

文章编号: 1001-0920(2011)04-0623-06

具有磁场效应的 ρ -间隔核学习机

陶剑文^{1,2}, 王士同¹

(1. 江南大学 信息工程学院, 江苏 无锡 214122; 2. 浙江工商职业技术学院 信息工程学院, 浙江 宁波 315012)

摘要: 针对模式分类问题, 提出一种具有磁场效应的 ρ -间隔核学习机 (ρ -MKLM), 旨在寻求一个具有磁场效应的最优超平面, 受其吸引, 使得一类模式离该平面的距离尽可能的小, 而另一类模式受其排斥, 离该平面的间隔尽可能的大, 从而最大可能地实现模式分类. 通过引入一个可调节的磁场强度 q , 减小一类模式的数据分布, 从而提升分类性能. 分别采用人工数据和实际数据进行实验, 所得结果显示, ρ -MKLM 在一类和二类模式分类上的性能均优于或等同于相关方法.

关键词: 磁场效应; 模式分类; 新奇检测; 支持向量机; 核方法

中图分类号: TP181

文献标识码: A

ρ -margin kernel learning machine with magnetic field effect

TAO Jian-wen^{1,2}, WANG Shi-tong¹

(1. School of Information Engineering, Southern Yangtze University, Wuxi 214122, China; 2. School of Information Engineering, Zhejiang Business Technology Institute, Ningbo 315012, China. Correspondent: TAO Jian-wen, E-mail: tjw@zjbt.net.cn)

Abstract: A ρ -margin kernel learning machine (ρ -MKLM) with magnetic field effect is proposed for pattern classification problem in this paper. The basic idea is to find a optimal superplane with magnetic field effect such that the distance between one class and the hyperplane is as small as possible due to the magnetic attractive effect, while at the same time the margin between the hyperplane and the other classes is as large as possible due to magnetic repulsion, thus implementing pattern classification as much as possible. Moreover, a magnetic field density q is introduced to compact the data distribution of one class, thus improving the classification performance of ρ -MKLM even more. Experimental results obtained with synthetic and real data show that the proposed algorithms are effective and competitive to other related diagrams in such cases as two-class and one-class pattern classification respectively.

Key words: magnetic field effect; pattern classification; novelty detection; support vector machine; kernel approach

1 引言

模式分类旨在通过有限的训练样本学习一个分类器, 且该分类器对于未来数据必须具有良好的泛化能力^[1]. 目前已有多种模式分类方法, 其中, 支持向量机 (SVM) 及其相关变体是当前实现模式分类的主流方法之一^[2], 其通过最大化数据类间的分割间隔来达到强泛化能力. 最初 Vapnik 等人^[3]提出 SVM 主要用于模式分类 (C-SVC). C-SVC 根据一个支持向量的典型训练样本小子集来表示决策边界^[4]. 后来, Schölkopf 等人^[5]提出一种 SVM 的扩展变体 ν -SVM, 通过引入一个新的参数 ν 来控制支持向量个数的下界和训练误差的上界, 该参数能取代 SVM 中的参

数 C . 基于扩展 SVM 的大间隔思想, 利用一类支持向量机 (OCSVM)^[6]方法在高维特征空间中构建一个具有最大间隔的超平面, 用于分割正常和异常数据. 同样, 对于一类分类 (OCC) 问题, Tax 等人^[7]提出一种支持向量数据描述 (SVDD) 的方法, 以寻求一个包含所有目标样本的最小超球体, 其核化版本等价于基于核集理论的最小包含球 (MEB) 问题^[8]. Ye 等人^[9]综合 SVDD 中的小球体和 OCSVM 中的大间隔思想, 提出一种用于新奇检测的小球体和大间隔方法 (SSLM), 该方法在训练样本中利用了少量异常样本信息.

虽然大间隔分类方法 (典型代表是 SVM 及其变体) 在实际应用中已取得了较好的效果, 但这种绝对

收稿日期: 2010-01-19; 修回日期: 2010-09-20.

基金项目: 国家自然科学基金项目(60975027, 60903100); 宁波市自然科学基金项目(2009A610080).

作者简介: 陶剑文(1973-), 男, 副教授, 博士生, 从事模式识别与数据挖掘等研究; 王士同(1964-), 男, 教授, 博士生导师, 从事人工智能、机器学习等研究.

间隔的分类解很容易受到输入数据的仿射或伸缩等变换的干扰^[10],其原因是基于大间隔思想的分类方法只考虑数据类间的绝对间隔而忽视了类数据的分布性状,因此大间隔分类方法只能得到次优化性能.近年来,许多模式分类研究都将类间间隔和类内聚类性作为影响分类器分类性能的两个重要因素^[11,1,10],增大类间间隔和提高类内内聚性(或减小数据分布)能增强分类性能.

一般来说,在类球形数据分布(或某个核函数映射后呈球形数据分布)情况下,超球SVM方法较无约束线性SVM具有更严格的界^[1],但是数据呈球形分布本身是一个较为严格的约束,大多情况下,即使经过某个核函数映射后,输入数据的整体分布形态也并非是完全球形分布,利用一个最小包含超球很难实现一类数据的类内内聚性的最大化.对此,受磁场效应理论(MFET)和SSLM等思想的启发,结合超球SVM的类内内聚性和传统线性SVM的大间隔思想,本文提出一种新颖的具有磁场效应的 ρ -间隔核学习机(ρ -MKLM). ρ -MKLM旨在学习一个具有磁场效应的最优超平面,使得一类(或正类)尽可能靠近超平面,并使数据分布最小化;而另一类(或负类)距离超平面的 ρ -间隔则尽可能的大. ρ -MKLM的分割超平面犹如一个磁场强度为 q 的磁场,对一类模式进行吸引而对另一类模式进行排斥,从而使得 ρ -MKLM具有较强的模式分类性能.

2 ρ -MKLM

2.1 线性形式

对于一个包含 N 个模式的二类分类问题,设给定训练集 $T = \{(x_1, y_1), \dots, (x_N, y_N)\}$.其中: $x_i \in X \subset R^d (1 \leq i \leq m_1 + m_2 = N)$ 为输入数据; $y_i \in \{+1, -1\}$ 为类标签,且当 $1 \leq i \leq m_1$ 时, $y_i = 1$;当 $m_1 + 1 \leq i \leq m_1 + m_2$ 时, $y_i = -1$.又设其中一类(或正类)含有 m_1 个模式 $\{x_i, y_i\}_{i=1}^{m_1}$,另一类(或负类)含有 m_2 个模式 $\{x_i, y_i\}_{i=m_1+1}^{m_1+m_2}$.两类模式间的最小距离称为 ρ -MKLM的 ρ -间隔,如图1所示.

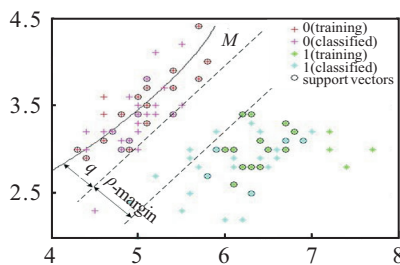


图1 具有磁场效应的 ρ -间隔学习机

为了正确分割该二类模式,且保证强的泛化能力, ρ -MKLM依据磁场效应原理构建一个超平面,使

得一类紧附于该平面,而另一类则与该平面的间隔最大化. ρ -MKLM优化问题可描述为

$$\min_{w, \rho, \xi, b} \frac{1}{2} \|w\|^2 - \nu\rho + \frac{1}{v_1 m_1} \sum_{i=1}^{m_1} \xi_i + \frac{1}{v_2 m_2} \sum_{j=m_1+1}^N \xi_j. \quad (1)$$

s.t.

$$w^T x_i + b \leq q + \xi_i, \quad 1 \leq i \leq m_1; \quad (2)$$

$$w^T x_j + b \geq q + (2 - q)\rho - \xi_j, \quad m_1 + 1 \leq j \leq N; \quad (3)$$

$$\xi \geq 0, \quad \rho \geq 0, \quad q > 0. \quad (4)$$

其中: $\xi = [\xi_1, \xi_2, \dots, \xi_N] \in R^d$ 为松弛向量; v, v_1, v_2 为3个正常数; $q \in (0, 2)$ 为磁场强度,是一可调参数.根据Lagrangian定理^[12],上述 ρ -MKLM原始问题的Lagrangian方程为

$$\begin{aligned} L(w, \rho, \xi, b, \alpha, \beta, \lambda) = & \frac{1}{2} \|w\|^2 - \nu\rho + \frac{1}{v_1 m_1} \sum_{i=1}^{m_1} \xi_i + \\ & \frac{1}{v_2 m_2} \sum_{j=m_1+1}^N \xi_j + \sum_{i=1}^{m_1} \alpha_i (w^T x_i + b - q - \xi_i) - \\ & \sum_{j=m_1+1}^N \alpha_j (w^T x_j + b - q - (2 - q)\rho + \xi_j) - \\ & \sum_{k=1}^N \beta_k \xi_k - \lambda\rho. \end{aligned} \quad (5)$$

其中: $\alpha_i \geq 0, \beta_i \geq 0, \lambda \geq 0$ 分别为Lagrangian乘子.在 $L(w, \rho, \xi, b, \alpha, \beta, \lambda)$ 方程中,分别对 w, ρ, ξ, b 等原始变量求偏导数,并令各偏导方程为零,求得原始问题的对偶形式为

$$\min_{\alpha \in R^n} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j. \quad (6)$$

s.t.

$$0 \leq \alpha_i \leq \frac{1}{v_1 m_1}, \quad 1 \leq i \leq m_1; \quad (7)$$

$$0 \leq \alpha_j \leq \frac{1}{v_2 m_2}, \quad m_1 + 1 \leq j \leq N; \quad (8)$$

$$\sum_{i=1}^N \alpha_i y_i = 0; \quad (9)$$

$$\frac{2\nu}{2 - q} \leq \sum_{i=1}^N \alpha_i \leq \frac{1}{v_1} + \frac{1}{v_2}; \quad (10)$$

$$w = - \sum_{i=1}^N \alpha_i y_i x_i. \quad (11)$$

对偶式(6)~(11)为一个二次规划问题^[12],具有与标准 ν -SVM相同的优化形式.因此, ρ -MKLM能直接利用解 ν -SVM的软件包来实现,从而使得 ρ -MKLM算法易于实现.

2.2 非线性形式

在非线性情况下, 通过一个满足 Mercer 条件的核函数(如径向基函数)对输入空间进行高维映射^[13], 然后在高维空间进行模式分类. ρ -MKLM 的核化形式为

$$\min_{w, \rho, \xi, b} \frac{1}{2} \|w\|^2 - \nu\rho + \frac{1}{v_1 m_1} \sum_{i=1}^{m_1} \xi_i + \frac{1}{v_2 m_2} \sum_{j=m_1+1}^N \xi_j. \quad (12)$$

$$\text{s.t. } w^T \phi(x_i) + b \leq q + \xi_i, \quad 1 \leq i \leq m_1; \quad (13)$$

$$w^T \phi(x_j) + b \geq q + (2 - q)\rho - \xi_j, \quad m_1 + 1 \leq j \leq N; \quad (14)$$

$$\xi \geq 0, \quad \rho \geq 0, \quad q > 0. \quad (15)$$

ρ -MKLM 核化对偶解为

$$\min_{\alpha \in R^n} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j). \quad (16)$$

$$\text{s.t. } 0 \leq \alpha_i \leq \frac{1}{v_1 m_1}, \quad 1 \leq i \leq m_1; \quad (17)$$

$$0 \leq \alpha_j \leq \frac{1}{v_2 m_2}, \quad m_1 + 1 \leq j \leq N; \quad (18)$$

$$\sum_{i=1}^N \alpha_i y_i = 0; \quad (19)$$

$$\frac{2\nu}{2 - q} \leq \sum_{i=1}^N \alpha_i \leq \frac{1}{v_1} + \frac{1}{v_2}; \quad (20)$$

$$w = - \sum_{i=1}^N \alpha_i y_i \phi(x_i). \quad (21)$$

其中 $K(u, v)$ 为某个符合 Mercer 条件的核函数^[13].

2.3 间隔 ρ 与偏置 b 的求解

考虑两个支持向量集合

$$SV_1 = \left\{ x_i \mid 0 < \alpha_i < \frac{1}{m_1 v_1}, 1 \leq i \leq m_1 \right\},$$

$$SV_2 = \left\{ x_j \mid 0 < \alpha_j < \frac{1}{m_2 v_2}, m_1 + 1 \leq j \leq N \right\}.$$

根据 KKT 条件^[12], 对于输入 $x_i \in SV_1 (1 \leq i \leq m_1)$, 式(2)变成一个所有松弛变量均为 0 的等式, 即

$$w^T \phi(x_i) + b = q, \quad 1 \leq i \leq m_1; \quad (22)$$

同理, 对于输入 $x_j \in SV_2 (m_1 + 1 \leq j \leq N)$, 式(3)也变成一个等式, 即

$$w^T \phi(x_j) + b = q + (2 - q)\rho, \quad m_1 + 1 \leq j \leq N. \quad (23)$$

设 $n_1 = |SV_1|, n_2 = |SV_2|, |\cdot|$ 表示集合基数. 由式(22)和(23), 可得

$$\rho^* = \frac{1}{2 - q} \left[\frac{M_2}{n_2} - \frac{M_1}{n_1} \right], \quad (24)$$

$$b^* = q - \frac{M_1}{n_1}. \quad (25)$$

其中

$$M_1 = \sum_{\alpha_i \in SV_1} \sum_{x_i, x_j \in SV_1} \alpha_i K(x_i, x_j),$$

$$M_2 = \sum_{\alpha_j \in SV_2} \sum_{x_i, x_j \in SV_2} \alpha_j K(x_i, x_j).$$

2.4 决策函数

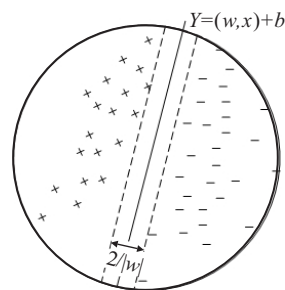
为了测试一个新模式 $x \in R^n$ 的类别, 利用 ρ -MKLM 判断其距磁场效应超平面的距离是否小于磁场效应强度 q , 即 ρ -MKLM 的决策函数为

$$f(x) = \text{sgn}(q - \|w^* T \phi(x) + b^*\|). \quad (26)$$

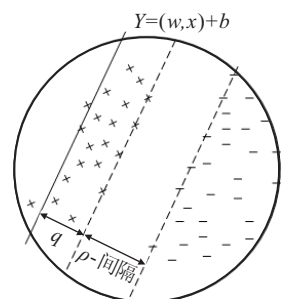
在求得优化的 α 值后, w^* 可由式(21)求得, b^* 可由式(25)求得.

3 性能分析

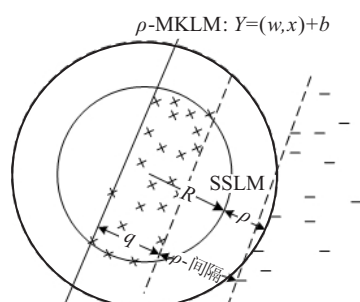
线性 SVM 目标函数为 $\min \frac{1}{2} \|w\|^2$, 类间间隔为 $\Delta = 2/\|w\|$, 通过最大化类间间隔以提高分类器的泛化能力^[3]. SSLM 通过目标函数 $\min R^2 - \nu\rho^2$ 来同时达到一类类内内聚和类间间隔最大化, 提高球形一类的分类能力^[9]. ρ -MKLM 通过目标函数 $\min \frac{1}{2} \|w\|^2 - \nu\rho$ 和一个可调节的磁场强度变量 q 来同时实现最大化类内内聚性和类间间隔. 为了说明 3



(a) SVM的最小包含球模型



(b) ρ -MKLM的最小包含球模型



(c) SSLM与 ρ -MKLM 的比较

图 2 SVM, ρ -MKLM, SSLM 三种方法的性能比较

种方法的不同泛化性能,构造一个最小包含球(MEB)以包含SVM和 ρ -MKLM的所有训练样本;而对于SSLM,MEB仅包含正常样本.图2(a)~图2(c)分别显示了这3种情况的分类模型.其中:“+”形数据为一类(或正类),“-”形数据为另一类(或负类).图2(a)显示,SVM仅通过类间间隔最大化来提升分类性能;图2(b)中, ρ -MKLM通过调节磁场强度 q (即使得某一类数据分布尽量最小)和最大化 ρ -间隔(即最大化类间间隔),明显提升了分类性能;图2(c)显示,虽然SSLM能实现一类(正类)的类内内聚性,但对于类线性分布数据, ρ -MKLM提升一类(正类)类内内聚性明显高于SSLM,从而使得 ρ -MKLM比SSLM具有更严格的分类边界.

对于 ρ -MKLM的分类性能,有如下定理:

定理1 ρ -MKLM能同时实现最大化类间间隔和正负类类内内聚性.

定理1表明, ρ -MKLM方法具有增强的模式分类性能,这可由后面的实验部分得以验证.

按照文献[9]的术语,本文称对应于Lagrangian乘子 $\alpha_i > 0$ 的训练样本 $x_i(1 \leq i \leq N)$ 为支持向量(SV),对应于松弛变量 $\xi_j > 0$ 的训练样本 $x_j(1 \leq j \leq N)$ 为间隔误差(ME).

定理2 设 m^+, m^- 分别为正类和负类的间隔误差数; s^+, s^- 分别为正类和负类的支持向量数; ρ -MKLM中参数 v, v_1, v_2, q 能同时控制支持向量的下界和间隔误差的上界.即存在如下关系:

$$m^+ \leq \frac{vm_1v_1}{2-q} \leq s^+, \quad (27)$$

$$m^- \leq \frac{vm_2v_2}{2-q} \leq s^-. \quad (28)$$

证明 根据KKT条件,当 $\rho > 0$ 时,有

$$\sum_{j=m_1+1}^N \alpha_j = v/(2-q);$$

当 $\xi_i > 0$ 时, $\sum_{i=1}^{m_1} \alpha_i = \frac{1}{v_1m_1}$ 对于所有正间隔误差成立.则下式成立:

$$\sum_{i=1}^{m_1} \alpha_i = v/(2-q) \geq \frac{m^+}{v_1m_1}. \quad (29)$$

由式(7)知,每个正支持向量对 $\sum_{i=1}^{m_1} \alpha_i$ 至多贡献 $1/(v_1m_1)$,故有

$$v/(2-q) = \sum_{i=1}^{m_1} \alpha_i \leq \frac{s^+}{v_1m_1}. \quad (30)$$

综合式(29)和(30),可证得式(27);同理,可证得式(28). □

定理2对于本文方法实验参数的优化选取具有指导意义.

4 实验结果及分析

通过与其他几种核学习方法进行比较来评价 ρ -MKLM算法的性能.对于二类模式分类问题,主要比较C-SVC, v -SVC与 ρ -MKLM算法的分类精度;对于一类模式分类(新奇检测)问题,重点比较OCSVM, SSLM和 ρ -MKLM算法的精度.

为了评价 ρ -MKLM在二类模式分类问题中的性能,本文采用1个人工数据集(香蕉数据集)和8个UCI实际数据集作为测试数据^[9].正类样本对应于UCI数据集中健康、正常或良性情况;负类样本对应于UCI数据集中疾病、异常或恶性情况.另外,为了进一步测试 ρ -MKLM在新奇检测问题中的性能,实验中还采用2个用于机器故障检测的数据集(Ball-bearing, Delft-pump)和1个心脏病数据集(Spectf)作为评测数据,按照文献[9]中实验设置,对这两个数据集分别随机抽取70%正类样本和一小部分负类样本,以使95%训练样本属于正类,仅有5%的训练样本属于负类.实验中所采用的数据集信息如表1所示.

表1 实验中所采用的数据集

Dataset	#pos	#neg	m_1	m_2	d
Banana	51	49	51	49	2
Wine	59	119	59	119	13
iris	50	100	50	100	4
Breast	458	241	458	241	9
Hepatitis	123	32	123	32	19
Diabetes	268	500	268	500	8
Liver	200	145	200	145	6
Heart	120	150	120	150	13
Biomed	127	67	127	67	5
Ball-Bearing	913	3237	640	40	32
Delft-Pump	376	1124	376	100	64
Spectf	95	254	95	10	44

4.1 实验参数设置

本文仿真实验环境为Matlab2009b,所有实验样本首先归一化为 $[-1, +1]$.实验中所有参数的协调通过网络搜索策略来选取^[13].实验算法均采用高斯核函数(RBF)

$$K(x, y) = \exp\left(-\frac{1}{\gamma}\|x-y\|^2\right).$$

其中: γ 值在网格 $\{\sigma^2/8, \sigma^2/4, \sigma^2/2, \sigma^2, \sigma^2*2, \sigma^2*4\}$ 中搜索选取, σ 为训练样本平均范数的平方根.对于C-SVC,惩罚参数 C 在 $\{0.01, 0.05, 0.1, 0.5, 1, 5, 10\}$ 中搜索选取;对于 v -SVC,参数 v 在 $\{0.01k, 0.1k\}$ 中搜索选取, k 为1~9之间的整数;对于OCSVM,参数 v 在 $\{0.01, 0.02, 0.04, 0.06, 0.08, 0.1, 0.2, 0.4, 0.6, 0.8, 1\}$ 中搜索选取;对于SSLM,参数 v 在 $\{10, 30, 50, 70, 90\}$ 中搜索选取, v_1 和 v_2 从 $\{0.001, 0.01\}$ 中搜索选取;对于 ρ -MKLM,根据参数定理,参数 v 在网格 $\{1, 5, 10, 15,$

20, 25, 30, 35, 40, 45, 50, 60, 80} 中搜索选取, v_1 和 v_2 从 {0.001, 0.01, 0.1} 中搜索选取.

人工生成 20 个数据集进行实验, 以评价磁场强度 q 对分类精度的影响. 磁场强度 q 对分类精度影响的分布曲线如图 3 所示. 不难看出, 优化的 q 值分布网格为 {0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8}. 一般情况下, 当 $q = 1$ 时, 分类精度能达到次优值, 故本实验对 q 值的选取策略为: 首先在 $q = 1$ 的情况下进行第 1 次测试; 然后在该值左右两端分别进行迭代搜寻; 直至寻到最优值.

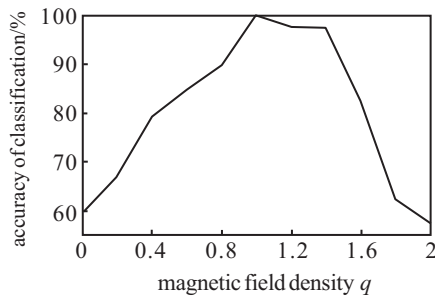


图 3 磁场强度 q 对分类精度影响

4.2 实验结论

4.2.1 人工数据

首先通过一个二维香蕉型人造数据集来比较 ρ -

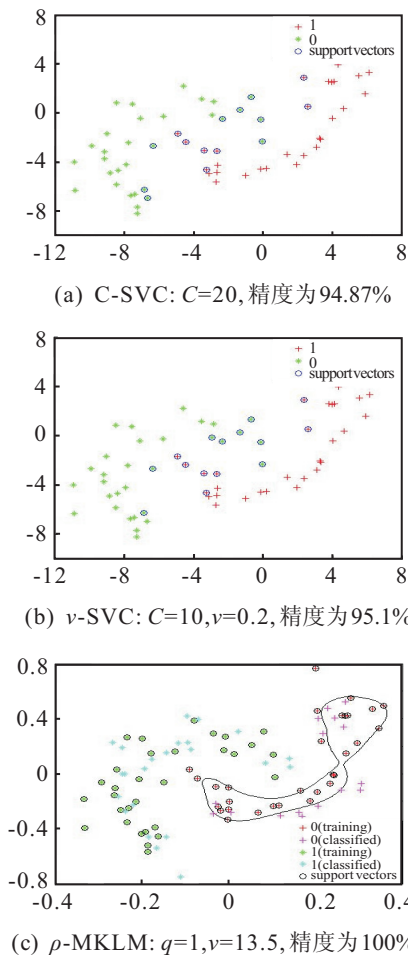


图 4 人造香蕉型数据集分类结果

MKLM方法与 C-SVC 及 v -SVC 方法的模式分类性能. 人造数据集样本数 $N = 100$, 正类数 #pos=46, 负类数 #neg=54. C-SVC, v -SVC 和 ρ -MKLM 三种分类方法的性能分别如图 4(a)~图 4(c) 所示.

4.2.2 实际数据集

通过执行 5 重交叉验证来搜索优化参数值, 并采用 g -means 度量来评价算法性能. 所有实验独立执行 10 次, 实验结果取平均值. 采用几何平均度量方法评价算法性能

$$g = \sqrt{a^+ \cdot a^-}, \quad (31)$$

其中 a^+ 和 a^- 分别为正类和负类分类精度^[9]. 该方法因同时考虑了正类与负类的分类效果而被广泛地用于处理不平衡数据集问题. 实验结果列于表 2 和表 3.

表 2 二类模式分类结果(平均结果与标准差)

Dataset	C-SVC	v -SVC	ρ -MKLM(参数 q 值)
Wine	86.51±3.67%	88.76±5.34%	98.6±4.5% (1)
iris	98.2±1.13%	98.3±0.31%	98.3±0.21% (1)
Breast	96.94±1.04%	96.5±2.12%	97.2±1.02% (1)
Hepatitis	79.48±7.13%	80.77±5.36%	88.5±3.5% (0.5)
Diabetes	65.67±2.87%	74.6±1.24%	74.6±2.61% (1)
Liver	44±14.64%	31±16.76%	71±7.18% (1.1)
Heart	85.3±1.58%	86.9±0.45%	883±5.17% (0.4)
Biomed	83.65±2.78%	82.64±4.12%	82.7±5.57% (1)

表 3 一类模式分类结果(平均结果与标准差)

Dataset	OCSVM	SSLML	ρ -MKLM(参数 q 值)
Ball-Bearing	82.3±5.53%	96.1±0.85%	96.3±0.56% (1.5)
Delft-Pump	85.11±13.12%	93.33±4.26%	95±3.61% (1)
Spectf	69.37±13.13%	72.75±4.76%	77.8±3.49% (0.6)

从表 2, 表 3 可以看出, 较其他几种方法(C-SVC, v -SVC, OCSVM, SSLML), ρ -MKLM 方法无论在二类分类还是一类分类(ND 问题)上均能取得较好的性能. 对于二类模式分类问题, 针对某些数据集(如 Wine, Liver 和 Hepatitis), ρ -MKLM 相对于经典的 C-SVC 和 v -SVC 具有明显的性能优势(如表 2 所示); 对于一类分类(ND 问题), ρ -MKLM 也能取得与 SSLML 方法相近的性能, 甚至在某些数据集(如 Delft-Pump, Spectf)上的实验结果明显优于 OCSVM 和 SSLML 方法(如表 3 所示). 由此说明, 在核诱导的高维特征空间构建一个具有磁场效应的 ρ 间隔超平面, 通过调节一个磁场强度 q , ρ -MKLM 不但能实现二类模式的分割, 还能解决一类模式分类(新奇检测)问题.

5 结论

受磁场效应理论的启发, 本文提出了一种具有磁场效应的大间隔核学习机 ρ -MKLM, 其融合了二类模式分类和新奇检测算法的思想, 使得其同时具备二类模式分类和一类模式分类(新奇检测)的优势.

在 Mercer 核诱导的特征空间, ρ -MKLM 构建一个 ρ 间隔超平面, 通过磁场效应使得一类(正类)因吸引而靠近超平面, 另一类则因排斥而尽量远离该超平面, 从而实现了不同模式间的最大可能分割; 同时, 通过引入一个可调的磁场强度 q , 进一步提升了 ρ -MKLM 学习机的分类能力. 人工数据和实际数据的实验结果均表明, ρ -MKLM 具有比其他相关方法更明显或相近的性能优势.

参考文献(References)

- [1] Koby Crammer, Mehryar Mohri, Fernando Pereira. Gaussian margin machines[C]. Artificial Intelligence and Statistics. Clearwater Beach, 2009: 105-112.
- [2] Burges C J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 955-974.
- [3] Vapnik V. The nature of statistical learning theory[M]. New York: Springer Verlag, 1995.
- [4] Vapnik V, Chapelle O. Bounds on error expectation for support vector machines[J]. Neural Computation, 2000, 12(9): 2013-2036.
- [5] Schölkopf B, Smola A J, Williamson R, et al. New support vector algorithms[J]. Neural Computation, 2000, 12(5): 1207-1245.
- [6] Schölkopf B, Smola A J. Learning with kernels[M]. Cambridge: MIT Press, 2002.
- [7] Tax D M J, Duin R P W. Support vector data description[J]. Machine Learning, 2004, 54(1): 45-66.
- [8] Chung Fulai, Deng Zhaohong, Wang Shitong. From minimum enclosing ball to fast fuzzy inference system training on large datasets[J]. IEEE Trans on Fuzzy System, 2009, 17(1): 173-184.
- [9] Wu Mingrui, Ye Jieping. A small sphere and large margin approach for novelty detection using training data with outliers[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31(11): 2088-2092.
- [10] Shivaswamy P K, Jebara T. Maximum relative margin and data-dependent regularization[J]. J of Machine Learning Research, 2010, 11(2): 747-788.
- [11] 文传军, 詹永照, 陈长军. 最大间隔最小体积球形支持向量机[J]. 控制与决策, 2010, 25(1): 79-83.
(Wen C J, Zhan Y Z, Chen C J. Maximal-margin minimal-volume hypersphere support vector machine[J]. Control and Decision, 2010, 25(1): 79-83.)
- [12] 邓乃扬, 田英杰. 支持向量机——理论、算法与拓展[M]. 北京: 科学出版社, 2009: 5-19, 81-132, 170-180.
(Deng N Y, Tian Y J. Support vector machine — Theory, algorithm and extension[M]. Beijing: Science Press, 2009: 5-19, 81-132, 170-180.)
- [13] Müller K R, Mika S, Ratsch G, et al. An introduction to kernel-based learning algorithms[J]. IEEE Trans on Neural Networks, 2001, 12(1): 181-201.

(上接第622页)

- [2] Vasilopoulos A V, Stamboullis A P. Modification of control chart limits in the presence of data correlation[J]. Quality Technology, 1978, 10(1): 20-30.
- [3] Holger Kramer, Wolfgang Schmid. Control charts for time series[J]. Nonlinear Analysis, Theory, Methods and Applications, 1997, 30(7): 4007-4016.
- [4] Marion R, Reynold J R, Lu Chao-wen. Control charts for monitoring process with autocorrelated data[J]. Nonlinear Analysis, Theory, Methods and Applications, 1997, 30(7): 4059-4067.
- [5] Schmid W. On the run length of a Shewhart chart for correlated data[J]. Statistical Papers, 1995, 36(5): 111-130.
- [6] Holger Kramer, Wolfgang Schmid. The influence of parameter estimation on the ARL of Shewhart type charts for time series[J]. Statistical Papers, 2000, 41(1): 173-196.
- [7] Yung-ming C, Tung-lung W. On average run length of control charts for autocorrelated processes[J/O]. Methodology and Computing in Applied Probability, 2009-11-27. DOI 10.1007/s11009-009-9160-4.
- [8] Fu J C, Spiring F A, Xie H H. On the average run length of quality control scheme using a Markov chain approach[J]. Statistics and Probability Letters, 2002, 56(2): 369-380.
- [9] Fu J C, Galit S, Chang Y M. A unified Markov chain approach for computing the run length distribution in control charts with simple or compound rules[J]. Statistics and Probability Letters, 2003, 65(2): 457-466.
- [10] Brook D, Evans D A. An approach to the probability distribution of cusum run length[J]. Biometrika, 1972, 59(2): 539-549.