

ϵ 不敏感损失函数支持向量机分类性能研究

杨俊燕, 张优云, 朱永生

(西安交通大学现代设计和转子轴承系统教育部重点实验室, 710049, 西安)

摘要: 将原先用于支持向量回归的 ϵ 不敏感损失函数引入到支持向量分类中, 提出 ϵ 不敏感损失函数支持向量分类算法(ϵ -SVC). 同标准支持向量分类方法(C-SVC)和最小二乘支持向量分类方法(LS-SVC)相比较, 试验结果表明: 当赋予参数 ϵ 一个足够大的接近于 1 的值时, ϵ -SVC 的分类正确率略低于 C-SVC 和 LS-SVC, 但是 ϵ -SVC 的训练、测试和参数选择的速度要高于 C-SVC 和 LS-SVC. 特别是对于大规模数据集, 这种优势将更加明显. 另外, 通过精确选择参数 ϵ 的值, ϵ -SVC 能够获得比 C-SVC 和 LS-SVC 更高的分类正确率, 但是训练、测试和参数选择的速度却随着 ϵ 的减小而降低.

关键词: ϵ 不敏感损失函数; 支持向量分类; 模式分类; 支持向量回归

中图分类号: TP183 **文献标识码:** A **文章编号:** 0253-987X(2007)11-1315-06

Classification Performance of Support Vector Machine with ϵ -Insensitive Loss Function

Yang Junyan, Zhang Youyun, Zhu Yongsheng

(Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The ϵ -insensitive loss function generally employed in support vector regression is introduced into support vector classification, and the support vector classification with ϵ -insensitive loss function (ϵ -SVC) is proposed. Compared with the standard support vector classification method (C-SVC) and the least square support vector classification method (LS-SVC), the experimental result indicates that the classification accuracy ratio of ϵ -SVC is slightly lower than that of C-SVC and LS-SVC when ϵ sufficiently approaches to 1, but the training, testing and parameter selecting rates of ϵ -SVC are superior to that of C-SVC and LS-SVC, especially for large scale problem. Through accurate search of the parameter ϵ , the ϵ -SVC is endowed with higher classification accuracy than C-SVC and LS-SVC, however, the training, testing and parameter selecting rates decrease with smaller ϵ .

Keywords: ϵ -insensitive loss function; support vector classification; pattern classification; support vector regression

支持向量机(SVM)是一种通用的学习算法, 它实现了由统计学习理论所提出的结构风险最小化原则^[1]. 由于其良好的推广能力, 近年来逐渐成为人工智能领域的研究热点. 在许多实际的应用中, SVM 表现出了优于学习算法的分类和预测能力^[2-3].

不同的损失函数可以构建不同的 SVM^[4]. 1992

年, Vapnik 等人提出了一种新颖的支持向量分类算法, 后人称其为标准支持向量分类算法(C-SVC). 该算法所采用的损失函数为 $L(y, f(x)) = (1 - yf(x))_+$, 其核心思想是通过引入非线性映射函数 $\varphi(x)$, 将原始模式空间映射到更高维的特征空间 Z , 在特征空间中构造最优分类超平面, 将高维空间中

的线性问题与低维空间中的非线性问题相对应,从而实现原始模式空间的分类^[4].

1999年,Suykens等人^[5]提出了一种改进版本的SVM——最小二乘支持向量机(LS-SVM),该算法采用的损失函数为二次损失函数 $L(y, f(x)) = (y - f(x))^2$.同C-SVC相比,LS-SVM将不等式约束用等式约束代替,方程的解通过求解一组线性方程组得到,而不是通过求解一个二次规划问题得到.但是,从最小二乘意义上讲,LS-SVM的解损失了稀疏性.

本文将原先用于支持向量回归(SVR)的 ϵ 不敏感损失函数引入到支持向量分类中,提出 ϵ 不敏感损失函数支持向量分类算法(ϵ -SVC),研究了 ϵ -SVC算法的分类性能,探讨了不同参数对其分类性能的影响.最后,将 ϵ -SVC算法同C-SVC和LS-SVC算法的分类性能进行了综合比较.

1 ϵ 不敏感损失函数支持向量机分类算法

为将估计指示函数(模式识别问题)中得到的结果推广到对实函数的估计(回归问题),引入 ϵ 不敏感损失函数^[4]

$$L(y, f(x, \alpha)) = L(|y - f(x, \alpha)|_\epsilon) \quad (1)$$

式中

$$|y - f(x, \alpha)|_\epsilon = \begin{cases} 0 & |y - f(x, \alpha)| \leq \epsilon \\ |y - f(x, \alpha)| - \epsilon & \text{其他} \end{cases} \quad (2)$$

与模式识别问题类似,将模式空间中的输入向量 x 通过映射函数 $\varphi(x)$ 将其映射到高维特征空间 Z 中,用特征空间中的线性函数集合

$$f(x, \alpha) = (\omega \varphi(x)) + b \quad (3)$$

估计回归函数.对于给定的训练数据集 $(y_1, x_1), \dots, (y_l, x_l)$,采用 ϵ 不敏感损失函数,则其约束优化问题可表示为^[4]

$$\begin{aligned} \min \quad & \frac{1}{2} \|\omega\|^2 + C \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{s. t.} \quad & \begin{cases} y_i - (\omega \varphi(x_i)) - b \leq \epsilon + \xi_i & i = 1, \dots, l \\ (\omega \varphi(x_i)) + b - y_i \leq \epsilon + \xi_i^* & i = 1, \dots, l \\ \xi_i, \xi_i^* \geq 0 & i = 1, \dots, l \end{cases} \end{aligned} \quad (4)$$

式中: C 为惩罚因子,用于实现经验风险和置信范围的折中; ξ_i 和 ξ_i^* 为松弛因子.式(4)所确定的约束优

化问题是一个典型的凸优化问题,可通过拉格朗日乘子法来求解.

对于一个 k 类分类问题,标签集 D 可以被定义为

$$D = \{d_1, d_2, \dots, d_k\} \quad d_i \in Z, d_1 < d_2 < \dots < d_k \\ k \geq 2 \quad i = 1, 2, \dots, k \quad (5)$$

在SVR中, y_i 为实数值,而对于分类问题, y_i 则取 k 个离散的值.当把分类问题的标签值看作实数值时,就可以在分类问题的训练样本集上构建SVR模型,并且可以通过式(6)来确定一个未知样本 x 的所属分类

$$y = \begin{cases} d_1 & \text{if } f^*(x) \leq d_1 \\ d_i & \text{if } d_i \leq f^*(x) \leq (d_i + d_{i+1})/2 \\ d_{i+1} & \text{if } (d_i + d_{i+1})/2 < f^*(x) \leq d_{i+1} \\ d_k & \text{if } f^*(x) \geq d_k \end{cases} \quad (6)$$

特别是对于两类分类问题($D = \{-1, 1\}$),式(6)可以写为

$$y = \text{sgn}(f^*(x)) \quad (7)$$

2 标准数据集和试验

本文从Statlog^[6]和UCI^[7]机器学习数据库中选择了9个数据集来对 ϵ 不敏感损失函数的分类性能进行研究.在本文中,将所有的训练数据中的每个属性都归一化到-1和1之间,测试数据集根据同样的线性变换进行归一化.表1给出了9个标准数据集的属性.

对于 ϵ -SVC、C-SVC和LS-SVC分类算法,核函数都采用径向基函数核,它们都具有共同的参数 C 和 γ (RBF核参数).本文采用网格搜索方法^[8]对这3种支持向量分类算法的参数进行选择,其中 C 和 γ 的取值都为 $[2^{-15}, 2^{-14}, \dots, 2^{14}, 2^{15}]$,这样在每个数据集上都要计算 $31 \times 31 = 961$ 次参数组合.对于训练集和测试集都可用的数据集,在进行参数选择的时候采用训练集的70%进行训练,剩下的30%进行校验,接着用最好的校验正确率所对应的参数对整个训练集进行训练,并在测试集上进行测试,然后取测试集上的分类正确率 r 作为对分类算法分类性能的评价标准.对于只有训练集的数据集,在训练集上做20次10重交叉检验,取最好的平均交叉检验分类正确率 r 作为对分类算法分类性能的评价标准.本文所用的试验平台为:Opteron 144 CPU, 2 048 MB内存, GCC编译器, Redhat Linux FC4操

作系统.

表 1 标准数据集

数据集	训练数据数	测试数据数	类别数	属性数
australian	690	0	2	14
breast-cancer	683	0	2	10
astroparticle	3 089	4 000	2	4
iris	150	0	3	4
wine	178	0	3	13
DNA	2 000	1 186	3	180
satimage	4 435	2 000	6	36
letter	15 000	5 000	26	16
shuttle	43 500	14 500	7	9

另外,文中所有的程序均基于 LibSVM^[9] 和 LSSVMLab^[10].

3 ϵ 不敏感损失函数支持向量机多类分类方法

一般情况下,存在 2 种多类分类方法:一种方法是在一个优化问题中直接考虑所有类别的数据进行直接多类分类;另外一种方法是将多类分类问题分解成几个 2 类分类问题来考虑. 式(6)表明, ϵ -SVC 可以被直接用于 2 类和多类分类问题,本文称这种多类分类方法为直接多类分类法. 对于 2 类分类问题,式(7)显示只需要考虑回归函数 $f^*(x)$ 的符号,而不需要考虑到其精确的值. 换句话说,就是回归函数 $f^*(x)$ 的精度对 ϵ -SVC 的分类正确率影响不大,因此参数 ϵ 可以被赋予一个足够大的接近于 1 的值. 根据 SVR 理论,大的 ϵ 值可以有效地降低支持向量的数量,提高 SVM 的训练和测试速度. 对于多类分类问题,式(6)可以用于直接确定类别,但是必须取小的 ϵ 值来确保回归函数的精度. 对于采用多个 2 类分类器来实现多类分类的方法,存在一对一、一对多和 DAG 等几种实现方法. Hsu 在文献[11]中指出,一对一和 DAG 方法更加适合于在实践中使用. 本文从简单实现的角度出发,仅考虑一对一方法. 一对一方法需构建 $k(k-1)/2$ 个 2 类分类器,因此对于每一个 2 类分类器, ϵ 都可以赋予一个较大的接近于 1 的值,以充分地利用 ϵ -SVC 用于 2 类分类的优点.

表 2 给出了 2 种多类分类方法在 4 个标准数据

集上的分类结果. 从表中可以看出,即使 ϵ 取比较大的值,一对一方法仍然能够在 4 个标准数据集上取得较高的分类正确率. 对于 ϵ -SVC 直接多类分类方法,随着 ϵ 的减小, ϵ -SVC 直接多类分类的正确率逐渐提高,但是在 DNA、satimage 和 glass 数据集上, ϵ -SVC 直接多类分类的正确率仍然要比一对一方法的分类正确率低. 因此,在本文中采用一对一方法进行多类分类.

表 2 2 种多类分类方法的分类性能比较

数据集	ϵ	直接法		一对一法	
		(C, γ)	r	(C, γ)	r
iris	0.95	(-, -)	0.340	$(2^{12}, 2^{-9})$	0.963
	0.45	$(2^{-2}, 2^{-2})$	0.953	$(2^7, 2^{-7})$	0.980
	0.10	$(2^{11}, 2^{-6})$	0.985	$(2^{10}, 2^{-12})$	0.981
wine	0.95	(-, -)	0.304	$(2^{-5}, 2^{-2})$	0.969
	0.45	$(2^{-1}, 2^{-2})$	0.927	$(2^7, 2^{-11})$	0.989
	0.10	$(2^0, 2^{-3})$	0.991	$(2^6, 2^{-7})$	0.997
DNA	0.95	(-, -)	0.236	$(2^{-2}, 2^{-13})$	0.953
	0.45	$(2^{15}, 2^{-15})$	0.780	$(2^2, 2^{-10})$	0.958
	0.10	$(2^{10}, 2^{-13})$	0.824	$(2^3, 2^{-9})$	0.960
satimage	0.95	(-, -)	0.231	$(2^{-9}, 2^1)$	0.911
	0.45	$(2^2, 2^{-1})$	0.777	$(2^4, 2^1)$	0.920
	0.10	$(2^1, 2^0)$	0.811	$(2^2, 2^0)$	0.922

注: (-, -) 表示在所有参数对 (C, γ) 上获得了相同分类正确率.

4 参数对 ϵ -SVC 分类算法分类性能的影响

对于 SVM, 模型参数是影响其性能的主要因素. 不适当的参数会降低 SVM 的分类性能, 出现过学习或欠学习的现象. 本节主要讨论 ϵ -SVC 的 3 个参数 C 、 γ 和 ϵ 对其分类性能的影响. 另外, 本节中所有的试验都是在 letter 数据集上做的.

图 1 给出了 ϵ -SVC 在所有 (C, γ) 对上的分类正确率, 其中 ϵ 的取值分别为 0.05、0.65、0.85 和 0.95. 从图 1 中可以看出, 4 个图形中都有一个最大值平面, 并且随着 ϵ 的增大, 最大值平面的面积也在增大. 根据 SVR 理论, 小的 ϵ 会导致 SVM 出现过学习现象, 这就要求一个较大的 C 来正确地分类数据, 图 1a 正好说明了这个问题. 随着 ϵ 的增大, 过学习问题逐渐得到克服, 一个小的惩罚系数也能使 SVM 取得较好的分类正确率, 见图 1d.

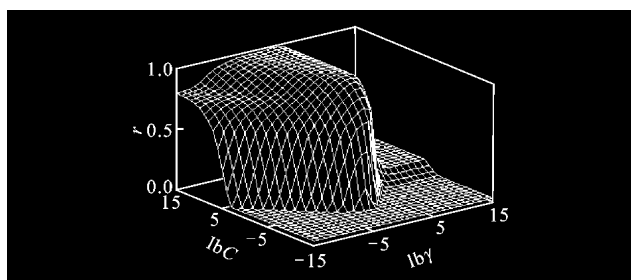
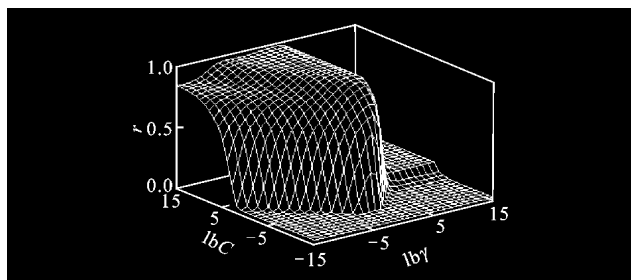
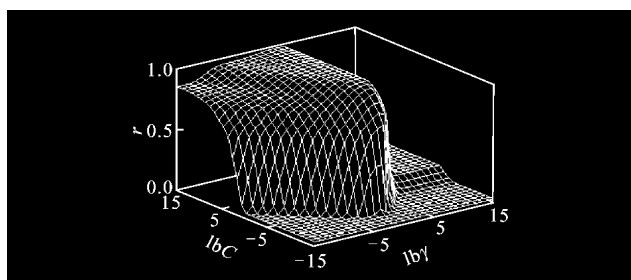
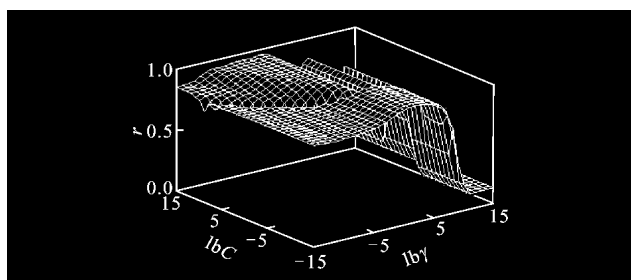
(a) $\epsilon=0.05$ (b) $\epsilon=0.65$ (c) $\epsilon=0.85$ (d) $\epsilon=0.95$ 图1 ϵ -SVC在所有 (C, γ) 对上的分类正确率

图2给出了 ϵ -SVC在不同的参数 C 和 ϵ 下的分类正确率 r 、支持向量数 N_{SV} 、训练时间 t_r 和测试速度 t_e ,其中 $C=[2^{-4}, 2^{-2}, 2^0, 2^3, 2^6, 2^8, 2^{10}, 2^{12}]$; $\epsilon=[0.01, 0.05, 0.1, 0.15, \dots, 0.9, 0.95]$;核参数 $\gamma=2^{-1}$.从图2a中可以看出,当参数 ϵ 从0.01增大到0.9时, ϵ -SVC的分类正确率随着 ϵ 的增大而增大,但是当 ϵ 大于0.9时, ϵ -SVC的分类正确率却有所下降.另外,当 C 从 2^{-4} 增大到 2^0 时, ϵ -SVC的分类正确率随着 C 的增大而增大,但是当 C 大于 2^0 时, ϵ -SVC的分类正确率对参数 C 和 ϵ 的变化不是很敏

感.图2b给出了 ϵ -SVC的支持向量数随参数 C 和 ϵ 变化的情况,可以看出 ϵ -SVC的支持向量数随着 ϵ 的增大而减少.比较图2b和图2d可以看出,这2个图形的变化趋势非常相似,说明支持向量的数目决定了 ϵ -SVC的测试速度.图2c给出了 ϵ -SVC的训练速度随 C 和 ϵ 的变化情况,可以看出随着 ϵ 的减小, ϵ -SVC的训练时间呈指数增长,特别是对于 C 大于 2^3 的情况.从上面的分析可以得出,小的 ϵ 使得 ϵ -SVC的测试和训练速度增加,同时由于出现过学习现象,需要一个大的 C 值来正确地分类数据,但是大的 C 值却使得 ϵ -SVC的训练时间急剧增加.因此,对于letter数据集,适当的 C 值将处于 2^0 和 2^3 之间,适当的 ϵ 值将处于0.8和0.85之间.

5 ϵ -SVC、C-SVC和LS-SVC分类算法比较

本节主要探讨 ϵ -SVC、C-SVC和LS-SVC分类算法在9个标准数据集上的分类正确率、训练速度、测试速度和参数选择速度,其中 ϵ -SVC算法的 ϵ 值设定为0.95.

表3给出了 ϵ -SVC、C-SVC和LS-SVC分类算法在9个标准数据集上的分类正确率,最高的分类正确率在表中以黑体显示.试验结果显示,在9个标准数据集上, ϵ -SVC的分类正确率要略低于C-SVC和LS-SVC,且C-SVC在大多数数据集上都获得了最高的分类正确率.

表4给出了 ϵ -SVC、C-SVC和LS-SVC分类算法在9个标准数据集上的训练时间、测试时间和参数选择时间 t_{ps} .需要注意的是,表中仅仅给出了最优模型的训练和测试时间.试验结果显示:在大多数标准数据集上, ϵ -SVC获得了最快的训练速度、测试速度和参数选择速度;LS-SVC的训练速度、测试速度和参数选择速度要明显低于 ϵ -SVC和C-SVC;同C-SVC相比, ϵ -SVC的训练速度提高了16.67%~86.93%,测试速度提高了18.25%~68.93%,参数选择速度提高了12.49%~87.27%;特别是对于大规模数据集DNA、satimage、letter和shuttle, ϵ -SVC的速度确实具有很大的优势.

为了进一步考察 ϵ -SVC在不同参数 ϵ 下的分类性能,采用网格搜索法对3个参数 C 、 γ 和 ϵ 同时进行搜索,选取最优的参数作为对 ϵ -SVC分类性能的评价,其中 $\epsilon=[0.05, 0.15, \dots, 0.85, 0.95]$, C 和 γ 的取值范围和前面相同.表5给出了考虑 ϵ 时, ϵ -

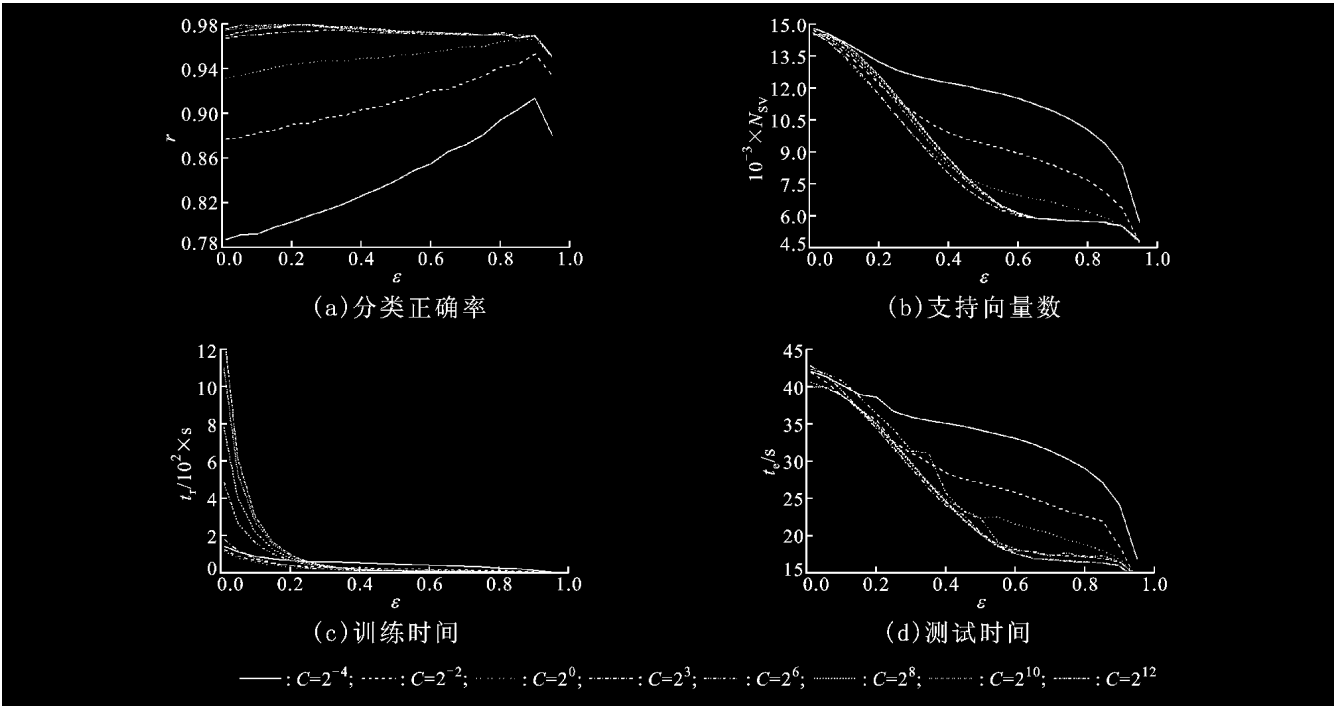


图 2 ϵ -SVC 在不同 C 和 ϵ 下的分类正确率、支持向量数、训练时间和测试速度

表 3 3 种分类算法在 9 个标准数据集上的分类正确率($\epsilon=0.95$)

数据集	ϵ -SVC		C-SVC		LS-SVC	
	(C, γ)	r	(C, γ)	r	(C, γ)	r
australian	$(2^{-6}, 2^{-4})$	0.859	$(2^{-5}, 2^{-2})$	0.861	$(2^5, 2^{12})$	0.868
breast-cancer	$(2^{-7}, 2^{-8})$	0.972	$(2^{-3}, 2^{-3})$	0.971	$(2^3, 2^5)$	0.971
astroparticle	$(2^{-5}, 2^2)$	0.970	$(2^1, 2^2)$	0.972	$(2^2, 2^{-1})$	0.968
iris	$(2^{12}, 2^{-9})$	0.963	$(2^3, 2^{-4})$	0.973	$(2^{11}, 2^1)$	0.973
wine	$(2^{-5}, 2^{-2})$	0.969	$(2^0, 2^{-2})$	0.990	$(2^{-2}, 2^1)$	0.991
DNA	$(2^{-2}, 2^{-13})$	0.953	$(2^3, 2^{-10})$	0.956	$(2^0, 2^7)$	0.954
satimage	$(2^{-9}, 2^1)$	0.911	$(2^5, 2^1)$	0.921	$(2^3, 2^{-1})$	0.916
letter	$(2^{-13}, 2^2)$	0.956	$(2^1, 2^2)$	0.981	$(2^{11}, 2^{-2})$	0.969
shuttle	$(2^{-13}, 2^{12})$	0.999	$(2^{15}, 2^1)$	0.999		

表 4 3 种分类算法在 9 个标准数据集上的训练时间、测试时间和参数选择时间($\epsilon=0.95$)

数据集	ϵ -SVC			C-SVC			LS-SVC		
	t_r	t_e	t_{ps}	t_r	t_e	t_{ps}	t_r	t_e	t_{ps}
australian	5.0		14 047.4	12.1		12 800.7	11.57		12 707.5
breast-cancer	1.04		4 098.39	2.34		7 043.19	13.26		13 818.7
astroparticle	0.15	0.13	1 306.96	0.18	0.16	1 984.07	4.25	1.24	8 140.87
iris	0.13		189.16	0.12		442.05	0.6		548.15
wine	0.13		418.05	0.32		864.34	1.07		960.60
DNA	1.21	0.94	8 749.83	2.17	1.52	15 563.8	2.55	3.92	12 404.8
satimage	0.7	0.64	8 580.75	3.25	2.06	16 683.9	11.98	11.42	44 477.1
letter	3.57	14.1	103 448	27.31	22.34	185 834.0	93.85	293.22	1 219 067
shuttle	3.37	0.57	32 316.8	7.36	0.63	253 935.1			

SVC 在 7 个标准数据集上的分类性能. 从表中可以看出, ϵ -SVC 能够取得比其他几种分类算法更高的分类正确率, 但是训练速度和测试速度却有所下降. 因为同时在 3 个参数上进行搜索, 参数选择的速度必然要比 2 个参数的选择慢很多.

表 5 考虑 ϵ 时 ϵ -SVC 在 7 个标准数据集上的分类性能

数据集	ϵ	(C, γ)	r	t_r/s	t_e/s
australian	0.25	$(2^{-3}, 2^{-8})$	0.866	7.83	
breast-cancer	0.45	$(2^2, 2^{-6})$	0.972	1.00	
astroparticle	0.75	$(2^{-1}, 2^2)$	0.972	0.21	0.15
iris	0.35	$(2^{14}, 2^{-6})$	0.985	1.69	
wine	0.05	$(2^2, 2^{-4})$	0.998	0.68	
DNA	0.85	$(2^{-1}, 2^{-9})$	0.960	1.93	1.45
satimage	0.65	$(2^{-1}, 2^1)$	0.922	2.82	1.55

6 结 论

分类问题可以被看作回归问题的一个特例. 不同的损失函数构成了不同的 SVM. 本文将原先用于 SVR 的损失函数—— ϵ 不敏感损失函数用于分类问题, 提出了 ϵ 不敏感损失函数支持向量分类算法. 对于 2 类分类问题, 由于只关心回归函数的符号, 而不考虑其回归精度, 因此参数 ϵ 可以被赋予一个足够大的接近于 1 的值. 对于多类分类问题, 本文讨论了 2 种方法: 直接法和一对一方法. 试验结果显示, 直接法的分类正确率要明显低于一对一方法. 接着讨论了参数 C 、 γ 和 ϵ 对 ϵ -SVC 分类性能的影响. 在 letter 数据集上的试验显示: 小的 ϵ 会导致 SVM 出现过学习现象, 这就要求一个较大的惩罚系数 C 来正确地分类数据; 随着参数 ϵ 的增大, 过学习问题逐渐得到克服, 一个小的惩罚系数 C 也能使 SVM 取得较好的分类正确率; ϵ -SVC 的支持向量数随着 ϵ 的增大而减少; 随着 ϵ 的减小, ϵ -SVC 的训练时间呈指数增长, 特别是当 C 较大时. 最后, 在 9 个标准数据集上对 ϵ -SVC、C-SVC 和 LS-SVC 3 种分类算法的分类性能进行了对比, 试验结果显示: 当 ϵ 足够大时, ϵ -SVC 的分类正确率要略低于 C-SVC 和 LS-SVC, 但是训练速度、测试速度和参数选择速度却有较大提高, 特别是在大规模数据集上的优势更加明显; 通过适当选择 ϵ , ϵ -SVC 能够取得比 C-SVC 和

LS-SVC 略高的分类正确率, 但是训练速度、测试速度和参数选择速度却明显下降.

参考文献:

- [1] Vapnik V N. 统计学习理论[M]. 许建华, 张学工, 译. 北京: 电子工业出版社, 2004.
- [2] 朱永生. 支持向量机及其在机械故障模式识别中的应用[D]. 西安: 西安交通大学机械工程学院, 2003.
- [3] Yang J Y, Zhang Y Y. Application research of support vector machines in condition trend prediction of mechanical equipment [M]// Lecture Notes in Computer Science: 3498. Berlin: Springer, 2005: 857-864.
- [4] Vapnik V N. The nature of statistical learning theory [M]. New York: Springer, 1995.
- [5] Suykens J A K, Brabanter J D, Lukas L, et al. Weighted least squares support vector machines: robustness and sparse approximation [J]. Neurocomputing, 2002, 48(1): 85-105.
- [6] Michie D, Spiegelhalter D J, Taylor C C. Machine learning, neural and statistical classification [EB/OL]. Englewood Cliffs, USA: Prentice Hall, 1994 [2007-02-15]. <http://www.ncc.up.pt/liacc/ML/statlog/datasets.html>.
- [7] Blake C L, Merz C J. UCI repository of machine learning databases [EB/OL]. Irvine, USA: Univ. California, Dept. Inform. Comput. Sci., 1998 [2007-02-15]. <http://www.ics.uci.edu/?mlearn/MLRepository.html>.
- [8] Hsu C W, Chang C C, Lin C J. A practical guide to support vector classification [EB/OL]. Taiwan: Taiwan Univ., Dept. Comput. Sci. Inform. Eng. [2007-03-10]. <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.
- [9] Chang C C, Lin C J. LIBSVM: a library for support vector machines [EB/OL]. [2007-03-10]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [10] Suykens J A K, Gestel T V, Brabanter J D, et al. Least squares support vector machines [EB/OL]. Singapore: World Scientific, 2002 [2007-03-10]. <http://www.esat.kuleuven.ac.be/sista/lssvmlab/>.
- [11] Hsu C W, Lin C J. A comparison of methods for multiclass support vector machines [J]. IEEE Transactions on Neural Networks, 2002, 13(2): 415-425.

(编辑 葛赵青 荆树蓉)