet the conditions $\xi_i \geq 0$ and Eq. (3):

$$\min : \tau(\boldsymbol{W}) = \frac{1}{2}\|\boldsymbol{W}\|^2 + C\sum_{i=1}^{k}\xi_i \tag{4}$$

In Eq. (4), the first parameter $\boldsymbol{W}$ maximizes the minimum distance from the hyperplane to both separated classes, the second parameter ξ_i minimizes the separation error, the constant C controls the tradeoff between the training error ξ_i and the separating margin $\boldsymbol{W}$. We can solve this constrained optimization problem by adopting Lagrange multipliers. The result is:

$$\begin{aligned} L(\boldsymbol{W}, b, \alpha, \gamma) = \frac{1}{2}\|\boldsymbol{W}\|^2 + C\sum_{i=1}^{k}\xi_i - \\ \sum_{i=1}^{k}\alpha_i(y_i(\boldsymbol{Wx} + b) - 1 + \xi_i) - \sum_{i=1}^{k}\gamma_i\xi_i \end{aligned} \tag{5}$$

Eq. (5) must meet the minimum of $\boldsymbol{W}, b$ and the maximum of α_i, γ_i. This way we can obtain Eq. (6) via the well known Karush–Kuhn–Tucker (KKT) conditions.

$$\left.\begin{aligned} \sum_{i=1}^{k}\alpha_i y_i = 0 \\ \boldsymbol{W} = \sum_{i=1}^{k}\alpha_i y_i x_i \end{aligned}\right\} \tag{6}$$

The optimization problem described by Eqs. (4–6) can be converted into the following dual problem, in which the maximization function is：

$$\left.\begin{aligned} \max : W(\alpha) = -\frac{1}{2}\sum_{i,j=1}^{k}\alpha_i\alpha_j y_i y_j (x_i \bullet x_j) + \sum_{i=1}^{k}\alpha_i \\ \text{s.t.}: 0 \leq \alpha_i \leq C,\ i = 1,\ 2,\ \cdots,\ k \\ \sum_{i=1}^{k}\alpha_i y_i = 0 \end{aligned}\right\} \tag{7}$$

The corresponding decision function is:

$$f(x) = \sin(\sum_{i=1}^{k}\alpha_i y_i (x_i \bullet x_j) + b) \tag{8}$$

Given the nonlinear classification problem, we adopt a kernel function ϕ to map the training data into a high dimensional feature space, in which a separating hyperplane must be found to make the training set linearly separable. The corresponding hyperplane is $W\phi(x) + b = 0$ and the decision function is：

$$f(x) = \sin(\sum_{i=1}^{k}\alpha_i y_i (\phi(x) \bullet \phi(x_i)) + b) \tag{9}$$

SVM theory only considers the inner product operation $K(x, y) = \phi(x) \bullet \phi(y)$ in a high dimensional feature space. Because it does not need to solve the function ϕ directly，the problem to calculate the product $\phi(x) \bullet \phi(x_i)$ in Eq. (9) is tackled skillfully. We call $K(x, y)$ a kernel function. The corresponding optimization problem in Eq. (7) is then changed into：

$$\left.\begin{aligned} \max : W(\alpha) = -\frac{1}{2}\sum_{i,j=1}^{k}\alpha_i\alpha_j y_i y_j K(x_i, x_j) + \sum_{i=1}^{k}\alpha_i \\ \text{s.t.}: 0 \leq \alpha_i \leq C,\ i = 1,\ 2,\ \cdots,\ k \\ \sum_{i=1}^{k}\alpha_i y_i = 0 \end{aligned}\right\} \tag{10}$$

The corresponding decision function is：

$$f(x) = \sin(\sum_{i=1}^{k}\alpha_i y_i K(x_i, x_j) + b) \tag{11}$$

There are several approaches to dispose of the optimization problems mentioned above. We have adopted a SVM toolbox based on the QP algorithm.

3 Headstream Analysis Algorithm Based on SVM

3.1 SVM model for headstream analysis

3.1.1 Training samples

For a distinct comparison, we use the same training data as reference [5], shown in Table 1, which is composed of 10 gushing water samples derived from two aquifers. Here we specify sample 1, 2, 3 as headstream Ⅰ, labelled (−1); 4, 5, 6 as headstream Ⅱ, labeled (+1); 7, 8, 9, 10 as test samples. Each sample was tested for nine kinds of ions and a pH value.

Table 1 Water quality analysis

Sample No.	Water class	Ion content (mmol/L)									pH	SVM decision results	
		Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	Cl^{-}	HCO_3^{-}	SO_4^{2-}	NO_3^{-}	F^{-}			
1	Ⅰ	0.69	0.4	20	0.2	1.31	12.37	7.6	0.04	0.21	8.4	Ⅰ	−1
2	Ⅰ	0.58	0.5	20.87	0.1	1.65	12.18	8.12	0.11	0.21	8.7	Ⅰ	−1.102
3	Ⅰ	0.58	0.46	22.09	0.15	1.45	11.7	7.96	0.1	0.16	8.78	Ⅰ	−1.198
4	Ⅱ	1.94	1.83	4.35	0.26	0.92	5.07	1.94	0.11	0.14	7.9	Ⅱ	1.228
5	Ⅱ	2.21	2.42	3.57	0.1	1.16	5.13	2.44	0.11	0.13	7.5	Ⅱ	1.297
6	Ⅱ	1.44	1.92	6.09	0.15	1.07	5.56	2.55	0.1	0.13	7.56	Ⅱ	1
7	Unknown	0.44	0.57	20	0.1	1.6	11.12	8.01	0.11	0.19	8.8	Ⅰ	−0.955
8	Unknown	0.82	0.6	19.85	0.23	1.18	11.8	7.7	0.02	0.18	8.12	Ⅰ	−0.955
9	Unknown	1.5	1.55	6.87	0.05	1.16	5.75	2.81	0.11	0.13	7.7	Ⅱ	0.896
10	Unknown	1.82	1.79	6.22	0.14	1.1	5.54	2.58	0.12	0.13	7.58	Ⅱ	0.987

3.1.2 Determination of the parameters of SVM

The SVM model used in this paper has two parameters: the constant C and the kernel function K. When $C \geq 1$, the training results of the SVM model are all the same. So we selected $C=1$. The common kernel functions adopted by SVM are as follows[10]: 1) a linear kernel, $K(x,y)=(x \bullet y)$; 2) a polynomial kernel: $K(x,y)=(x \bullet y+1)^d$, d is a natural number; 3) a Gaussian RBF (radial basis function) kernel: $K(x,y)=\exp\dfrac{-(x-y)^2}{2\sigma^2}$, $\sigma>0$; 4) and the sigmoid kernel: $K(x,y)=S(a(x \bullet y)+t)$, where S represents the sigmoid function and a and t are constants. For different kernel functions and parameters, the corresponding SVM model training results are given in Table 2.

Table 2 SVM training results with different kernel functions and parameters

The kernel functions and parameters	Avg CV error	Stddev
Linear kernel	0	0
Polynomial kernel, $d \leq 5$	0	0
Polynomial kernel, $5<d\leq 10$	50.0%	52.70
Gaussian RBF kernel, $0.1\leq б \leq 10$	0	0

As shown in Table 2, the selection of a linear kernel can make a good classification result. If a polynomial kernel is selected, the parameter d must be less than or equal to 5. The Gaussian RBF kernel also does well, but its parameter б is hard to determine. Whether large or small, the generalization property of the SVM model is affected heavily. By common experience, the linear kernel is preferable to other high dimensional kernels, so we selected a linear kernel.

3.1.3 Training result of the SVM model

Given $C=1$ and the linear kernel, the training results of the SVM model are as follows:

the sample no. of support vectors: 1, 6;

the support vector $\boldsymbol{W}^*$=[0.005 6, 0.011 4, −0.103 4, −0.000 4, −0.001 8, −0.050 6, −0.037 5, 0.000 5, −0.000 6, −0.006 2];

bias parameter b^*=2.026 5;

decision function:

$$f(x)=\sin(W^* x+b^*) \tag{12}$$

3.1.4 An extended analysis of decision results

We used Eq. (12) to predict the unknown samples and concluded that all the prediction results given in Table 1 are correct. Table 3 provides a comparison between the support vector ($\boldsymbol{W}^*$) trained by the SVM model and the indices weights (w_i, i=1, 2, ……) trained by the ANN model in reference [5].

Table 3 Comparison between support vector and ANN weights

Decision method	Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	Cl^{-}	HCO_3^{-}	SO_4^{2-}	NO_3^{-}	F^{-}	pH
	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
Support vector	0.005 6	0.011 4	−0.103 4	−0.000 4	−0.001 8	−0.050 6	−0.037 5	0.000 5	−0.000 6	−0.006 2
Ann weights	0.502 8	−0.104 8	0.143 7	1.155 1	0.429 5	0.088 1	−0.039 9	0.949 2	0.993 7	−0.475 8

As shown in Table 3, the ANN method can obtain correct prediction results, but with a frail generalization power. The ANN weights do not denote the importance of the indices correctly. Specifically, the K^+, NO^- and F^- ions with larger weights are not the preferable indices to analyze the water samples. If tested with more samples, the ANN model would arrive at an incorrect decision, because the capacity of the training set is small, which is an unsolved puzzle in ANN theory.

Compared with the ANN model, each component of the support vector trained by the SVM model can denote the importance of each index independently. Sorting the components of the support vector ($\boldsymbol{W}^*$) by their absolute values from maximum to minimum, we know the preferable analytical indices are Na^+ (−0.103 4) and HCO_3^- (−0.050 6) for their larger absolute values, which agrees with practical situations. Moreover, we can test the water samples just by Na^+ and HCO_3^- ions.When we trained the SVM model again in which the features of the input training set include only Na^+ and HCO_3^-, the new decision function obtained was as follows:

$$f(x)=\sin(\boldsymbol{W}\boldsymbol{x}+b),$$

$$\boldsymbol{W}=[-0.116\,0,\ -0.056\,8],\quad b=2.022$$

Using the above function to predict the unknown water samples, we know that all the decision results are correct. This becomes a valuable reference in the selection of water quality indices.

3.2 H-SVMs model for multi-headstream analysis

3.2.1 Training samples

The above SVM model is adopted only to analyze binary classes of water samples. In fact, there are often more than two aquifers in coal mines; usually, a complicated hydraulic interaction exists between them. How to recognize a multi-headstream is a difficulty in the coal mining industry, which has not been satisfactorily solved as yet.

The next problem is how to extend the SVM model and adapt it to multi-headstream recognition. For convenience, we use the training data in reference [6], listed in Table 4. There are 35 headstream samples all together, divided into 4 classes. The second limestone and Ordovician aquifer was labeled Ⅰ; the eighth limestone aquifer labeled Ⅱ; the roof sandstone aquifer labeled Ⅲ and the quaternary aquifer (in which the primary ingredient of gravel stone is limestone) labeled Ⅳ. Six types of ions were selected as the decision indices.

3.2.2 Construction of a SVM classification tree

The recognition of a multi-headstream depends on k-class SVMs. The common k-class SVM methods are given in reference [11] as follows: (1) one-versus-rest SVMs (1-v-r SVMs), (2) one-versus-one SVMs (1-v-1 SVMs), (3) error correcting output codes SVMs (ECOC SVMs), (4) directed acyclic graph SVMs (DAG-SVMs) and (5) hierarchical SVMs (H-SVMs). For the gushing water analysis, the H-SVMs is preferable. We used H-SVMs to construct k-class SVM models with the same parameters as mentioned above in the training process, i.e., $C=1$ and a linear kernel was selected. The training results show that the H-SVMs method is to construct a classification tree based on a serial of bi-class SVMs, in which the tree is a binary tree structure. Each leaf node of the tree represents a class of water samples; the decision rules are the possible routes from tree root to leaves, shown in Fig. 1. For the H-SVMs method, we can construct 15 kinds of classification trees. We tried to find a preferable tree according to a minimum of classification errors and a maximum separating distance. The optimal classification tree is shown in Fig. 1.

Table 4 Samples of multi-headstream

No.	Ion content (mmol/L)						Class
	Na^++K^+	Ca^{2+}	Mg^{2+}	Cl^-	SO_4^{2-}	HCO_3^-	
1	11.98	76.15	15.56	8.5	26.9	292.84	Ⅰ
2	19.34	65.73	18.48	10.64	67.24	239.19	Ⅰ
3	11.5	84.57	24.81	19.86	82.61	253.83	Ⅰ
4	19.78	52.5	16.29	9.93	37.66	229.43	Ⅰ
5	35.1	46.2	17.6	35.8	43.2	212.9	Ⅰ
6	44.88	73.24	24.8	24.07	85.97	303.56	Ⅰ
7	10.29	61.23	29.33	12.16	47.46	309.85	Ⅱ
8	10.64	59.3	28.4	12.59	34.7	291.68	Ⅱ
9	8.0	69.3	26.39	10.96	43.88	295.24	Ⅱ
10	6.45	63.43	24.1	9.24	41.9	266.34	Ⅱ
11	8.3	63.5	26.9	11.19	43.85	282.52	Ⅱ
12	7.1	63	24.7	7.35	37.8	266.13	Ⅱ
13	7.7	67.1	39	8.82	46.5	281.57	Ⅱ
14	7.0	68.7	24.9	11.7	43.77	282.16	Ⅱ
15	17.85	62.96	17.28	6.68	23.31	284.57	Ⅱ
16	13.59	61.59	18.85	6.68	23.57	276.69	Ⅱ
17	10.0	63.87	32.83	4.06	65.09	295.87	Ⅱ
18	12.69	69.39	29.38	13.64	34.54	325.08	Ⅱ
19	98.10	3.1	1.10	23.50	43.84	638.70	Ⅲ
20	207.35	34.75	11.16	23.78	46.54	558.82	Ⅲ
21	311.75	16.25	2.04	33.58	20.56	736.76	Ⅲ
22	303.12	10.24	8.55	32.84	17.47	773.45	Ⅲ
23	304.82	5.77	3.61	40.77	53.00	628.96	Ⅲ
24	257.23	0.00	0.00	27.22	12.24	428.71	Ⅲ
25	502.45	0.00	2.48	29.04	9.79	1105.8	Ⅲ
26	309.33	0.00	0.00	29.03	0.00	562.17	Ⅲ
27	358.58	10.22	3.72	32.68	14.69	691.17	Ⅲ
28	9.10	86.5	31.8	22.40	57.8	348.31	Ⅳ
29	13.25	99.2	31.1	29.85	83.00	361.12	Ⅳ
30	9.20	106.7	39.1	40.10	69.80	402.10	Ⅳ
31	17.30	98.2	20.6	20.24	53.20	354.40	Ⅳ
32	4.68	69.14	22.93	26.67	13.38	251.26	Ⅳ
33	19.58	74.67	16.92	24.46	27.62	272.94	Ⅳ
34	19.90	70.47	16.78	18.40	10.79	294.47	Ⅳ
35	20.54	51.73	16.04	24.34	12.34	236.00	Ⅳ

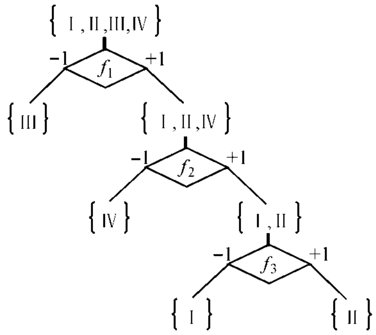

Fig. 1 SVM classification tree for the recognition of multi-headstream

3.2.3 Multi-headstream decision rules and an extended analysis

The decision rules obtained by the SVM classification tree are given in Table 5. The combination of the states calculated by the decision functions $f_i(x)$ denotes the predicted classes of the test samples, $f_i(x) = \sin(w_i \boldsymbol{x} + b_i),\ i = 1,\ 2,\ 3$. The parameters of $f_i(x)$ are given in Table 6. Using the decision rules in Table 5 to predict the test samples in ref. [6], we obtained the results shown in Table 7.

Table 5 Decision rules to recognize multi-headstreams

Class	f_1	f_2	f_3
Ⅰ	+1	+1	−1
Ⅱ	+1	+1	+1
Ⅲ	−1		
Ⅳ	+1	−1	

Table 6 Values of the parameters (w_i, b_i) of the decision function f_i

Decision function	$Na^+ + K^+$	Ca^{2+}	Mg^{2+}	Cl^-	SO_4^{2-}	HCO_3^-	b_i
	w_{i1}	w_{i2}	w_{i3}	w_{i4}	w_{i5}	w_{i6}	
f_1	−0.005 5	0.003 5	0.001 3	0.000 5	0.001 4	−0.004 5	2.320 8
f_2	0.101 0	−0.096 2	−0.016 9	−0.239 7	0.090 9	0.001 5	6.531 4
f_3	−0.050 0	−0.192 0	0.128 7	−0.037 5	−0.003 7	0.065 4	−6.532 3

Table 7 Prediction and verification of the headstream of gushing water using the SVM classification tree

Sample	$Na^+ + K^+$	Ca^{2+}	Mg^{2+}	Cl^-	SO_4^{2-}	HCO_3^-	f_1	f_2	f_3	Primary class	Predicted class	Result
1	23.76	66.4	19.59	18.13	57.26	255.29	+1	+1	−1	Ⅰ	Ⅰ	True
2	9.97	64.45	26.84	9.59	40.53	288.14	+1	+1	+1	Ⅱ	Ⅱ	True
3	294.75	8.93	3.63	30.27	24.24	680.51	−1			Ⅲ	Ⅲ	True
4	14.19	81.96	24.41	25.81	40.99	315.08	+1	−1		Ⅳ	Ⅳ	True

Compared with the decision rules in ref. [6], the SVM classification tree is more intuitive, with a simple structure, convenient for calculation and it is easy to extract decision rules. As well, the components (w_i) of the support vector ($\boldsymbol{W}$) can denote the importance of each ion index to the headstream recognition. Since the absolute values of component w_{11} (−0.005 5) and w_{16} (−0.004 5) are larger than the others of the support vector w_1 in the decision function f_1, as shown in Table 6, we can say that the related ions $Na^+ + K^+$ and HCO_3^- are the preferred indices in water sample Ⅲ, which agrees with the fact that water sample Ⅲ is derived from the roof sandstone aquifer. Water samples Ⅰ, Ⅱ and Ⅳ are derived from the limestone aquifer and are different from water sample Ⅲ, which agrees with the practice of first separating Ⅲ from Ⅰ, Ⅱ and Ⅳ. The results also prove that the choice of the classification tree is right. Next, we use f_2 to separate water sample Ⅳ from Ⅰ and Ⅱ. This occurs prior to the separation of the samples between the quaternary aquifer and the limestone aquifer (the eighth limestone aquifer and the second limestone aquifer); the representative ions are $Na^+ + K^+$ and Cl^-. Finally, we use f_3 to distinguish water samples Ⅰ and Ⅱ where the representative ions are Ca^{2+} and Mg^{2+}. This agrees with the experience that water sample Ⅰ has a lower hardness and Ⅱ a higher hardness.

It is evident that using the H-SVMs model to analyze the headstreams of gushing water can not only obtain correct results but can also make a sophisticated analysis of the indices, which is a valuable supplement to other methods.

4 Conclusions

1) The SVM is a valid method to analyze the headstream of gushing water in coal mines and the H-SVMs a preferable method to recognize multi-headstreams. The training parameter C of the SVM model is greater or equal to 1 and the preferable kernel function is linear.

2) The headstream of gushing waters in coal mines shows a good linear separability. The feasibility to distinguish the headstream depends on the quantity and separability of hydrogeochemical indices. Different aquifers have distinct ion indices, whose weights can be denoted by the support vector $\boldsymbol{W}$ in decision function of the SVM model, which is a valuable reference for selecting hydrogeochemistry

indices.

References

[1] Zhou X L, Zheng S S, Yang G Y. Synthetical analysis on the sources of water inrush into no.2 mine of Nantong mining area. *Journal of China University of Mining & Technology*, 1997, 2(26): 60–62. (In Chinese)

[2] Yue M. Related analysis of gray system applied to determinate water resources of mine water inrush. *Coal Science and Technology,* 2000, 4(30): 37–39. (In Chinese)

[3] Li M S, Yu Y L, Lu F G, et al. A fuzzy comprehensive evaluation model of mine shaft water resource outburst in Yaoqao coal mine. *Site Investigation Science and Technology*, 2001(2): 16–20. (In Chinese)

[4] Wang Y M, Jiao L M. Judging mine inflow source by analyzing water quality. *Safety in Coal Mines*, 2001, 10(32): 12–14. (In Chinese)

[5] Feng L J, Li J S, Shao G Q. The application of adaline in recognition of mine water quality types. *Coal Geology & Exploration,* 2002, 4(30): 35–37. (In Chinese)

[6] Zhang X L, Zhang Z X, Peng S P. Application of the second theory of quantification in identifying gushing water sources of coal mines. *Journal of China University of Mining & Technology,* 2003, 3(32): 251–254. (In Chinese)

[7] Zhang H R, Zhou R F, Guo D Z, et al. Investigation on predicting roof water gush in coal mines based on multi-factor analysis. *Journal of China University of Mining & Technology*, 2005, 1(34): 112–116. (In Chinese)

[8] Lu F, Liu X, Liu Q. The theory of gray relative analysis and it's new research. *Journal of Wuhan University of Technology*, 2000, 2(22): 41–43. (In Chinese)

[9] Vapnik V N. *The Nature of Statistical Learning Theory.* New York: Springer-Verlag, 2000.

[10] Deng N Y, Tian Y J. *A Novel Method of Data Mining-Support Vector Machine.* Beijing: Science Press, 2004. (In Chinese)

[11] Liu Z G, Li D R, Qin Q Q, et al. An analytical overview of methods for multi-category support vector machines*. Computer Engineering and Applications*, 2004, 7(40): 10–13. (In Chinese)