

支持向量机在分类中的应用

陆 波 尉询楷 毕笃彦

(空军工程大学工程学院, 西安 710038)

摘 要 通过引入结构风险最小化原则和最优分类面的概念,介绍了支持向量机及其用于非线性分类的基本原理和训练算法,并选用不同的核函数及参数对一组线性不可分的两类样本进行了划分识别,得到了较好的效果,并对结果进行了分析说明,展望了支持向量机的发展趋势。

关键词 支持向量机 最优分类面 非线性分类 二次规划

中图分类号: TP18 文献标识码: A 文章编号: 1006-8961(2005)08-1029-07

Applications of Support Vector Machines in Classification

LU Bo, WEI Xun-kai, BI Du-yan

(The Engineering Institute, Air Force Engineering University, Xi'an 710038)

Abstract This paper introduces the theory and training algorithm of the support vector machine which is applied in nonlinear classification and recognition by the way of bringing in the concept such as structural risk minimization principle and optimal hyperplane, then a set of nonlinear binary samples are successfully classified by using different kernel functions, followed by discussion to the results. After that current multi-class classification algorithms and application areas are reviewed. Finally future developments are prospected.

Keywords support vector machine, optimal hyperplane, nonlinear classification, quadratic programming

1 引言

基于数据的机器分类是现代智能技术中一个非常重要的方面。现有的机器分类方法其共同理论基础之一是统计学。传统统计学研究的是样本数量趋于无穷大时的渐进理论,现有的方法也多是基于此假设,但在实际问题中,样本数量往往是有限的,因此很多分类方法在实际应用中不尽人意。支持向量机是由 Vapnik 等人在 20 世纪 90 年代中期提出的一种机器学习算法,它以其良好的理论背景,从结构风险最小化原则出发为机器学习提供了一个崭新的方向。支持向量机是从线性可分情况下的最优分类面发展而来的,通过将输入空间映射到一个高维内积空间中,解一个线性约束的二次规划问题得到全

局最优解,有效避免了“维数灾难”,保证了收敛速度,而且不存在局部极小值问题。因此在解决小样本、非线性及高维模式识别问题中具有特有的优势。本文利用支持向量机技术对解决模式识别中的非线性分类问题进行了应用研究,并综述了现阶段支持向量机分类算法的一些新发展和应用。

2 基于支持向量机的分类

2.1 结构风险最小化原则

传统的分类方法只有在样本数目足够多的前提下,其性能才有理论上的保证。在实际应用中,人们通常采用“经验风险最小化(ERM)原则”,即通过求经验风险最小值来近似代替期望风险最小值。然而仍不能取得理想的效果,这是因为从期望风险最小化

收稿日期:2004-02-23; 改回日期:2004-12-29

第一作者简介:陆波(1979 ~),男。现为空军工程大学信号与信息处理专业硕士研究生。主要研究方向为图像处理与模式识别。
E-mail: yueboting@163.com

到经验风险最小化并无可靠的理论依据,只是人们直观上合理的想当然的做法^[1]。为此,Vapnik 等人推出了一种在有限样本情况下的较完善的机器学习理论体系——统计学习理论,它被认为是目前针对小样本统计估计和预测学习的最佳理论。统计学习理论提出了一种新的策略:把指示函数集^[2]构造为一个函数子集序列,使各个子集按照 VC 维^[3]的大小排列;在每个子集中寻找最小经验风险,在子集间折中考虑经验风险和置信范围^[2],以取得实际风险的最小。这种思想被称为“结构风险最小化(SRM)原则”,如图 1 所示。基于支持向量机的分类方法就是这种思想的具体实现,即设计函数集的某种结构使每个子集中都能取得最小的经验风险,然后选取适当的子集使置信范围最小,则这个子集中使经验风险最小的函数就是用于分类的最优判别函数。

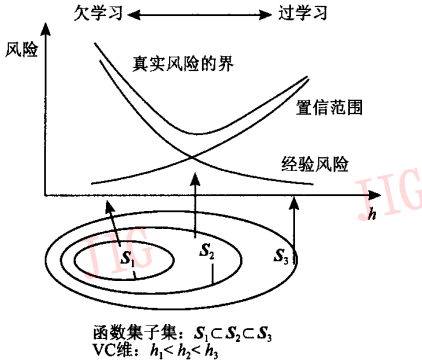


图 1 结构风险最小化示意图
Fig.1 Structure risk minimization

2.2 最优分类面

支持向量机理论是从线性可分情况下的最优分类面提出的,其基本思想可由图 2 所示的两类线性可分情况说明。

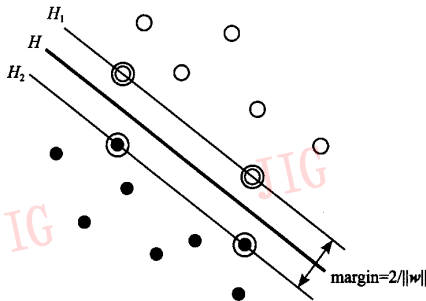


图 2 最优分类面示意图
Fig.2 Optimum separation plane

图中实心点和空心点分别表示两类训练样本, H 为分类线, H_1, H_2 分别为过各类样本中离分类线最近的点且平行于分类线的直线, H_1 和 H_2 之间的距离叫做分类间隔(图中用 margin 表示)。所谓最优分类线就是要求分类线不但能将两类无错误地分开,而且要使分类间隔最大。推广到高维空间,最优分类线就成为最优分类面。设分类线方程为 $x \cdot w + b = 0$, 将其归一化,使得对线性可分样本集 $(x_i, y_i), i = 1, \dots, n; x \in R^d; y \in \{+1, -1\}$, 满足:

$$y_i [(x_i \cdot w) + b] - 1 \geq 0 \quad (1)$$

此时分类间隔为 $2/\|w\|$, 因此使间隔最大等价于使 $\|w\|^2$ 最小, 满足式(1)且使 $\|w\|^2$ 最小的分类面就是最优分类面。过两类样本中离分类面最近的点且平行于最优分类面的超平面 H_1, H_2 上的训练样本(即使式(1)中等号成立的样本)叫做支持向量, 这是因为它们支撑了最优分类面, 如图 2 中用圆圈标出的点所示。利用 Lagrange 优化方法可以把上述最优分类面问题转化为其对偶问题: 即在以下约束条件下

$$\sum_{i=1}^n y_i a_i = 0 \text{ 和 } a_i \geq 0, i = 1, \dots, n \quad (2)$$

对 a_i 求解下列函数的最大值

$$Q(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j (x_i \cdot x_j) \quad (3)$$

a_i 为与每个样本对应的 Lagrange 乘子。这是一个不等式约束下二次函数的寻优问题, 存在唯一解, 解中将只有一部分 a_i 不为 0, 其对应的样本就是支持向量。解上述问题后得到的最优分类函数是:

$$f(x) = \text{sgn} \{ (w^* \cdot x) + b^* \} = \text{sgn} \left(\sum_{i=1}^n a_i^* y_i (x_i \cdot x) + b^* \right) \quad (4)$$

$\text{sgn}()$ 为符号函数。由于非支持向量对应的 a_i 均为 0, 因此式中的求和实际上只对支持向量进行, b^* 是分类阈值, 可用任一支持向量由式(1)求得。

上述最优分类面是在线性可分的前提下讨论的, 而实际情况往往是线性不可分的(某些样本不满足式(1))。此时, 可在式(1)中增加一个松弛项 $\xi \geq 0$, 则式(1)变成:

$$y_i [(x_i \cdot w) + b] - 1 + \xi_i \geq 0 \quad (5)$$

当样本 x_i 满足式(1)时, ξ 为 0; 否则 $\xi > 0$, 表示此样本为造成线性不可分的点。对式(5)利用 Lagrange 乘法及对偶原理可得到线性不可分条件下的对偶问题: 即在以下约束条件下,

$$\sum_{i=1}^n y_i a_i = 0, 0 \leq a_i \leq C, i = 1, \dots, n \quad (6)$$

对 a_i 求解式(3)的最大值。用与求解线性可分时同样的方法求解这一优化问题便可得到线性不可分情况下的最优分类面(也称为广义最优分类面),其结果与线性可分情况下得到的式(4)相同,只是 a_i 的取值范围变为 $[0, C]$, 其中 C 为事先选定的大于 0 的常数,它实际上起着对错分样本控制惩罚程度的作用,实现在错分样本数与算法复杂度之间的折衷,也称为惩罚系数。

2.3 非线性分类问题的支持向量机方法

以上讨论得到的最优和广义最优分类函数式(4)中只含有待分类样本与训练样本中的支持向量的内积运算 $(x \cdot x_i)$ 。可见,要解决一个特征空间中的最优线性分类问题,只需进行这个空间中的内积运算即可。由于很多实际模式识别问题在其定义的空间中不是线性可分的,对此可采用一个非线性变换 φ 将输入映射到一个特征空间 z , 其中 $z = \varphi(x)$ 。在特征空间 z 中,数据是线性可分的,故在 z 中可找到一个线性分类函数,它对应于输入空间中的一个非线性分类函数,要求解此时的最优分类函数涉及计算内积 $(z_i \cdot z_j) = \sum_{k=1}^m \varphi_k(x_i) \varphi_k(x_j)$, 若直接求解,不仅需要知道非线性映射 φ 的形式,而且计算量随着特征空间维数的增加呈指数增长,使求解难度加大甚至不能求解。因为优化问题和分类函数中涉及内积运算 $(z_i \cdot z_j)$, 根据 Hilbert-Schmidt 原理,可找到满足 Mercer 条件^[4]的核函数 $K(x_i, x_j)$, 使 $K(x_i, x_j) = (z_i \cdot z_j)$ 。则优化函数变为

$$Q(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j K(x_i, x_j) \quad (7)$$

相应的分类函数也变为

$$f(x) = \text{sgn} \left(\sum_{i=1}^n a_i^* y_i K(x_i, x) + b^* \right) \quad (8)$$

这就是解决非线性分类问题的支持向量机,其形式结构近似于神经网络,如图 3 所示。

不难看出,支持向量机的基本思想可概括为:首先通过非线性变换将输入空间映射到一个高维空间,然后在这个高维空间中求取最优线性分类面,非线性变换是通过定义适当的内积函数(核函数)实现的。选择不同形式的核函数就可以生成不同的支持向量机。常用的核函数有以下几种形式^[5]:

(1) 多项式核函数

$$K(x, x_i) = (1 + (x \cdot x_i))^p, p = 1, \dots, n \quad (9)$$

(2) 高斯径向基核函数

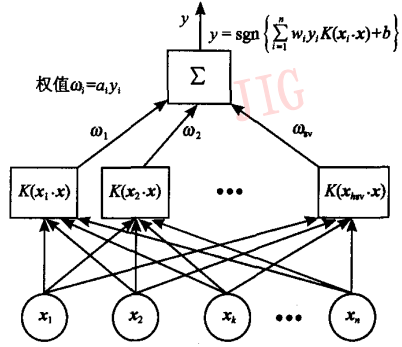


图 3 非线性支持向量机的一般结构

Fig. 3 Structure of nonlinear support vector machines

$$K(x, x_i) = e^{-\frac{\|x-x_i\|^2}{2\sigma^2}} \quad (10)$$

(3) 两层前馈神经网络核函数

$$K(x, x_i) = \tanh(v(x, x_i) - \theta) \quad (11)$$

2.4 支持向量机的实现

支持向量机的训练本质上是解决一个二次规划(QP)问题,即上述式(6)、式(7)。该问题可改写成以下矩阵形式:

给定训练样本 $(x_i, y_i), i = 1, \dots, n$; 核函数 $K(x_i, x_j)$ 和调节参数 C , 求二次型

$$Q(A) = A^T \mathbf{1} - \frac{1}{2} A^T H A \quad (12)$$

的极小值并满足下列约束条件:

$$A^T \mathbf{y} = 0, 0 \leq a_i \leq C \quad (13)$$

其中, $H = [h_{ij}]$, $h_{ij} = y_i y_j K(x_i, x_j), (i, j = 1, \dots, n)$; $A = (a_1, a_2, \dots, a_n)$; $\mathbf{y} = (y_1, y_2, \dots, y_n)$; $\mathbf{1}$ 为单位阵。

现有的 Chunking^[6], Osuna^[7], SMO^[8] 等训练算法均能有效地求解以上问题。本文采用了 Osuna 训练算法。其基本思想为:将一个 QP 问题分解成一系列较小规模的 QP 问题迭代求解,每次只处理部分训练数据,当所求的解都满足最优化(Kuhn-Tucker)条件^[9]时,便得到了优化问题的最优解。具体步骤如下:

(1) 从训练样本集中取出 m 个 (m 大于支持向量的数目) 样本构成工作集 B , 将其余 $n - m$ 个样本组成的集合记为 N , N 中所有样本的 Lagrange 乘子 a 均置为 0。

(2) 用 QP 方法对 B 求解最优化问题,得到支持向量(Lagrange 乘子 a 不为 0 的向量),并构成一个分类器。

(3) 用该分类器测试集合 N 中的样本,若 N 中

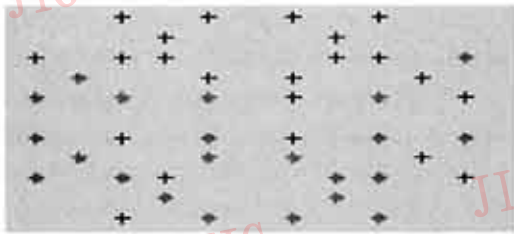
所有样本均满足最优化条件或 N 为空集则结束, 否则继续。

(4) 将 N 中至少一个不满足最优化条件的样本放入 B 中, 同时从 B 中取出同样数量的样本 (Lagrange 乘子 a 为 0) 放入 N 中; 返回步骤 2。

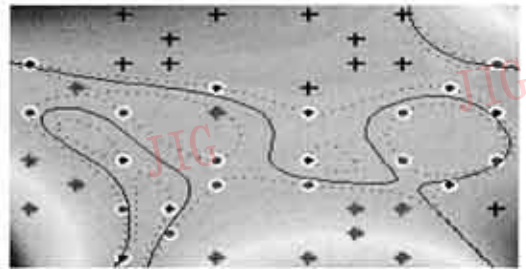
2.5 实例

在此分别采用多项式核函数和高斯径向基核函

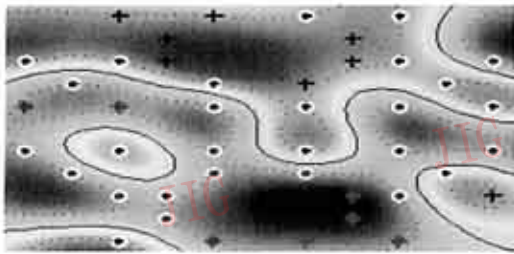
数对一组线性不可分的两类样本进行分类识别, 得到的结果如图 4 所示。两类样本分别用“+”和“*”表示。图中用圆圈圈出的样本为支持向量, 可以看出, 用支持向量机求得的最优分类线(实曲线)均能将样本正确地划分开。所选用的核函数、参数及分类结果如表 1 所示。



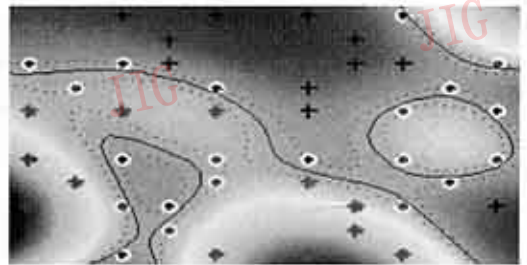
(a) 两类样本数据的分布



(b) 采用 6 阶多项式核函数
($P=6, C=Inf$)



(c) 高斯径向基核函数
($\sigma=0.5, C=10$)



(d) 高斯径向基核函数
($\sigma=1, C=Inf$)

图 4 对线性不可分样本的分类结果

Fig. 4 Classification results of overlapping data

表 1 两种支持向量机及不同参数的分类结果

Tab. 1 Classification results

支持向量机类型	核函数中的参数	支持向量个数	测试误差 (%)
多项式	$P=3, C=Inf$	28	3.66
高斯径向基	$\sigma=1, C=Inf$	28	0.04
	$\sigma=0.5, C=10$	21	0.90

由于 C 值表征的是经验风险与 VC 维(即推广能力)之间的权衡, 而当 C 取无穷 (Inf) 时, 相当于仅考虑其推广能力。从表中可以看出, 在相同的惩罚 C 作用下, 径向基函数具有更好的分类结果和最佳的推广能力, 且其卓异的分类精度是传统方法难以企及的。另外, 选择不同的参数也将导致分类精度的不同, 关于参数的选择可参考 Cross-validation 方

法^[10], 在此不作详述。

3 多分类算法

由于现实世界中遇到的问题多为多类问题, 而且由于支持向量机是针对二分类问题设计的算法, 因此如何将多类问题转化成为二分类的组合就成为支持向量机研究的热点问题^[11-16]。

3.1 一对余型多类支持向量机方法

训练集 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 其中 $x_i \in \mathbf{R}^n, y_i \in \{1, 2, \dots, k\}, i = 1, \dots, n$, 算法描述为对于 k 分类问题, 取所有 $y = i$ 和 $y \neq i$ 的训练样本可以构建 $M = C_k^2 = k$ 个训练子集 T_i 。对 $i = 1, \dots, k$, 则由第 i 类和其余 $(M - 1)$ 类训练数据便可训练式(8)

所示二值分类器,分别使用 SVM 二值分类算法对这 M 个训练集进行学习,便可以构造 M 个分类器。

测试阶段,将测试样本输入到 M 个支持向量机,得到决策函数集 $\{g_i(x)\}$ 和判别函数集 $\{q_i(x)\}, i=1, \dots, k$, 则将其归为决策函数 $g_i(x)$ 最大的所对应的类别指示值。

3.2 一对一型多类支持向量机方法

对于 k 分类问题,取所有 $y=i$ 和 $y=j$ 的训练样本构建 $M=C_k^2$ 个训练子集 T_{ij} 。对 $i=1, \dots, k, j=i+1, \dots, k$, 则由第 i 类和第 j 类数据训练数据便可训练二值分类器,分别使用 SVM 二值分类算法对这 M 个训练集进行学习,便可以构造 M 个分类器。测试阶段的判别采用投票决策法 (Voting strategy), 用所有的 M 个分类器对测试样本 x 进行分类,在第 i 类和第 j 类之间分类时,记 v_{ij} 为投票函数,若该分类器判断 x 属于 i 类,则 i 类的票数加 1, 否则 j 类的票数加 1。最后将测试样本 x 归为得票数最多的类别。

3.3 DDAG 型多类支持向量机方法

对于一对一算法,Plat 等人提出了一种新的学习构架:决策导向无环图 (decision directed acyclic graph, DDAG)。对于 k 分类问题, DDAG 也含有 $k(k-1)/2$ 个二值分类器,而在对多个一对一分类器进行组合的过程中,引入了图论中的有向无环图 (DAG) 的思想。训练方法与一对一方法基本类似,测试时按照 DDAG 图进行计算。

3.4 K 类支持向量机方法

K 类支持向量机算法与一对多方法类似,它也是构造 k 个支持向量机决策函数,不同的是所有支持向量机的参数是通过求解一个二次规划问题获得。

3.5 LS-SVM 型多类支持向量机方法

Suykens 将二分类 LSSVM 方法推广到了多类别问题,本质上还是将二次规划的求解问题转化成为线性规划问题。

3.6 ECOC 型多类支持向量机方法

ECOC 编码是由 Bose 和 Chaudhuri 在 1960 年提出的一种分布式输出码,1963 年, Duda 将其应用于模式识别领域中,1995 年 Dietterich 和 Bakiri 将 ECOC 引入到多类别模式识别问题中。

通过一类对余类,或者一类对一类的方式构造二分类问题,当然也可以把奇数类看作是正类 (+1),把偶数类看作是负类 (-1),或者采用某种特定适宜实际分类问题情况的构造方法,这样就可

以得到一系列的两类问题,总共可以得到 L 个决策函数集 $\{f_1, \dots, f_L\}$ 。按照设定的对应规则,可以将每一类与一系列长度为 L 、值为 +1 或者 -1 的编码相对应,按照类别的顺序就可以得到一个 k 行 L 列的编码表。

由 +1 或者 -1 组成的编码矩阵记为 $M_{k \times L}$, 其中, k 为类别数, L 为编码长度也就是分类器的个数。当 $m_{ij} = +1(-1)$ 时,表示此样本相对于第 i 类而言是作为正(负)类训练第 j 个分类器 f_j 。这样就可以得到决策函数集 $f(x) = \{f_1(x), f_2(x), \dots, f_L(x)\}$, 而测试过程中,对于测试样本 x ,只需要计算各个分类器的输出向量与各类别向量的距离,使其最小的类就是样本所属的类。

4 基于支持向量机分类的应用

4.1 与新理论的综合应用

随着混沌理论、小波理论、模糊理论、粗糙集理论等一批新兴的智能学科的兴起,支持向量机算法在应用过程中也出现了一些新的发展趋势。

(1) 与混沌理论的综合应用

采用混沌型支持向量机的时序分析框架,利用混沌特征参数选择合适的预测器结构,然后在进行预测分析时取得了很好的效果^[17]。

(2) 与小波理论的综合应用

基于小波和支持向量机的滚动轴承状态监测方法,利用小波能量提取信号特征,然后利用支持向量机进行分类,取得了较为满意的结果^[17]。

(3) 与模糊理论的综合应用

模糊一对多方法和模糊一对一方法,有效地消除了不可分边界,较好地提高了多类支持向量机的分类精度^[18-21]。

(4) 与粗糙集理论的综合应用

采用粗糙集进行特征属性约减的多类支持向量机方法,并将其应用到城市污水监控中去,取得了较好的效果,并具有一定的容错性^[22]。

4.2 应用领域

近来,随着支持向量机理论的完善和快速发展,支持向量机在众多领域得到了广泛的应用,例如手写体识别^[22],人脸识别与检测,计算机系统入侵监测、文本分类,基因、蛋白质分析,语音识别,图像分类,故障诊断^[23]等。

5 支持向量机的发展趋势

对于支持向量机算法,目前尚存在许多需要解决的问题,主要有以下几方面需要着重研究:

(1)核的选取 目前主要侧重于构造核的方法的研究上,而如何选择合适的核则是支持向量机算法理论上需要解决的主要问题。关于支持向量机的核构造目前有很多文献可以参阅,其中文献[24]有较详细的介绍。

(2)多类(变量)问题的研究 目前对于模式分类问题已有的算法主要有:一对一方法(one-against-one)、一对多方法(one-against-all)、成对分类法(pair wise)、纠错码型(error code correcting)、决策树型(decision tree)、无环导向图型(DDAG)以及球形支持向量机(sphere-structured SVM),也有一些采用组合支持向量机进行决策分类的新方法,如采用 Logistic 回归进行 SVM 组合的方法等。如何选择恰当的组合方法获得满意的推广性能是支持向量机用于多分类问题中需要深入研究的方向。

(3)模型选择 核函数确定的情况下,目前最常用的支持向量机推广误差的估计方法是 K 层交叉检验法(cross-validation)和留一法(leave-one-out, LOO),常用的 LOO 推广误差的评定界主要有: Jiaakkola-Haussler 上界、Joachimes 上界,以及 Vapnik 上界,其中, Vapnik 上界由于其计算简单,所以实际使用得比较多,但是这些界都存在一个问题:估计得到的界往往大于实际中的情况,使得有时获得的推广界根本无法使用。这也是支持向量机推广性估计中亟需解决的重要问题。

(4)高效、快速的实现方法 目前主要的支持向量机实现方法有:内点算法(interior point), Newton 梯度法,以及求解大型问题的 Chunking 选块算法,分解算法(Decomposing),序贯最小优化方法(SMO, sequential minimal optimization)等,对于大规模、超大规模数据集这些算法在一定程度上都不能够满足快速计算的要求,因此开发适合于大型、超大型数据集的海量计算方法就显得十分重要。

参考文献(References)

1 BIAN Zhao-qi, ZHANG Xue-gong. Pattern Recognition [M]. Beijing: Tsinghua University Press, 2000. [边肇祺,张学工. 模式识别[M]. 北京:清华大学出版社,2000.]

2 Vapnik V N. The Nature of Statistical Learning Theory [M]. New York: Springer, 1995.

3 Vapnik V, Levin E, Le Cun Y. Measuring the VC-dimension of a learning machine [J]. Neural Computation, 1994, 8(5): 51 ~ 76.

4 Burges C. A tutorial on support vector machines for pattern recognition [J]. Data Mining and Knowledge Discovery, 1998, 2(2): 1 ~ 47.

5 Osuna E, Freund R. Support vector machines [A]. In: Training and Applications [R], Technical Report Alm-1602, Cambridge MA, USA: MIT, 1997, 2: 527 ~ 531.

6 Vapnik V. Estimation of Dependence Based on Empirical Data [M], New York: SpringerVerlag, 1982.

7 Osuna E, Freund R. An improved training algorithm for training support vector machines [A]. In: Proceedings of IEEE Workshop on Neural Networks for Signal Processing [C], New York, 1997, 1: 252 ~ 256.

8 Platt J. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines [M]. Cambridge: The MIT Press, 1998: 169 ~ 182.

9 Fletcher R. Practical Methods of Optimization [M]. New York: Wiley and Sons Press, 1987.

10 Alexander Smola. Regression Estimation with Support Vector Learning Machines [EB/OL]. <http://www.first.gmd.de/smola,2000-06-05>.

11 Allwein L, Robert E. Reducing multiclass to binary: A unifying approach for margin classifiers [J]. Journal of Machine Learning Research, 2000, 1(1): 113 ~ 141.

12 Pairwise. Classification and Support Vector Machines [M]. Cambridge: The MIT Press, 1999.

13 Platt J C, Cristianini N, Shawe-Taylor J. Large margin DAGs for multiclass classification [J]. Advances in Neural Information Processing Systems, 2000, 12(12): 547 ~ 553.

14 Suykens J A K, Vandewalle J. Multiclass Least Squares Support Vector Machines [R], TR1999-012. Washington DC, America: Washington Institute of Technology, 1999.

15 Kindermann J, Leopold E, Paass G. Multi-Class Classification with Error Correcting codes [EB/OL]. <http://www.first.gmd.2004-11-04>.

16 Chih-Wei Hsu, Chih-Jen Lin. A comparison of methods for multiclass support vector machines [J]. IEEE Transactions on Neural Networks, 2001, 13(2): 415 ~ 425.

17 LI Ying-hong, WEI Xun-kai, LIU Jian-xun. Engineering Applications of Support Vector Machines [M]. Beijing: Publishing House of Ordnance Industry, 2004. [李应红,尉询楷,刘建勋. 支持向量机的工程应用[M]. 北京:兵器工业出版社,2004.]

18 Inoue T, Abe S. Fuzzy Support Vector Machines for Pattern Classification [R]. TR2001-005, WashingtonDC, America: Washington Institute of Technology

19 Lin Chunfu, Wang Shengde. Fuzzy support vector machines [J]. IEEE Transactions on Neural Networks, 2002, 13(3): 464 ~ 471.

20 Huang Han-Pang, Liu Yi-Hung. Fuzzy support vector machines for pattern recognition and data mining [J]. International Journal of

- Fuzzy Systems, 2002, 14(3):25 ~ 28.
- 21 FAN Xi-wei, DU Shu-xing. Rough support vector machine and its application to waste water treatment processes [J]. Control and Decision, 2004, 19(5): 573 ~ 576. [范昕伟, 杜树新. 粗 SVM 分类方法及其在污水处理过程中的应用[J]. 控制与决策, 2004, 19(5): 573 ~ 576.]
- 22 WEI Xing-guo. Research on Handwritten Numerical Digits Recognition Based on Kernel-based Learning Algorithm [D]. Nanjing: Nanjing University of Science and Technology, 2003. [魏兴国. 基于核方法的手写体数字识别 [D]. 南京: 南京理工大学, 2003.]
- 23 WEI Xun-kai, LI Ying-hong. Applications of support vector machines in aeroengine fault diagnosis [J]. Journal of Aerospace Power, 2004, 19(6): 844 ~ 848. [尉询楷, 李应红. 支持向量机在航空发动机故障诊断中的应用[J]. 航空动力学报, 2004, 19(6): 844 ~ 848.]
- 24 Ayat N E, Cheriet M. A new support vector machine kernel for pattern recognition, application to digit image recognition [A]. In: Proceedings of the 6th International Conference on Document Analysis and Recognition [C], Seattle, USA, 2001: 1215 ~ 1219.