基于预分类的FSVM

申丰山
(郑州大学信息工程学院 郑州 450001)

摘 要 模糊支持向量机(fuzzy support vector machine,FSVM)通过为每个样例设置模糊化训练参数,达到抑制离群点及噪声数据对分类器不利影响的目的。提出了基于预分类的FSVM,每个样例的模糊权重通过关联于该样例的预分类面来确定。该方法不仅考虑了各个样例在未来分类中的作用效果,还考虑了分类器对离群点及噪声数据的敏感性。这样确定的模糊权重能使SVM根据离群点及噪声数据的影响情况决定抑制强度,减少或避免无视数据具体特征的盲目抑制。在IDA、UCI等标准数据集上的实验验证了所提方法的合理性和有效性。

关键词 模糊支持向量机(FSVM),预分类面,模糊权重,敏感性

中图法分类号 TP391.4 **文献标识码** A

FSVM Based on Pre-classification

SHEN Feng-shan
(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract A fuzzy support vector machine(FSVM) reduces the bad effect of outliers or noises on the classifier by using different training parameters for different training examples. This paper proposed an FSVM based on pre-classification, in which fuzzy weight for each training example is computed according to a classification hyperplane created for the training example. The method considers not only the effect of each individual training example, but also the sensitivity of the classifier to the outliers or noises. SVM with such fuzzy weights is able to make proper suppression for the training data. Experimental results on the standard training datasets from IDA and UCI repositories show the reasonableness and effectiveness of the proposed algorithm.

Keywords Fuzzy support vector machine(FSVM), Pre-classification hyperplane, Fuzzy weight, Sensitivity

1 引言

支持向量机(SVM)[1]作为一种优秀的分类工具受到了人们的广泛关注[2,3]。但是SVM的训练过程对训练集中远离本类中心的噪声及离群点很敏感,被错分的离群点会影响分类面的产生,进而影响分类器的泛化性能。对此,人们提出模糊支持向量机(Fuzzy support vector machine,FSVM)[4-8]方法以削弱噪声及离群点的不利影响。文献[4,5]最早提出了FSVM基本模型,文献[6]提出了双边带权FSVM模型(B-FSVM),文献[7]提出了软SVM模型(S_SVM),文献[8,9]提出了用于非平衡类学习的模糊支持向量机(FSVMs-CIL)。尽管这些FSVM的构成形式存在一定差异,但是最终影响其性能的关键因素是模糊权重函数。模糊权重函数是估计样本作用及其属于噪声/离群点的可能性大小,进而决定各个样本点训练参数值大小的一种计算方法。目前可供选择的模糊权重函数种类非常有限,基本上局限于基于样本点到样本中心距离的计算方法及其变种[10-12]。而这种模糊权重仅与其中一类样本有关,不能反映两类样本之间的关系,用于训练分类器时,容易导致过抑制,不仅不能改进反而会降低分类器性能。本文提出了基于预分类的FSVM。该方法通过关联于各个样例的预分类面评估相应样例在未来分类中的作用效果,同时也考虑分类器对离群点及噪声数据的敏感性。综合两种因素产生的模糊权重能够使FSVM根据训练集中样例个体的情况和两类样例集之间的分离关系确定抑制强度,减少或避免盲目抑制,稳定分类器工作性能。实验验证了所提方法的合理性和有效性。

2 SVM和FSVM

由Vapnik[1]等提出的SVM自两类数据的分类问题发展而来,可以扩展到多类问题。SVM在经由核函数$K(x_i,x_j)=\Phi(x_i)\cdot\Phi(x_j)$映射的特征空间中工作,其中$\Phi:R^d\to F$是从$d$维输入空间到一个可能的高维特征空间的一个映射,"$\cdot$"表示特征空间中的内积运算。

对于一个数据集为$\{(x_i,y_i)\}_{i=1}^{l}$的二分类问题,$x_i\in R^d$为代表第i个训练数据的d维特征向量,$y_i\in\{-1,1\}$是x_i的类标签。用于模式分类的SVM试图在特征空间中找到一个具有最大间隔的超平面,即最优分类面$H_{w,b}:w\cdot\Phi(x)+b=0$,用以分开两类训练数据。该分类面可以通过解决如下二次规划问题获得:

$$\min\frac{1}{2}\|w\|^2+C\sum_{i=1}^{l}\xi_i \quad (1)$$
$$\text{s.t. } y_i(w\cdot\Phi(x_i)+b)\geqslant 1-\xi_i,\ \xi_i\geqslant 0,\ \forall i=1,\cdots,l$$

到稿日期:2012-11-10 返修日期:2013-04-03 本文受国家自然科学基金(61070137, 60933009)资助。

申丰山(1970-),男,博士,副教授,主要研究方向为模式识别与机器学习、图像处理、软件工程,E-mail:iefsshen@zzu.edu.cn。

式中，C 是控制间隔宽度与训练误差的折中参数，称为惩罚因子。较大的 C 降低 SVM 的训练误差并且形成较窄的间隔。ξ_i 为松弛变量。式(1)的 Lagrangian 对偶问题是：

$$\max W(\alpha)=\sum_{i=1}^{l}\alpha_i-\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i y_i\alpha_j y_j\Phi(x_i)\cdot\Phi(x_j)$$
$$\text{s.t.}\ \sum_{i=1}^{l}\alpha_i y_i=0,0\leqslant\alpha_i\leqslant C\ \text{for}\ i=1,\cdots,l \tag{2}$$

根据最优性条件可以得到如下形式的决策函数：

$$f(x)=w\cdot\Phi(x)+b=\sum_{i=1}^{l}\alpha_i y_i K(x,x_i)+b \tag{3}$$

式中，α_i 是二次规划问题(2)的解，x 是要分类的输入向量，b 为偏置，可以根据 Karush-Kuhn-Tucker(KKT)条件得到。如果 $f(x)\geqslant 0$，则 x 分为正类，否则分为负类。系数 $\alpha_i>0$ 对应的输入向量 x_i 称为支持向量。实际上，w 仅仅依赖于支持向量。

FSVM 通过模糊化配置训练参数达到降低离群点及噪声影响的目的。在 FSVM 中，训练集中每个样例都关联一个表征其重要性的模糊权重 $s_i(0<s_i\leqslant 1)$。假定 $\{(x_i,y_i,s_i)\}_{i=1}^{l}$ 是两类分类问题的一个训练集，其中 $x_i\in\mathcal{R}^d$，d 为训练数据的维数，$y_i\in\{-1,1\}$ 是 x_i 的类别标签，s_i 是 x_i 的模糊权重。则 FSVM 算法被形式化为寻找如下优化问题的解：

$$\min\frac{1}{2}\|w\|^2+C\sum_{i=1}^{l}s_i\xi_i$$
$$\text{subject to}\ y_i(w\cdot\Phi(x_i)+b)\geqslant 1-\xi_i,\xi_i\geqslant 0,i=1,\cdots,l \tag{4}$$

较小的 s_i 将降低 ξ_i 的影响，从而使得相应的训练数据 x_i 被看得不太重要，由 x_i 产生的错误也会比较小。式(4)的对偶形式为：

$$\max W(\alpha)=\sum_{i=1}^{l}\alpha_i-\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i y_i\alpha_j y_j K(x_i,x_j)$$
$$\text{subject to}\ \sum_{i=1}^{l}\alpha_i y_i=0,0\leqslant\alpha_i\leqslant s_i C,i=1,\cdots,l \tag{5}$$

如果使 $s_i=1$，则 FSVM 便成为 SVM。

目前较为常见的中心距离型模糊权重计算方法如下：

$$s_i=\begin{cases}1-(\|x_+-x_i\|/(r_++\delta)), & \text{若}\ x_i\in\text{类 1}\\ 1-(\|x_--x_i\|/(r_-+\delta)), & \text{若}\ x_i\in\text{类 2}\end{cases} \tag{6}$$

其中，s_i 是数据 x_i 的模糊权重，x_+ 和 x_- 分别是类 1 和类 2 的均值，$\delta>0$ 用于避免 $s_i=0$。r_+、r_- 分别是类 1 和类 2 的半径：

$$r_+=\max_{\{x_i:x_i\in\text{类1}\}}\|x_+-x_i\| \tag{7}$$

$$r_-=\max_{\{x_i:x_i\in\text{类2}\}}\|x_--x_i\| \tag{8}$$

3 基于预分类的 FSVM

传统 FSVM 通常无法估计训练集的可分性特征和各个样例对分类器的作用效果，因此其对分类结果的控制是盲目的。对此，本文提出的基于预分类的 FSVM 首先要估计各个样例在未来分类中的作用效果和分类器对离群点及噪声数据的敏感性，在此基础上确定相应模糊权重的大小。样例作用效果和分类器对离群点及噪声数据敏感性的估计都是通过为各个样例分别建立的预分类面来进行的。预分类面是点超平面的一种。

定义 1(点超平面、点分类面、点误分率、最优点分类面)

令 $X=\{x_i\}_{i=1}^{l}$ 为输入向量集，$y_i\in\{-1,1\}(i=1,\cdots,l)$ 为 X 中输入向量对应的输出。在由核函数 $K(x_i,x_j)=\Phi(x_i)\cdot\Phi(x_j)$ 定义的特征空间中，$f_i(x_k)=w_i\cdot\Phi(x_k)+b_i$，其中 $x_k\in X$ 且 $f_i(x_i)=0$ 定义了通过 $\Phi(x_i)(x_i\in X)$ 的超平面，称为点超平面。显然通过一点的超平面可以有无穷多个。点超平面用于分类时称为点分类面。点分类面对 X 中两类数据的误分率称为 $\Phi(x_i)$(或 x_i)的点误分率，其计算公式为：

$$e_i=\frac{1}{2l}\left(l+\sum_{k=1}^{l}sign(y_k f_i(x_k))\right) \tag{9}$$

其中，$sign(z)=\begin{cases}1, & \text{if}\ z\geqslant 0\\ -1, & \text{if}\ z<0\end{cases}$。

一个点的各个点分类面中，具有最小点误分率的点分类面称为该点的最优点分类面。

最优点分类面由于紧贴着其中一类样本，因此不具有 SVM 所拥有的最大间隔属性，用于分类时不能保证优良的泛化性能。但是各个点的最优点分类面对于样本集的点误分率与样例所处位置关系十分密切。具有较小点误分率的样例位于两类数据之间的边界区域，其所属类别具有相对较大的不确定性，其中如果包含离群点及噪声数据，则可能影响最优分类面的生成和 SVM 的泛化性能。因此对于这部分边界样例需要进行适度抑制，在训练 SVM 时，赋予它们相对较小的模糊权重，而对远离边界、类别较为确定的样例赋予相对较大的模糊权重或者直接赋为 1。

点误分率起到了模糊权重的作用。点误分率是通过设计分类面即最优点分类面来尝试分类而得到的。因此这种计算模糊权重的方法称为预分类方法。

为了避免过抑制，进一步考虑分类器对离群点及噪声数据的敏感性。若分类器不容易受到离群点及噪声数据的影响，则样例免于抑制或减轻抑制，否则适度加强抑制。分类器对离群点及噪声数据的敏感性采用 $(1-\min_{i=1,\cdots,l}e_i^*)$ 来度量，e_i^* 表示通过 $\Phi(x_i)(x_i\in X)$ 的最优点分类面所具有的点误分率，即最小点误分率。这样，实际上以最终分类器可能达到的最高分类精度来估计数据的可分性，或者说估计边缘样例对分类器的干扰性。

综合考虑单个样例作用效果和分类器敏感性两种因素后，样例 $x_k\in X$ 的模糊权重根据下式来计算：

$$s_k=\min(e_k^*-\min_{i=1,\cdots,l}e_i^*+1-\min_{i=1,\cdots,l}e_i^*,1) \tag{10}$$

这种模糊权重具有如下特点：

当数据集完全可分时，标准 SVM 往往就能产生理想的分类结果，基于预分类的 FSVM 将在很大程度上转化为标准 SVM，因为此时训练集中每个样例都将获得接近于 1 或等于 1 的模糊权重。

当数据集不是完全可分时，可能的离群点及噪声数据的模糊权重会适度降低，样例作用会受到适当的抑制，而其它样例仍然获得接近于 1 或等于 1 的模糊权重。这样的权重结构使得 FSVM 对样例的抑制作用被控制在合理范围内，有利于 FSVM 维持稳定、良好的分类效果。

这种通过预学习获得初步知识、再根据所获知识完善学习模型的方法与人类的学习行为极为相似，这是人类在生产实践中自然形成的一种行之有效的学习方法，在很多问题上具有普适性。文献[15]的鲁棒学习最小二乘支持向量机即采用了类似的思想来提高支持向量回归机的学习性能。其具体方法是训练 LS-SVM 模型，根据模型预测误差更新 SVM 权值，重复上述过程，直至满足终止条件。这是一个多次学习的过程，自然会消耗较多的学习时间。本文方法仅预学习一次，预学习代价远小于 SVM。

使用式(10)计算模糊权重意味着需要构造最优点分类面，这将为分类器的学习带来额外高昂的预处理代价。本文通过如下方法构造预分类面，将模糊权重计算代价维持在较低水平上，实验将表明这种方法的有效性和本文思想的合理性。

令 M 为正的输入向量集，N 为负的输入向量集。o_M 和 o_N 是对应于 M 和 N 的两个预分类面构造参考点，由下式来产生：

$$o_M=(c_M-c_N(1-g))/g \tag{11}$$

$$o_N=(c_N-c_M(1-g))/g \tag{12}$$

其中，c_M 和 c_N 分别为正、负类中心，即

$$c_M=\frac{1}{|M|}\sum_{i=1}^{|M|}\Phi(m_i)(m_i\in M) \tag{13}$$

$$c_N=\frac{1}{|N|}\sum_{j=1}^{|N|}\Phi(n_j)(n_j\in N) \tag{14}$$

$g\in(0,1]$是由实验确定的一个常数，$|\cdot|$表示集合基数。

对于任意 $m\in M$，其预分类面 H_m 对应的决策函数通过下式来构造：

$$f_m(x)=(\Phi(m)-o_M)^{\mathrm{T}}(\Phi(x)-\Phi(m)) \tag{15}$$

上标“T”表示转置。任意数据 $n\in N$ 的预分类面 H_n 对应的决策函数通过下式来构造：

$$f_n(x)=(\Phi(n)-o_N)^{\mathrm{T}}(\Phi(x)-\Phi(n)) \tag{16}$$

输入向量 $m\in M$ 对应的点误分率为：

$$e_m=\frac{1}{2(|M|+|N|)}(|M|+|N|+\sum_{x\in M}sign(f_m(x))+\sum_{x\in N}sign(-f_m(x))) \tag{17}$$

同理，输入向量 $n\in N$ 对应的点误分率为：

$$e_n=\frac{1}{2(|M|+|N|)}(|N|+|M|+\sum_{x\in N}sign(f_n(x))+\sum_{x\in M}sign(-f_n(x))) \tag{18}$$

以式(17)、式(18)所得点误分率代替式(10)中的点误分率。

基于预分类的 FSVM 模糊权重计算方法可以总结如下。

算法 1 基于预分类的 FSVM 模糊权重

输入：正类输入向量集 $M=\{m_1,\cdots,m_{|M|}\}$ 和负类输入向量集 $N=\{n_1,\cdots,n_{|N|}\}$

输出：$x_k\in M\cup N(k=1,\cdots,|M\cup N|)$ 的模糊权重 $\{s_k\}_{k=1}^{|M\cup N|}$ ($x_k\in M\cup N$)

过程：

for i=1 to |M| do

$$e_{m_i}=\frac{1}{2(|M|+|N|)}(|M|+|N|+\sum_{j=1}^{|M|}\mathrm{sign}(f_{m_i}(m_j))+\sum_{k=1}^{|N|}\mathrm{sign}(-f_{m_i}(n_k)))$$

end for

for i=1 to |N| do

$$e_{n_i}=\frac{1}{2(|M|+|N|)}(|N|+|M|+\sum_{j=1}^{|N|}\mathrm{sign}(f_{n_i}(n_j))+\sum_{k=1}^{|M|}\mathrm{sign}(-f_{n_i}(m_k)))$$

end for

for k=1 to |M∪N| do

$s_k=e_{x_k}-\min_{x_i\in M\cup N}e_{x_i}+1-\min_{x_i\in M\cup N}e_{x_i}$;

if $s_k>1$

$s_k=1$;

end if;

end for

return $\{s_k\}_{k=1}^{|M\cup N|}$。

容易知道算法的时间复杂度为 $O(l^3)$，其中 $l=|M|+|N|$ 为训练数据总数。时间复杂度从理论上提供了算法效率评估的一种参考，实际运行时间则比较客观地反映了算法实际响应速度和算法实用性。本文算法实际运行时间远小于 SVM。因为本文算法的基本操作是数量较为确定的简单算术运算，而 SVM 的每一次迭代都包含着极为复杂的二次规划寻优计算。

下文中将使用上述算法的 FSVM 即基于预分类的 FSVM，称为 CIFSVM(使用分类信息的 FSVM)。CIFSVM 的训练步骤可以总结如下：

1)使用算法 1 计算训练集中各个输入向量 x_k 的模糊权重 s_k；

2)将模糊权重 s_k 应用于式(5)的 FSVM；

3)求解 FSVM 得到决策函数。

4 实验结果

本节将对所提方法的有效性进行实验研究与分析，待比较的算法包括标准 SVM、使用中心距离型模糊权重的 SVM(称为 DFSVM，FSVM using center distance fuzzy weight)和本文提出的基于预分类的 FSVM，即 CIFSVM。实验所用数据集包括来自智能数据分析(Intelligent Data Analysis，IDA)库和 UCI 机器学习库[13]的一些公共数据集，其中 IDA 数据集主要从 UCI、DELVE 和统计与逻辑学习算法 STATLOG 等几个基准库中搜集而来，并且将其中原来不是二类可分的分类问题随机划分为两类，进一步对所有集合实施了预处理，以便于人们检验自己的算法。本文所用 IDA 数据集的信息如表 1 所列。在使用 IDA 数据集实验时，在每种数据中随机选取 5 个数据集进行交叉验证测试，即进行 5 次实验，每次以其中一个数据集作为训练集，以其余 4 个数据集作为测试集，最后取 5 次实验的平均值作为算法在该种数据上的实验结果。

表 1 IDA 数据集信息及其参数设置

数据集	训练样例数	测试样例数	属性	γ	C
banana	400	400 * 4	2	1	1
diabetis	468	468 * 4	8	0.50	128
flare_solar	666	666 * 4	9	0.25	64
german	700	700 * 4	20	0.05	128
image	1300	1300 * 4	18	0.05	128
thyroid	140	140 * 4	5	0.10	128
waveform	400	400 * 4	21	0.06	1

实验所用 UCI 数据集包括：Bupa、Glass、(Statlog)Heart、PageBlocks、Pima、Waveformnoise 和 CreditApproval，其中 Bupa 是 BUPA 医疗研究中心搜集的以血液检测的 5 项指标加上 1 项每天饮酒量构成的反映肝病特征的一些数据。Glass 包含 6 种玻璃的氧化物含量信息，样本的每个属性具有实型值。(Statlog)Heart 包含了与心脏病有关的信息。PageBlocks 搜集了文档页面布局特征信息。Pima 收集了印第安人 Pima 部落关于糖尿病的一些特征。Waveformnoise 包含了 3 种加噪波信息。CreditApproval 是关于客户信用评价的样本集。Glass、PageBlocks 和 Waveformnoise 为多类问题，使用时以一对多方式转化为二类问题。UCI 数据集信息如表 2 所列。

表 2　UCI 数据集信息及其参数设置

数据集	样例数	属性	类数	σ	C
Bupa	345	6	2	20	8
Glass	214	8	2	0.5	16
(Statlog) Heart	270	13	2	32	1
PageBlocks	5473	10	2	40	128
Pima	768	8	2	62	4
Waveformnoise	5000	40	2	16	2
CreditApproval	653	15	2	128	1

表 3 和表 4 分别给出了算法在 IDA 和 UCI 数据集上的预测精度。

表 3　算法在 IDA 数据集上的预测精度(%)

数据集\分类器	SVM	DFSVM	CIFSVM
banana	90.61	89.46	91.25
diabetis	87.20	78.98	88.04
flare_solar	61.79	67.59	68.07
german	87.99	84.44	91.63
image	87.94	93.96	97.60
thyroid	95.96	96.14	98.39
waveform	91.15	89.55	91.20

表 4　算法在 UCI 数据集上的预测精度(%)

数据集\分类器	SVM	DFSVM	CIFSVM
Bupa	66.20	65.84	67.24
Glass	68.95	68.14	69
(Statlog) Heart	63.33	59.81	64.07
PageBlocks	87.63	91.53	92.45
Pima	72.94	70.59	74.28
Waveformnoise	84.76	88.78	89.78
CreditApproval	66.96	57.46	67.34

在 UCI 数据集上实验时，将每组数据集随机分为大小基本相等的 5 个子集进行类似于 IDA 数据集的交叉验证测试，取 5 次结果的平均值作为该组数据实验的最终结果。在所有实验中，径向基核函数被用于分类器训练，因为径向基核函数能够产生复杂程度各异的分类边界，具有较好的拟合能力。在使用径向基核函数时需要确定核参数 γ 和惩罚因子 C 的值。对于 IDA 数据集，核参数 γ 基于 $1/(\lambda\times$维数)试探选取；对于 UCI 数据集，核参数 γ 基于 $1/(2\sigma^2)$试探选取，惩罚因子 C 在{0.5,1,2,4,8,16,32,64,128}范围内选取。实验确保不同算法在同一数据集上使用相同的核参数 γ 和惩罚因子 C。CIFSVM 算法中的参数 g 固定为 0.01。

从表 3、表 4 可以看出，CIFSVM 在所使用的全部 14 个数据集上的预测精度均高于 SVM 和 DFSVM，并且 CIFSVM 不是基于最佳条件(包括 g 值及模糊权重)获得此结果的，由此可见其有效性和合理性。DFSVM 在 flare_solar、image、thyroid、PageBlocks 和 Waveformnoise 共 5 个数据集上的预测精度高于 SVM，而 SVM 在其余 9 个数据集上的预测精度高于 DFSVM。因此，与 DFSVM 相比，SVM 仍然保持着相对的优势。实验结果表明，合适的模糊权重有利于 SVM 维持良好的泛化性能，而不合适的模糊权重会向训练过程传递错误信息并干扰训练，以至于降低 SVM 的分类性能。因此，调整 SVM 训练参数配置方法存在损害分类器工作性能的风险。传统 SVM 虽然在训练参数配置上显得较为简单和保守，但是却能在很大程度上减少或避免分类器性能发生异常。

结束语　恰当配置训练参数有助于提高 SVM 抵御离群点及噪声数据不利影响的能力。本文在综合考虑单个样例作用效果和分类器对离群点及噪声数据敏感性的基础上提出了基于预分类的 FSVM，该方法通过分别关联于各个样例的预分类面评估相应样例在未来分类中的作用效果，并且考虑分类器对离群点及噪声数据的敏感性，在此基础上形成模糊权重和基于预分类的 FSVM。按照这种参数配置方法训练的 SVM 能够根据不同训练集的可分性特征自适应地对可能的离群点及噪声数据适度地进行抑制，使 SVM 维持稳定、良好的工作性能。

近年来，粗糙集理论在机器学习领域获得了有效应用，成为获取样本知识的一种新工具，基于粗糙集的 SVM[14] 有望更好地处理机器学习中的不确定性知识。

参 考 文 献

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

[2] Burges C. A tutorial on support vector machines for pattern recognition [J]. Data Mining and Knowledge Discovery, 1998, 2 (2): 121-167

[3] Malon C, Uchida S, Suzuki M. Mathematical symbol recognition with support vector machines [J]. Pattern Recognition Letters, 2008, 29(9): 1326-1332

[4] Lin Chun-fu, Wang Sheng-de. Fuzzy support vector machine [J]. IEEE Trans. Neural Netw., 2002, 13(2): 464-471

[5] Huang H P, Liu Y H. Fuzzy support vector machines for pattern recognition and data mining [J]. International Journal of Fuzzy Systems, 2002, 4(3): 826-835

[6] Wang Yong-qiao, Wang Shou-yang, Lai K K. A New Fuzzy Support Vector Machine to Evaluate Credit Risk [J]. IEEE Transactions on Fuzzy Systems, 2005, 13(6): 820-831

[7] Liu Y, Zheng Y F. Soft SVM and Its Application in Video-Object Extraction [J]. IEEE Transactions on Signal Processing, 2007, 55(7): 3272-3282

[8] Batuwita R, Palade V. FSVM-CIL: Fuzzy support vector machines for class imbalance learning [J]. IEEE Transactions on Fuzzy Systems, 2010, 18(3): 558-571

[9] 秦传东，刘三阳，张市芳. 基于不平衡数据分类的一种平衡模糊支持向量机[J]. 计算机科学，2012，39(6)：188-190

[10] 吴青，刘三阳，杜喆. 基于边界向量提取的模糊支持向量机方法[J]. 模式识别与人工智能，2008，21(3)：332-337

[11] 王熙照，崔芳芳，鲁淑霞. 密度加权近似支持向量机[J]. 计算机科学，2012，39(1)：182-184

[12] Jiang X F, Zhang Y, Lv J C. Fuzzy SVM with a new fuzzy membership function [J]. Neural Comput & Applic, 2006, 15: 268-276

[13] Newman D J, Hettich S, Blake C L, et al. UCI Repository of machine learning databases [OL]. http://www.ics.uci.edu/mlearn/MLRepository.html, 1998

[14] Chen De-gang, He Qiang, Wang Xi-zhao. FRSVMs: Fuzzy rough set based support vector machines[J]. Fuzzy Sets and Systems, 2010, 161: 596-607

[15] 张淑宁，王福利，尤富强，等. 基于鲁棒学习的最小二乘支持向量机及其应用[J]. 控制与决策，2010，25(8)：1169-1177