

矿井涌水水源识别的 MMH 支持向量机模型

闫志刚^{1, 2}, 白海波²

(1. 中国矿业大学 环测学院, 江苏 徐州 221116; 2. 中国矿业大学 深部岩土力学与地下工程国家重点实验室, 江苏 徐州 221008)

摘要: 提出一种新的多水源判别的 H 支持向量机模型。推导 H 支持向量机的理论推广误差公式, 发现确保高优先级节点的推广性能是提高 H 支持向量机性能的有效途径; 设计基于 SVM 最大间隔逐层分类、最小间隔逐层聚类构造 H 支持向量机的新方法, 以各支持向量机节点的分类间隔为分类、聚类指标, 通过 TopDown, BottomUp 两种方式混合构造 H 支持向量机, 即 MMH 支持向量机。实验效果表明, MMH 支持向量机结构简单、泛化能力强, 不仅能正确区分各类水源, 而且其层次结构能很好地反映各水源的层次关系。判别函数的法向量还可以指示各含水层水质化验指标的权重, 为矿井涌水水源识别提供了新的科学方法。

关键词: 采矿工程; 矿井涌水; 多水源识别

中图分类号: TD 45

文献标识码: A

文章编号: 1000 - 6915(2009)02 - 0324 - 06

MMH SUPPORT VECTOR MACHINES MODEL FOR RECOGNIZING MULTI-HEADSTREAM OF WATER INRUSH IN MINE

YAN Zhigang^{1, 2}, BAI Haibo²

(1. School of Environment and Spatial Informatics, Chinese University of Mining and Technology, Xuzhou, Jiangsu 221116, China;
2. State Key Laboratory for Geomechanics and Deep Underground Engineering, Chinese University of Mining and Technology, Xuzhou, Jiangsu 221008, China)

Abstract: A novel hierarchy support vector machines(H-SVMs) model is presented to recognize the headstreams of water inrush in coal mine. Firstly, an analytical model is deduced to analyze the generalization power of H-SVMs. According to the results, a feasible approach is put forward to improve the performance of H-SVMs to guarantee the performances of each SVM node, whose position is located at a high level. Secondly, a novel method is presented to build H-SVMs, i.e. MMH-SVMs(maximal margin hierarchical SVMs), taking the separating margins of each SVM node as indices for classification and clustering, using TopDown and BottomUp routes from top to bottom to classify the input samples at each SVM node by maximal separating margin and from bottom to top clustering the input samples by minimal separating margin. Experimental results show MMH-SVMs have a simple structure, and a good generalization performance. It can predict the headstreams of water inrush correctly; and its tree structure can also denote the hierarchy of headstreams. Moreover, the normal vector parameter W in each SVM decision function can describe the weights of discrimination indices of the headstreams of water inrush, in which a novel scientific method is introduced to predict the headstream of water inrush in coal mine.

Key words: mining engineering; water inrush in coal mine; multi-headstream recognition

收稿日期: 2008 - 10 - 06; **修回日期:** 2008 - 11 - 20

基金项目: 国家重点基础研究发展计划(973)项目(2007CB209406); 国家自然科学基金重点项目(50634050); 国家自然科学基金资助项目(40802061); 中国博士后科学基金(20080441081)

作者简介: 闫志刚(1974 -), 男, 博士, 1997年毕业于中国矿业大学地质工程专业, 现任副教授, 主要从事空间数据处理、矿井突水预测等方面的教学与研究。E-mail: zhg-yan@163.com

1 引言

水源识别是突(涌)水水量预测的前提, 是矿井水害防治工作的基础, 水文地球化学方法是判别突(涌)水水源的有效方法^[1~4], 利用该方法判别水源时, 首先, 根据各含水层的水文长观孔、突水监测点采集若干水源样本, 测定各样本的水化学指标, 然后, 根据确知水化学组分的样本建立水源判别的目标函数, 再利用目标函数对待定水源进行识别。常用的判别方法可归结为多元统计、灰色关联分析、神经网络以及支持向量机(SVM)等。张许良等^[1]利用多元统计分析中的数量化理论建立了水源判别的线性模型; 杨永国和黄福臣^[2]研究了水源判别的非线性方法: 灰色关联分析与 BP 神经网络; 上述方法均基于传统的统计学习理论, 要求有大的样本数目支撑, 但这个前提条件通常难以满足。SVM 是基于结构风险最小化原理的新兴的统计学习方法, 更适合于小样本的识别问题, 具有较优的推广能力。姜谔男和梁冰^[3]研究了水源分析的 LS-SVM 模型, 取得了很好的预测效果, 但模型通过径向基函数将水源识别因子映射到高维空间, 无法对各因子进行有效分析, 并且, 在处理多类水源时, 采用了编码运算, 隐藏了水源识别的细节。闫志刚等^[4]认为水源间有良好的线性可分性, 提出了水源识别的线性 SVM 模型。本文进一步提出了多水源分析的 MMH 支持向量机(maximal margin H-SVMs)模型, 推导了模型的理论推广误差公式, 研究了新的建模方法, 并用于实际的水源识别。

2 支持向量机简介

对于模式识别问题^[5], 就是找到一个可计算的识别函数 $y = f(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^n$, $y \in \{-1, 1\}$, 对于给定的 k 个样本 $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$, $\mathbf{x} \in \mathbf{R}^n$, 来找到一个可将样本分离的超平面(决策平面), 即 $\mathbf{W}\mathbf{x} + \mathbf{b} = 0$, $\mathbf{W} \in \mathbf{R}^n$, $\mathbf{b} \in \mathbf{R}$, 对应的识别函数为

$$f(\mathbf{x}) = \text{sgn}[(\mathbf{W}\mathbf{x}) + \mathbf{b}] \quad (1)$$

决策平面应满足约束为

$$y_i(\mathbf{W}\mathbf{x}_i + \mathbf{b}) \geq 1 - \xi_i \quad (i = 1, 2, \dots, k) \quad (2)$$

最佳决策面应满足 2 类样本到决策面的最小距

离为最大, 这样, 分类问题变成满足 $\xi_i \geq 0$ 和式(2)的最小化问题, 即

$$\min(\tau(\mathbf{W})) = \frac{1}{2} \|\mathbf{W}\|^2 + C \sum_{i=1}^k \xi_i \quad (3)$$

式(3)中的第一项使 2 类样本到决策面的最小距离为最大, 第二项使误差为最小; 常数 C 对二者折中。这个带有约束的优化问题, 可以利用拉格朗日优化方法转化为其对偶问题:

$$\left. \begin{aligned} \max(\mathbf{W}(\alpha)) &= -\frac{1}{2} \sum_{i,j=1}^k \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) + \sum_{i=1}^k \alpha_i \\ \sum_{i=1}^k \alpha_i y_i &= 0 \\ (0 \leq \alpha_i \leq C, \quad i &= 1, 2, \dots, k) \end{aligned} \right\} \quad (4)$$

相应的分类函数可化为

$$f(\mathbf{x}) = \text{sgn} \left[\sum_{i=1}^k \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{x}) + \mathbf{b} \right] \quad (5)$$

对于非线性可分的情况, 可使用一个非线性函数 ϕ , 把数据映像到一个高维特征空间, 再在高维特征空间建立优化超平面, 相应的超平面方程为 $\mathbf{W}\phi(\mathbf{x}) + \mathbf{b} = 0$, 分类函数变为

$$f(\mathbf{x}) = \text{sgn} \left[\sum_{i=1}^k \alpha_i y_i (\phi(\mathbf{x}) \cdot \phi(\mathbf{x}_i)) + \mathbf{b} \right] \quad (6)$$

通常, 函数 ϕ 难以确定, SVM 理论只考虑高维特征空间的点积运算 $K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}) \cdot \phi(\mathbf{y})$, 而不直接使用函数 ϕ , $K(\mathbf{x}, \mathbf{y})$ 称为核函数, 分类函数变为

$$f(\mathbf{x}) = \text{sgn} \left[\sum_{i=1}^k \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + \mathbf{b} \right] \quad (7)$$

3 H 支持向量机

利用 SVM 仅能识别 2 类水源, 识别多类水源需要多类 SVMs 算法, 层次 SVMs(hierarchical SVMs, H-SVMs)就是其中一种, 见图 1。H-SVMs 首先将所有类别分成 2 个子类, 再将子类进一步划分成 2 个次级子类, 直到得到若干单独的类别为止, 这样就得到一棵倒立的二叉分类树, 样本类别的判别规则就是由树根到叶子节点经过的路径。

3.1 现有 H-SVMs 的构造方法

目前, 构造 H-SVMs 均采用 TopDown 策略(从根节点到叶子节点), 逐层将输入类别分为 2 类, 就

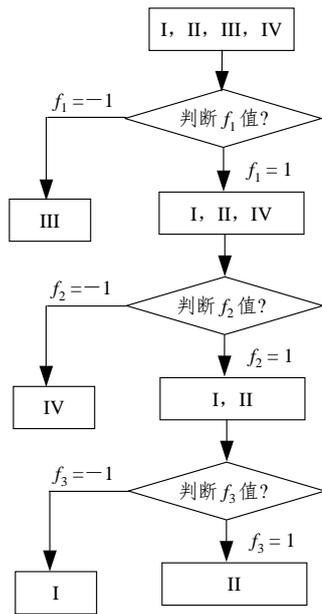


图 1 多水源判别的 MMH-SVMs 分类树

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

各决策节点如何分配输入类别有如下几种方式:

(1) 根据领域知识分配

该方法^[6]一般由熟悉领域问题的专家胜任, 根据问题领域的层次关系设计 H-SVMs 树, 树形结构合理, 符合实际; 但该方法在无专家知识的情况下将不再适用。

(2) 根据样本的空间分布分配

该方法^[7~9]先将输入样本聚为 2 类, 然后按一定的策略在决策节点分配输入类别, 一般采取“winner-take-all”(胜者通吃)的规则, 也就是输入类别的归属由聚类结果中包含该类样本最多的聚类中心决定。该方法减轻了对专家的依赖, 在领域知识不明显的情况下, 可以构造出合理的 H-SVMs, 但该方法的实际效果与聚类算法有关, 不同算法的偏差较大, 并且会出现“无胜出”或“全胜出”而导致某侧子树为空的现象。

(3) 逐层优选

该方法^[10]实质也是聚类分析, 只是在每次聚类时, 仅优选出某类作为叶子节点, 将其余各类输入到下一层继续优选, 而优选的标准可以是类凸包间的欧氏距离, 分类精度等。该方法构造的是一偏态树, 优点是训练速度、分类速度快, 缺点是偏态树

中节点层次多, 误差累积效应大, 并且, 每次优选一类的策略并不一定合理。

(4) 形态优先

该方法^[11]优先设计树形, 再在各节点分配样本类别, 尽量使 H-SVMs 树接近正态。该方法可以减少节点层次, 加快分类速度, 在类别数很大时, 比较实用, 但该方法仅仅考虑树形, 可能与实际有出入。

通过上述分析可知, 虽然构造 H-SVMs 的方法很多, 但如何针对矿井水源识别问题构造合理的 H-SVMs 仍需要专门研究。

3.2 H-SVMs的推广性能分析

根据定理^[12], 假设感知器决策树(perceptron DT, PDT)可以将 m 个样本分开, 其中, 决策节点个数为 K , 设各决策节点 i 的分类间隔为 γ_i , 则该决策树推广误差以概率 $1 - \delta$ 为

$$\varepsilon(m, K, \delta) \leq$$

$$\frac{2}{m} \left[65R^2 D' \lg(4em) \lg(4m) + \lg \frac{(4m)^{K+1} \binom{2K}{K}}{(K+1)\delta} \right] \quad (8)$$

其中,

$$D' = \sum_{i=1}^K \frac{1}{\gamma_i^2}$$

式中: R 为样本的最小包球半径, $\frac{1}{K+1} \binom{2K}{K}$ 为 PDT 的结构计数。

H-SVMs 是一种特殊的决策树, 对于类别数为 K 的 H-SVMs, 其决策结点个数为 $K - 1$, 结构计数为 $\prod_{i=1}^{K-1} (2i - 1)$, 重复式(8)的证明过程, 可得 K 类 m 个样本被 H-SVMs 线性可分时的理论推广误差上界

$$\varepsilon(m, K, \delta) \leq$$

$$\frac{2}{m} \left(65R^2 D' \lg(4em) \lg(4m) + \lg \frac{(4m)^K \prod_{i=1}^{K-1} (2i - 1)}{\delta} \right) \quad (9)$$

其中,

$$D' = \sum_{i=1}^{K-1} \frac{1}{\gamma_i^2}$$

由式(9)可知, H-SVMs 推广误差的界与 K , m , R , D' 有关。当样本类别数 K 、样本数 m 、样本最小包球半径 R 增大时, H-SVMs 的推广误差界也增大, 并且, 推广误差界受 K , m 的影响大。因此, 减小样本类别数、样本数是提高 H-SVMs 性能的根本途径。

对于给定的样本集, K , m , R 不再变化, 则 H-SVMs 的推广能力只与 D' 有关: D' 越小, 则 H-SVMs 的推广误差界越小。对给定样本集的 H-SVMs 分类树来说, 由于 D' 是各决策节点分类间隔平方的倒数和, 因此, 要使 D' 减小, 则每一决策节点的分类间隔 γ_i 都要取最大, 而 SVM 本身就是使 γ_i 最大的分类器, 所以只要树的形态确定了, 则 D' 的大小也随之确定。这就说明 H-SVMs 的推广能力不仅依赖于每个 SVM 节点的推广能力, 更依赖于它的体系结构。

现在的问题是如何确定最佳的树形来获得最小的 D' 。在 H-SVMs 中, 上层决策节点将位于该节点的样本分为两类, 不同的分类结果必然影响到下层节点的分类间隔, 可见, γ_i 之间存在约束关系, 则求 D' 的最小值问题就是在约束条件下的极值问题。但 γ_i 间的约束是随机的, 对 D' 的影响不确定, 也就是说将无法确定牺牲本层节点的分类间隔 γ_i 会换来整体最优, 即 $D' = \sum_{i=1}^K \frac{1}{\gamma_i^2}$ 的最小。因此, 稳妥起

见, 在每一决策节点分配输入样本类别时, 应优先考虑使本层 SVM 节点的分类间隔 γ 最大。

至此, 可以认为: 对于特定的分类问题, 应该存在推广能力最优的 H-SVMs, 但能否构造出性能较优的 H-SVMs 是存在风险的。当采用 TopDown 策略构造 H-SVMs 时, 越是上层的节点, 优先级越高, 确保优先级高的节点的推广性能较优, 才能增大构造较优 H-SVMs 的可能性。

4 矿井水源识别的 MMH-SVMs 模型

4.1 建模方法

根据上面的分析结果, 本文将按照从上到下, 优先保证高层节点的推广性能的原则构造 H-SVMs。当 SVM 线性可分时, 它的推广能力可由分类间隔 γ 来衡量, γ 越大则 SVM 的推广能力越好, 本文基此设计了新的适合于矿井水源分析的 H-SVMs 建模方法:

4.1.1 最大间隔逐层分类法

(1) 步骤 1。选择使 SVM 分类间隔最大的分类方法将输入样本按类别分为 2 类, 若输入样本的类别数为 K , 则分类方法共有 $2^{(K-1)} - 1$ 种。

(2) 步骤 2。对步骤 1 得到的 2 样本集分别递归执行, 当样本类别数为 1 时, 则中止。

由于最大间隔逐层分类的思想来源于式(9)的结论, 因此可以得到推广性能较优的 H-SVMs, 但该方法在样本类别数较大时, 运算将很困难, 因为在最坏情况下, 该方法训练的 SVM 分类器个数与类别 K 的关系为 $O(2^{(K-1)})$ 。为了解决该问题, 可以采用下面的简化方法。

4.1.2 最小间隔逐层聚类法

(1) 步骤 1。利用 SVM 对输入样本的任意 2 类进行分类, 将分类间隔最小的 2 类聚为一类, 若输入样本类别为 K , 则所需训练的 SVM 分类器个数为 $K(K-1)/2$ 。

(2) 步骤 2。对聚类后的样本递归执行, 直到所有类别聚为一类, 即 H-SVMs 的根。

该方法构造 H-SVMs 分类树所训练的 SVM 分类器个数与类别 K 的关系为 $O((K-1)^2)$ 。

最大间隔逐层分类采用了 TopDown 策略, 最小间隔逐层聚类则是 BottomUp 策略, 容易证明两类方法所构 H-SVMs 并不一致。最大间隔分类的推广性能较优, 最小间隔聚类的时间效率较高, 2 类方法可以混用, 通过聚类尽快约简样本类别, 以减少训练复杂度; 通过分类可以获得推广性能好的决策节点, 聚类与分类阶段则是针对样本类别数的折中。本文称上述建模方法构造的 H-SVMs 为最大间隔 H-SVMs。

4.2 实例分析

为方便比较, 使用张许良等^[1]的数据, 见表 1, 共采集到 39 个水源样本, 水样来源于该矿的 4 个主要含水层, 前 35 个水样做为训练样本, 来构造判别模型, 后 4 个做为模型的测试样本。I 为二灰和奥陶纪含水层, II 为八灰含水层, III 为顶板砂岩含水层, IV 为第四系含水层(砂砾石成分以石灰岩为主), 选取 6 种离子组合作为判别指标。

H-SVMs 各节点均采用线性 SVM, 线性 SVM 的训练参数 C 是对分类间隔与分类错误率的折中, C 过大或者过小均影响 SVM 的推广性能, 通过多次实验最后确定 $C = 1$ 。4 类水源对应的所有树结构有 15 种, 用本文提出的最大间隔逐层分类、最小间

隔逐层聚类方法分别求对应的 H-SVMs 结构，两种方法所得结果一致，如图 1 所示。

由图 1 得到的多水源判别规则见表 2，判别函数 f_1, f_2, f_3 的状态组合确定了水样的类别。判别函数参数见表 3。利用表 2 的判别规则对表 1 中的检验水样进行判别，结果全部正确。

表 1 多水源样本^[1]
Table 1 Samples of multi-headstream^[1]

序号	离子含量/(mmol · L ⁻¹)						水源类别
	Na ⁺ +K ⁺	Ca ²⁺	Mg ²⁺	Cl ⁻	SO ₄ ²⁻	HCO ₃ ⁻	
1	11.98	76.15	15.56	8.50	26.90	292.84	I
2	19.34	65.73	18.48	10.64	67.24	239.19	I
3	11.50	84.57	24.81	19.86	82.61	253.83	I
4	19.78	52.50	16.29	9.93	37.66	229.43	I
5	35.10	46.20	17.60	35.80	43.20	212.90	I
6	44.88	73.24	24.80	24.07	85.97	303.56	I
7	10.29	61.23	29.33	12.16	47.46	309.85	II
8	10.64	59.30	28.40	12.59	34.70	291.68	II
9	8.00	69.30	26.39	10.96	43.88	295.24	II
10	6.45	63.43	24.10	9.24	41.90	266.34	II
11	8.30	63.50	26.90	11.19	43.85	282.52	II
12	7.10	63.00	24.70	7.35	37.80	266.13	II
13	7.70	67.10	39.00	8.82	46.50	281.57	II
14	7.00	68.70	24.90	11.70	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	10.00	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	3.10	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	86.50	31.80	22.40	57.80	348.31	IV
29	13.25	99.20	31.10	29.85	83.00	361.12	IV
30	9.20	106.70	39.10	40.10	69.80	402.10	IV
31	17.30	98.20	20.60	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
36	23.76	66.40	19.59	18.13	57.26	255.29	I
37	9.97	64.45	26.84	9.59	40.53	288.14	II
38	294.75	8.93	3.63	30.27	24.24	680.51	II
39	14.19	81.96	24.41	25.81	40.99	315.08	IV

表 2 水源判别规则
Table 2 Discrimination rules to recognize multi-headstream

类别	f_1	f_2	f_3
I	+1	+1	-1
II	+1	+1	+1
III	-1		
IV	+1	-1	

表 3 判别函数 f_i 的参数值(w_i, b_i)
Table 3 Values of the parameters(w_i, b_i) of the discrimination function f_i

判别函数变量	f_1	f_2	f_3
$W_{i1}(\text{Na}^+\text{+K}^+)$	-0.005 5	0.101 0	-0.050 0
$W_{i2}(\text{Ca}^{2+})$	0.003 5	-0.096 2	-0.192 0
$W_{i3}(\text{Mg}^{2+})$	0.001 3	-0.016 9	0.128 7
$W_{i4}(\text{Cl}^-)$	0.000 5	-0.239 7	-0.037 5
$W_{i5}(\text{SO}_4^{2-})$	0.001 4	0.090 9	-0.003 7
$W_{i6}(\text{HCO}_3^-)$	-0.004 5	0.001 5	0.065 4
b_i	2.320 8	6.531 4	-6.532 3

与张许良等^[1~3]的水源判别模型比，MMH-SVMs 分类树更直观，结构简单，计算方便，并且各支持向量 W_i 的分量还可以表征相应的水化学指标在水源判别中的权重，见表 2。判别函数 f_1 的支持向量 W_1 的分量 $W_{11}(-0.005 5)$ ， $W_{16}(-0.004 5)$ 较其他分量，说明对应的 $\text{Na}^+\text{+K}^+$ ， HCO_3^- 是水源 III 指示离子，与 III 顶板砂岩水，类型为 HCO_3^- 和 $\text{K}^+\text{+Na}^+$ 相吻合；水源 I，II，IV 为灰岩类水，III 与 I，II，IV 不同类型水的优先分开符合实际。接下来，利用 f_2 将水源 IV(第四系水)与水源 I(二灰水)，II(八灰水)分开，也就是将第四系含水层与灰岩含水层优先分开，指示离子为 $\text{Na}^+\text{+K}^+$ 和 Cl^- ；最后用 f_3 区分 I(二灰)与 II(八灰)水源，指示离子为 Ca^{2+} ， Mg^{2+} ，也符合 I 为低硬度水，II 为高硬度水的实际

情况。

可见, MMH-SVMs 不仅能将多水源正确分类, 还能对分类指标进行分析, 弥补了其他方法的不足。

5 结论与探讨

(1) MMH-SVMs 非常适合于矿井水源识别, 不仅预测精度较高, 而且其层次结构能很好地反映水源类型的层次关系, 各子 SVM 的法向量还能反映水源识别因子的权重。

(2) 本例的水源识别数据具有良好的线性可分性, 可以用分类间隔来描述各子 SVM 的推广能力, 这也是 MMH-SVMs 建模的基础。但对非线性可分 SVM, 目前仍缺乏描述其推广能力的确切理论, 也影响了 MMH-SVMs 的实用性, 这需要进一步研究。

参考文献(References):

- [1] 张许良, 张子戎, 彭苏萍. 数量化理论在矿井突(涌)水水源判别中的应用[J]. 中国矿业大学学报(自然科学版), 2003, 32(3): 251 - 254.(ZHANG Xuliang, ZHANG Ziwu, PENG Suping. Application of second theory of quantification to identifying water inrush source of coal mines[J]. Journal of China University of Mining and Technology, 2003, 32(3): 251 - 254.(in Chinese))
- [2] 杨永国, 黄福臣. 非线性方法在矿井突水水源判别中的应用研究[J]. 中国矿业大学学报(自然科学版), 2007, 36(3): 283 - 286.(YANG Yongguo, HUANG Fuchen. Water source determination of mine inflow based on nonlinear method[J]. Journal of China University of Mining and Technology(Natural Science), 2007, 36(3): 283 - 286.(in Chinese))
- [3] 姜谱男, 梁冰. 地下水化学特征组分识别的粒子群支持向量机方法[J]. 煤炭学报, 2006, 31(3): 310 - 313.(JIANG Annan, LIANG Bing. Particle swarm optimization support vectors machine method of identifying standard components of ions of groundwater[J]. Journal of China Coal Society, 2006, 31(3): 310 - 313.(in Chinese))
- [4] 闫志刚, 杜培军, 郭达志. 矿井涌水水源分析的支持向量机模型[J]. 煤炭学报, 2007, 32(8): 842 - 847.(YAN Zhigang, DU Peijun, GUO Dazhi. SVM models for analyzing the headstreams of mine water inrush[J]. Journal of China Coal Society, 2007, 32(8): 842 - 847.(in Chinese))
- [5] VAPNIK V, LADIMIR N. The nature of statistical learning theory[M]. New York: Springer Verlag, Inc., 2000.
- [6] 徐启华, 师军. 一种新型多分类支持向量算法及其在故障诊断中的应用[J]. 系统仿真学报, 2005, 17(11): 2 766 - 2 768.(XU Qihua, SHI Jun. New multi-class support vector algorithm and its application to fault diagnosis[J]. Acta Simulata Systematica Sinica, 2005, 17(11): 2 766 - 2 768.(in Chinese))
- [7] CASASENT D, WANG Y C. A hierarchical classifier using new support vector machines for automatic target recognition[J]. Neural Networks, 2005, 18(5 - 6): 541 - 548.
- [8] HAO P Y, CHIANG J H, TU Y K. Hierarchically SVM classification based on support vector clustering method and its application to document categorization[J]. Expert Systems with Applications, 2007, 33(3): 627 - 635.
- [9] 张国云, 章兢. 一种新的分裂层次聚类 SVM 多值分类器[J]. 控制与决策, 2005, 20(8): 931 - 934.(ZHANG Guoyun, ZHANG Jing. A novel SVM multi-class classifier based on divisive hierarchical clustering[J]. Control and Decision, 2005, 20(8): 931 - 934.(in Chinese))
- [10] 唐发明, 王仲东, 陈绵云. 支持向量机多类分类算法研究[J]. 控制与决策, 2005, 20(7): 746 - 749.(TANG Faming, WANG Zhongdong, CHEN Mianyun. On multi-class classification methods for support vector machines[J]. Control and Decision, 2005, 20(7): 746 - 749.(in Chinese))
- [11] LEI H S, GOVINDARAJU V. Half-against-half multi-class support vector machines[C]// Proceedings of IAPR International Workshop on Multiple Classifier Systems. Monterrey, CA: [s. n.], 2005: 156 - 164.
- [12] BENNETT K, CRISTIANINI N, SHAWE-TAYLOR J, et al. Enlarging the margin in perceptron decision trees[J]. Machine Learning, 2000, 41(3): 295 - 313.