

文章编号: 1001-0920(2008)07-0745-06

决策树支持向量机多分类器设计的向量投影法

厉小润¹, 赵光宙¹, 赵辽英²

(1. 浙江大学 电气工程学院, 杭州 310027; 2. 杭州电子科技大学 计算机应用技术研究所, 杭州 310018)

摘要: 针对如何有效地设计决策树支持向量机(SVM)多类分类器的层次结构这个关键问题,提出一种基于向量投影的类间可分性测度的设计方法,并给出一种基于该类间可分性测度设计决策树 SVM 多分类器层次结构的方法。为加快每个 SVM 子分类器的训练速度且保持其高推广性,将基于向量投影的支持向量预选取方法用于每个子分类器的训练中。通过对 3 个大规模数据集和手写体数字识别的仿真实验表明,新方法能有效地提高决策树 SVM 多类分类器的分类精度和速度。

关键词: 多分类器; 决策树支持向量机; 向量投影; 类间可分性

中图分类号: TP391

文献标识码: A

Design of decision-tree-based support vector machines multi-class classifier based on vector projection

LI Xiaorun¹, ZHAO Guangzhou¹, ZHAO Liaoying²

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China; 2. Institute of Computer Application Technology, Hangzhou Dianzi University, Hangzhou 310018, China. Correspondent: ZHAO Liaoying, E-mail: zhaoly@zju.edu.cn)

Abstract: For the key issue that how to effectively design the hierarchical structure of the decision-tree-based support vector machines multi-class classifier, a method based on vector projection is proposed to measure inter-class separability, and a way is presented to design the hierarchical structure of the multi-class classifier based on the inter-class separability. Furthermore, in order to improve the training speed of each sub-classifier and maintain its high generalization ability, the method for pre-extracting support vectors based on vector projection is applied in each sub-classifier. Experiment results on three large-scale data sets and handwriting digits recognition indicate that the proposed method speeds up the decision-tree-based support vector machines multi-class classifiers and yields higher precision.

Key words: Multi-class classifiers; Decision-tree-based support vector machines; Vector projection; Inter-class separability

1 引言

支持向量机(SVM)是 Vapnik 等人^[1]提出的一种基于统计学习理论的机器学习方法,它较好地解决了非线性、高维数、局部极小点等问题,并且对于特征相关性和稀疏性不敏感,已成为继神经网络之后机器学习的一个新的研究热点。目前,SVM 已广泛地应用于人脸识别、人脸检测、指纹识别和时间序列预测等方面,并取得了良好的应用效果。

然而,SVM 是针对两类分类问题提出的,如何将其有效地推广到多类分类是当前的研究热点。目

前,SVM 解决多类识别问题的方法可分为两大类:整体法和分解法。整体法是在所有训练样本上一次性求解一个大的二次规划问题,同时将多类分开^[2]。这种方法思路虽然简单,但求解最优化问题的过程中使用的变量非常多,计算复杂度过高而不适用。分解法主要针对多个两类分类问题进行多类分类,包括 1-a-r (one-against-rest)^[3]方法,1-a-1 (one-against-one)^[4]方法,决策导向非循环图方法(DDAGSVM)^[5]和基于决策树的 SVM 多类分类方法(DTB SVM)^[6-9]。对于 N 类分类问题,

收稿日期: 2007-03-27; 修回日期: 2007-09-07.

基金项目: 浙江省自然科学基金项目(Y106085)。

作者简介: 厉小润(1970—),男,浙江东阳人,副教授,博士生,从事模式识别、智能控制的研究;赵光宙(1946—),男,浙江东阳人,教授,博士生导师,从事非线性控制、预测控制等研究。

DTBSVM 分类器只需构造 $N - 1$ 个 SVM 分类器, 并且该方法不存在不可识别域, 分类时也不需要遍历所有的分类器. 大量研究和实验表明^[6-9], 基于决策树的 SVM 多类分类方法是当前用 SVM 解决多类识别问题方法中最优的一种. 但如何有效地设计决策树 SVM 多类分类器的层次结构仍是一个有待解决的问题.

类的可分性测度的设计是有效设计 SVM 多类分类器层次结构的基础. 本文提出了类间可分性测度设计的向量投影法, 并以此为依据设计决策树 SVM 多类分类器的层次结构. 理论分析和仿真实验表明, 文中设计的 SVM 多分类器能有效提高分类精度和速度.

2 基于决策树的 SVM 多分类器的结构

DTBSVM 分类器的结构有两种: 一种是每个二值 SVM 分类器由一个类与剩下的类构造分割面; 另一种是每个二值 SVM 分类器由多类与多类构造分割面. 本文主要讨论第 1 种 DTBSVM 分类器.

决策树 SVM 多类分类器的测试过程是: 从根节点开始计算决策函数, 根据值的正负决定下一节点, 如此下去, 直到到达某一叶子节点为止, 此叶子节点所代表的类别就是测试样本的所属类别. 根据该测试过程可知, 测试时并不一定要计算所有分类器判别函数, 并且越上层节点的分类性能对整个分类器模型的推广性影响越大. 因此, 在生成决策树的过程中, 应使最易分割的类最早分割出来, 即在决策树的上层节点处开始分割.

在 DTBSVM 分类器的结构设计中, 类间的可分性(或类间的相似性)是衡量两个类是否容易分割的依据, 能否有效度量类间可分性直接决定了子类划分结构质量的高低. 在样本空间中, 不同类别具有不同的类域(包含类训练集样本的最小凸集). 如果类别的类域不相交, 则类别之间的可分性越大, 相似性越小, 越容易分割; 若类别的类域存在交叠, 则交叠部分所含的样本数占总样本数比例越大, 类间的可分性越小, 相似性越大, 越不易分割.

下面以图 1 所示的二维空间的 4 分类问题为例来讨论决策树 SVM 多类分类器的不同层次结构. 图 1 中, 每个圆表示各类样本在二维输入空间的分布图, 圆心表示类中心. 对于该 4 分类问题, 可以有 12 种不同的决策树层次结构.

图 2 是针对该 4 分类问题的 3 种不同的决策树结构, 其中: 每个非叶子节点代表一个两类 SVM, 每个叶子节点代表一个类. 图 2(a) 是根据各类之间欧氏距离极小极大法确定的层次结构^[7]; 图 2(b) 是按超球体最小包含类体积由大到小而确定的层次结

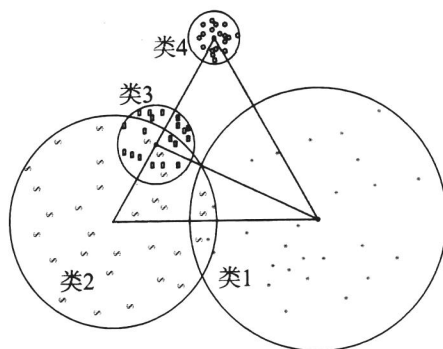
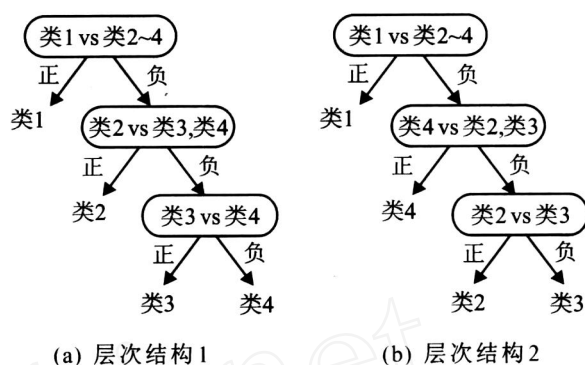


图 1 4 类样本分布图



(a) 层次结构 1

(b) 层次结构 2

(c) 层次结构 3

图 2 与图 1 对应的决策树 SVM 多分类器的层次结构^[8]; 图 2(c) 是根据类域之间的相交情况设计的层次结构.

从图 1 可看出, 类 4 与其他类都不相交, 因此是最易分割的类. 剩下 3 类中, 类 1 和类 3 不相交. 设 n_i 表示第 i 类的样本数, tr_{ij} 表示第 i 类和第 j 类相交的样本数, 记 $d_{12} = tr_{12} / (n_1 + n_2)$, $d_{23} = tr_{23} / (n_2 + n_3)$. 从图中可看出 d_{12} 小于 d_{23} , 因此类 1 为第 2 个易分割的类. 根据上述分析, 图 2(c) 对应的层次结构是最优的. 这表明, 对于某些问题而言, 根据类之间的相交情况来设计层次结构更为合理.

现在的问题是, 如何判断类别的类域是否相交; 如果不相交, 如何度量类间的可分性; 如果相交, 如何估计类间可能错分的样本数及度量类间的可分性. 向量投影法能较好地解决这个问题.

3 基于向量投影的类间可分性测度

向量投影的思想是李青等^[10]为了预选取两类样本的支持向量而提出的. 经研究发现, 通过向量投

影能较好地求出两类样本的类域之间相交的样本数.

为便于理解,先从线性空间的向量投影出发,给出其原理;然后将结论推广到非线性特征空间.

3.1 线性情况

首先给出一些定义.

定义 1 设第 i 类样本集 $X_i = \{x_1, x_2, \dots, x_{l_i}\}$, 定义第 i 类样本的平均特征为该类样本的样本中心点 m_i , 即

$$m_i = \frac{1}{l_i} \sum_{j=1}^{l_i} x_j. \tag{1}$$

定义 2 参考两类情况,将 m_1 到 m_2 的方向 $\overrightarrow{m_1 m_2}$ 定义为第 1 类模式的特征方向; m_2 到 m_1 的方向 $\overrightarrow{m_2 m_1}$ 定义为第 2 类模式的特征方向.

定义 3 假设样本点 x_i 到本类特征方向上的投影点为 x_i^o , 则 x_i^o 到本类样本中心 m_i 的欧氏距离

$$\overrightarrow{m_i x_i^o} = |x_i^o - m_i| \tag{2}$$

定义为特征距离.

用反证法容易证明以下定理:

定理 1 设样本集 $X_1 = \{x_1, x_2, \dots, x_{l_1}\}$ 和样本集 $X_2 = \{y_1, y_2, \dots, y_{l_2}\}$ 的样本中心距离 $d = |m_1 - m_2|$, 分别计算样本 x_i 和 y_j 的特征距离 $\overrightarrow{m_1 x_i}$ 和 $\overrightarrow{m_2 y_j}$. 令

$$r_1 = \max_{x_i \in X_1} (\overrightarrow{m_1 x_i}), \tag{3}$$

$$r_2 = \max_{y_j \in X_2} (\overrightarrow{m_2 y_j}), \tag{4}$$

则当 $r_1 + r_2 < d$ 时, 样本集 X_1 和 X_2 的类域是不相交的. 如果样本集 X_1 和 X_2 的类域相交, 则必有 $r_1 + r_2 \geq d$.

根据定理 1, 从类间可分性测度值越小所表示类的可分性越好的思路出发, 定义样本的可分性测度.

定义 4 当 $r_1 + r_2 < d$ 时, 定义类间可分性测度为

$$se_{12} = se_{21} = d - (r_1 + r_2). \tag{5}$$

当 $r_1 + r_2 \geq d$ 时, 设 X_1 中满足 $\overrightarrow{m_1 x_i} \geq d - r_2$ 且 $\overrightarrow{m_2 x_i} < d$ 的样本数为 tr_1 , X_2 中满足 $\overrightarrow{m_2 y_j} \geq d - r_1$ 且 $\overrightarrow{m_1 y_j} < d$ 的样本数为 tr_2 , 则定义类间可分性测度为

$$se_{12} = se_{21} = (tr_1 + tr_2) / (l_1 + l_2). \tag{6}$$

3.2 非线性情况

引理 1^[10] 已知两个向量 $x = (x_1, x_2, \dots, x_n)$ 和 $y = (y_1, y_2, \dots, y_n)$, 经非线性映射作用, 映射到特征空间 H 中, 则这两个向量在特征空间的 Euclidean 距离为

$$d^H(x, y) = \sqrt{k(x, x) - 2k(x, y) + k(y, y)}. \tag{7}$$

根据引理 1, 可计算第 i 类样本和第 j 类样本在特征空间的中心距离

$$d^H = \sqrt{k(m_i, m_i) - 2k(m_i, m_j) + k(m_j, m_j)}. \tag{8}$$

引理 2 已知 3 个向量 $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ 和 $z = (z_1, z_2, \dots, z_n)$, 经非线性映射作用, 映射到特征空间 H 中, 并且样本集在 H 空间是线性可分的, 设 (z) 在向量 (x) (y) 上的投影为 (z^o) , 则在特征空间中 (z^o) 到 (x) 的 Euclidean 距离为

$$\overrightarrow{(x) (z^o)} = \frac{k(z, y) - k(z, x) - k(x, y) + k(x, x)}{\sqrt{k(x, x) - 2k(x, y) + k(y, y)}}. \tag{9}$$

根据引理 2, 可计算样本在特征空间的特征距离. 将线性空间的类间可分性测度定义推广到非线性特征空间.

定义 5 设样本集 $X_1 = \{x_1, x_2, \dots, x_{l_1}\}$ 和样本集 $X_2 = \{y_1, y_2, \dots, y_{l_2}\}$ 在特征空间的中心距离 $d^H = |m_1 - m_2|$, 分别计算样本 x_i 和 y_j 在特征空间中的特征距离 $\overrightarrow{(m_1) (x_i^o)}$ 和 $\overrightarrow{(m_2) (y_j^o)}$. 令

$$r_1 = \max_{x_i \in X_1} (\overrightarrow{(m_1) (x_i^o)}), \tag{10}$$

$$r_2 = \max_{y_j \in X_2} (\overrightarrow{(m_2) (y_j^o)}), \tag{11}$$

则当 $r_1 + r_2 < d^H$ 时, 定义类间可分性测度为

$$se_{12} = se_{21} = d^H - (r_1 + r_2). \tag{12}$$

当 $r_1 + r_2 \geq d^H$ 时, 设 X_1 中满足 $\overrightarrow{(m_1) (x_i^o)} \geq d^H - r_2$ 且 $\overrightarrow{(m_2) (x_i^o)} < d^H$ 的样本数为 tr_1 , X_2 中满足 $\overrightarrow{(m_2) (y_j^o)} \geq d^H - r_1$ 且 $\overrightarrow{(m_1) (y_j^o)} < d^H$ 的样本数为 tr_2 , 则定义类间可分性测度为

$$se_{12} = se_{21} = (tr_1 + tr_2) / (l_1 + l_2). \tag{13}$$

4 决策树 SVM 多类分类器的设计

定义了类间可分性测度, 然后就是如何根据该测度设计决策树 SVM 多类分类器的层次结构. 因为本文所讨论的多类分类器的每个二值 SVM 分类器是由一个类与剩下的类构造分割面, 而且类间可分性测度值越小表示类的可分性越好, 所以将与其他各类间的可分性测度值之和最小的类判为最易分割的类. 为便于实现, 可用一个数组表示分类器的层次结构, 如用数组 $L = [4, 1, 2, 3]$ 表示图 2(c) 的层次结构.

设有 c 类训练样本集 $S_j, j = 1, 2, \dots, c$; 共 l 个训

训练样本对 $\{(x_i, y_i)\}, i = 1, 2, \dots, l$ 作如下约定: 如果样本 $x_i \in S_j$, 则 $y_i = j$. 整个分类器设计过程分为两步: 分类器层次结构的确定和各子分类器的训练.

4.1 分类器层次结构的确定

Step1: 根据定义5, 对 c 类样本求出类间可分性测度 $se_{ij} = se_{ji}, i, j = 1, 2, \dots, c, i \neq j$. 构造一个表示可分性测度的对称矩阵

$$SE = \begin{bmatrix} 0 & se_{12} & \dots & se_{1,c-1} & se_{1,c} \\ se_{21} & 0 & \dots & se_{2,c-1} & se_{2,c} \\ \dots & \dots & \ddots & \dots & \dots \\ se_{c-1,1} & se_{c-1,2} & \dots & 0 & se_{c-1,c} \\ se_{c,1} & se_{c,2} & \dots & se_{c,c-1} & 0 \end{bmatrix}.$$

Step2: 定义数组 $D_no = [1, 2, \dots, c]$, 设 i 初始值为1, 令 $SE(k, :)$ 表示 SE 的第 k 行, 对 j 从1到 $c-2$, 重复以下过程:

1) 对 SE 的每一行分别求和, 使 $k_0 = \arg \min_{k=1, \dots, c-1-j} \text{sum}(SE(k, :))$, 则 $L(i) = D_no(k_0)$. 如有多个和同为最小, 则取其中最先得的一行和为最小.

2) 去掉 SE 的第 k_0 行和第 k_0 列, 去掉 D_no 的第 k_0 个元素, $i = i + 1$.

Step3: 最后,

$$L(c-1) = D_no(1), L(c) = D_no(2).$$

4.2 各子分类器的训练

定义结构数组 $node$, 用来存放每个节点的信息 (包括支持向量, 权值, 域值 b 等). 对 j 从1到 $c-1$, 重复以下过程:

Step1: 将第 $L(j)$ 类样本集作为第 j 个子分类器的正类, 将剩下的第 $L(j+1), \dots, L(c)$ 类样本集合并, 作为子分类器的负类.

Step2: 预选取边界向量. 过程为: 首先将第 $L(j)$ 类样本分别与剩下的第 $L(j+1), \dots, L(c)$ 类样本进行两类样本间基于向量投影的支持向量预选取^[10]; 然后分别进行合并, 作为正类和负类的训练样本集.

Step3: 训练第 j 个两类 SVM 子分类器, 计算相

应的域值, 得到第 j 个节点 $node(j)$ 的结构信息.

5 仿真研究

以 Statlog^[11] 数据库中的大规模多类别数据集和 MNIST 数据库中的手写体数字识别为例, 对本文方法, one-against-one, 欧氏距离极小极大方法^[7] 和超球体最小包含方法^[8] 进行仿真研究和比较. 运行环境为: 主频 2.8 GHz, 内存 1 GB, 仿真工具是 Matlab 7.1. 二值 SVM 分类器是在动态链接库 SVM_V0.51^[12] 的基础上编程实现的.

选用 RBF 核函数

$$k(x, z) = \exp\left(-\frac{\|x-z\|^2}{2\sigma^2}\right).$$

仿真时需要设置的参数包括: 核参数 σ^2 , 惩罚因子 C , 边界向量预选取方法^[10] 中用到的参数 μ 和 D .

5.1 Statlog 数据库的大规模数据识别

表1列出了数据集的一些信息. 参数的确定采用交叉验证法, 最后得到每种方法相对最优的参数值如表2所示. 其中 $\text{ones}(1, n)$ 表示 $1 \times n$ 的全1行向量, σ^2 对应的数组值表示每个子分类器的核参数.

表1 实验数据集

数据集	训练样本数	测试样本数	类别数	属性数
Letter	15 000	5 000	26	16
Satimage	4 435	2 000	6	36
Shuttle	43 500	14 500	7	9

表3是各种方法在取最优参数值时的实验结果. 其中: T_c 表示测试时间, T_s 表示训练时间, $\#SVs$ 表示各二值 SVM 分类器的支持向量总和, u_SVs 表示不同支持向量的个数. 对于 Letter 和 Satimage 数据集, one-against-one 的每个 SVM 分类器没有预选取支持向量; 而对于 Shuttle 数据集, one-against-one 的每个 SVM 分类器都先预选取支持向量, 且参数为 $\mu = 0.05, D = 5$.

从表3可看出, 对于 Letter 数据集, one-against-one 的分类精度最好, 但测试时间最长, 约是其他方法的7倍, 从速度和分类精度综合考虑, 本文

表2 最优参数值

数据集	(C, D, μ) ² (数组)		
	本文算法	超球体最小包含	欧氏距离极小极大
Letter	(8, 5, 0.09) 32 × ones(1, 25)	(8, 5, 0.05) 32 × ones(1, 25)	(8, 5, 0.01) 32 × ones(1, 25)
Satimage	(1.5, 5, 0.05) [1 524, 1 524, 1 524, 512, 1 524]	(1.5, 5, 0.05) [1 524, 1 524, 1 524, 1 524, 1 524]	(1.5, 5, 0.05) [1 524, 1 524, 1 524, 512, 1 524]
Shuttle	(2 ¹² , 5, 0.05) [2 048, 2 048, 512, 2 048, 2 048, 2 048]	(2 ¹² , 5, 0.01) [2 048, 512, 512, 512, 512, 512]	(2 ¹² , 5, 0.05) [2 048, 2 048, 512, 2 048, 2 048, 2 048]

表 3 最优参数下实验结果比较

数据集	T_x/s		识别率 / %	$\# SV_s$		T_c/s		u_SVs				
	one-against-one	超球体最小包含		欧氏距离极大极小	本文算法							
Letter	397	33 204	97.4	2 423	7 150	96.1	616	7 114	95.6	462	7 241	96.3
	348	7 750		50	5 042		45	4 982		53	5 041	
satimage	60	3 404	91.8	43	2 191	91.2	13	2 208	92	14	2 216	91.8
	35	1 510		17	1 428		13	1 529		13	1 554	
shuttle	472	1 095	98.9	1 923	743	90.8	523	1 102	99.8	470	1 131	99.8
	26	334		9	332		27	451		27	455	

表 4 预选取边界向量和不预选取边界向量的实验结果比较 (Letter 数据集)

方 法	T_x/s		识别率 / %	$\# SV_s$		T_c/s		u_SVs	
	不预选取边界向量			预选取边界向量					
本文方法	3 916	7 389	96.4	462	7 241	96.3			
	58	5 087		53	5 041				
超球体最小包含	4 533	7 251	96.3	2 423	7 150	96.1			
	45	5 091		50	5 042				

方法相对较好;对于 Satimage 数据集,欧氏距离极大极小法相对最优,本文方法次之;对于 Shuttle 数据集,欧氏距离极大极小法和本文方法并列最优,超球体最小包含法的识别率最低,但分类速度最快.表 3 的结果表明,从分类精度和速度综合考虑,本文方法是一种有效的多类分类方法.

表 4 是 Letter 数据集用本文方法和超球体最小包含方法在预选取边界向量(本文方法 $\mu = 0.09, D = 5$;超球体最小包含方法 $\mu = 0.05, D = 5$)和不预选取边界向量时的实验结果比较.从表 4 可看出,对于大规模的数据集,预选取支持向量能显著提高 SVM 分类器的训练速度,而识别率仅略有降低.

5.2 手写体数字识别

MNIST 数据库中有 60 000 个训练样本,10 000 个检验样本.本文仅取其中的一部分进行训练和检验,训练样本取 3 000 个,测试样本取 500 个.实验数据的特征为每个样本 28×28 维,共 10 类.对所有样本归一化,即 $x = x/255$.经仿真发现,每种方法相对最优的参数一样.RBF 核函数的参数为 $C = 1 000$;一个决策树支持向量机各分类器的参数²相等,均为 32.仿真时,1-a-1 方法没有预选取边界向量,其他几种方法都先预选取边界向量,参数为 $\mu = 0.5, D = 5$.表 5 给出了仿真结果.从表 5 可看出,本文方法的识别率最高,测试时间最短;1-a-1 方法的训练时间最短,但测试时间最长,并且识别率不如本文方法高.这进一步表明本文方法是一种有效的多类分类方法.

表 5 手写体数字识别的仿真结果比较

方 法	T_x/s	T_c/s	$\# SV_s$	u_SVs	识别率 / %
1-a-1	134	248	6 851	1 673	94.6
超球体最小包含	282	93	2 795	1 621	93.8
欧氏距离极大极小	279	88	2 742	1 552	93.8
本文方法	278	84	2 594	1 544	94.8

6 结 论

本文首先介绍了决策树 SVM 多分类器的原理,举例说明并分析了其几种不同的层次结构,在此基础上给出类间可分性测度的重要性.然后考虑到类的分布情况,提出了一种基于向量投影的类间可分性测度的设计方法,并给出一种基于该类间可分性测度设计决策树 SVM 多类分类器层次结构的方法.最后,通过对 3 个大规模数据集和 MNIST 手写体数字识别的仿真实验,表明本文方法可获得满意的识别效果.该方法涉及到的参数选择问题和边界向量的预选取方法还有待进一步深入研究.

参考文献 (References)

[1] Vapnik V. The nature of statistical learning theory[M]. New York: Springer, 1995.
 [2] Weston J, Watkins C. Multi-class support vector machines[R]. London: Royal Holloway University of London, 1998.
 [3] Bottou L, Cortes C, Denker J. Comparison of classifier methods: A case study in handwriting digit recognition [C]. Proc of the 12th IAPR Int Conf on Pattern

- Reconiton. Jerusalem: IEEE, 1994: 77-82.
- [4] Kebel U. Pairwise classification and support vector machines[C]. *Advances in Kernel Methods — Support Vector Learning*. Cambridge: MIT Press, 1999: 255-258.
- [5] Platt J, Cristianini N, Shawe-Taylor J. Large margin DAG's for multiclass classification[C]. *Advances in Neural Information Processing Systems 12*. Cambridge, 2000: 547-553.
- [6] Sahbi Hichem, Geman Donald, Perona Pietro. A hierarchy of support vector machines for pattern detection[J]. *J of Machine Learning Research*, 2006, 7(10): 2087-2123.
- [7] Hsu C W, Lin C J. A comparison of methods for multiclass support vector machines [J]. *IEEE Trans on Neural Network*, 2002, 13(2): 415-425.
- [8] 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): 746-754.
(Tang F M, Wang Z D, Chen J Y. On multiclass classification methods of support vector machines [J]. *Control and Decision*, 2005, 20(7): 746-754.)
- [9] Wang X D, Shi Z W, Wu C M, et al. An improved algorithm for decision-tree-based SVM[C]. *Proc of the 6th World Congress on Intelligent Control and Automation*. Dalian, 2006: 4234-4237.
- [10] 李青, 焦季成, 周伟达. 基于向量投影的支撑向量预选取[J]. *计算机学报*, 2005, 28(2): 145-152.
(Li Q, Jiao L C, Zhou W D. Pre-extracting support vector for support vector machine based on vector projection[J]. *Chinese J of Computers*, 2005, 28(2): 145-152.)
- [11] Michie D, Spiegelhalter D, Taylor C. *Machine learning, neural and statistical classification*[DB/OL]. (1994). <http://www.niadd.liacc.up.pt/old/statlog/datasets.html>.
- [12] Platt J C. Fast training of support vector machines using sequential minimal optimization [DB/OL]. (1998-10-22). <http://research.microsoft.com/~jplatt>.

(上接第 744 页)

- [6] 刘思峰, 党耀国, 方志耕. 灰色系统理论及其应用[M]. 第3版. 北京: 科学出版社, 2004: 1-8.
(Liu S F, Dang Y G, Fang Z G. *Grey system theory and its application*[M]. 3rd ed. Beijing: Science Press, 2004: 1-8.)
- [7] 刘思峰. 冲击扰动系统预测陷阱与缓冲算子[J]. *华中理工大学学报*, 1997, 25(1): 25-27.
(Liu S F. The trap in the prediction of a shock disturbed system and the buffer operator [J]. *J of Huazhong University of Science and Technology*, 1997, 25(1): 25-27.)
- [8] 刘思峰. 缓冲算子及其应用[J]. *灰色系统理论与实践*, 1992, 2(1): 45-50.
(Liu S F. Buffer operator and its application [J]. *Theories and Practices of Grey System*, 1992, 2(1): 45-50.)
- [9] 谢乃明, 刘思峰. 一种新的弱化缓冲算子[J]. *中国管理科学*, 2003, 11(增): 46-48.
(Xie N M, Liu S F. A new applicative weakening buffer operator[J]. *Chinese J of Management Science*, 2003, 11(S): 46-48.)
- [10] 党耀国, 刘思峰, 刘斌, 等. 关于弱化缓冲算子的研究[J]. *中国管理科学*, 2004, 12(2): 108-111.
(Dang Y G, Liu S F, Liu B, et al. Study on the weakening buffer operators and researches[J]. *Chinese J of Management Science*, 2004, 12(2): 108-111.)
- [11] 党耀国, 刘斌, 关叶青. 关于强化缓冲算子的研究[J]. *控制与决策*, 2005, 20(12): 1332-1336.
(Dang Y G, Liu B, Guan Y Q. Study on the strengthening buffer operators [J]. *Control and Decision*, 2005, 20(12): 1332-1336.)
- [12] 中国统计局. *中国统计年鉴*[M]. 北京: 中国统计出版社, 1997-2005.
(The Chinese statistics bureau. *Chinese statistics yearbook*[M]. Beijing: Chinese Statistics Publishing House, 1997-2005.)