

一种 SVM 增量学习算法 α -ISVM*

萧嵘, 王继成, 孙正兴, 张福炎

(南京大学 计算机软件新技术国家重点实验室, 江苏 南京, 210093);

(南京大学 计算机科学与技术系, 江苏 南京, 210093)

E-mail: cloud@graphics.nju.edu.cn

http://www.nju.edu.cn

摘要: 基于 SVM (support vector machine) 理论的分类算法, 由于其完善的理论基础和良好的试验结果, 目前已逐渐引起国内外研究者的关注. 深入分析了 SVM 理论中 SV (support vector, 支持向量) 集的特点, 给出一种简单的 SVM 增量学习算法. 在此基础上, 进一步提出了一种基于遗忘因子 α 的 SVM 增量学习改进算法 α -ISVM. 该算法通过在增量学习中逐步积累样本的空间分布知识, 使得对样本进行有选择地遗忘成为可能. 理论分析和实验结果表明, 该算法能在保证分类精度的同时, 有效地提高训练速度并降低存储空间的占用.

关键词: 支持向量机; 分类; 模式识别; 增量学习; 机器学习

中图法分类号: TP18 **文献标识码:** A

随着信息时代的到来, 特别是随着网络迅猛发展而出现的“信息爆炸”问题, 向传统的信息挖掘、知识获取技术提出了巨大的挑战. 一方面, 我们现在拥有的数据极大丰富, 其间蕴含的信息和知识具有很大的潜在价值; 另一方面, 信息的更新速度也达到了令人吃惊的地步. 因此, 具有增量学习功能的数据分类技术, 正逐渐成为当前信息的智能化发现和挖掘的关键技术之一. 与普通的数据分类技术相比, 增量学习分类技术具有显著的优越性, 这主要表现在两个方面: 一方面由于其无须保存历史数据, 从而减少了存储空间的占用; 另一方面, 由于其在新的训练中充分利用了历史的训练结果, 从而显著地减少了后继训练的时间.

现有的增量学习算法^[1-2]大多采用决策树算法和神经网络算法来实现, 一般在不同程度上具有以下两方面的缺点: 一方面由于缺乏对整个样本集期望风险的控制, 算法易于对训练数据产生过量匹配; 另一方面, 由于缺乏对训练数据有选择的遗忘淘汰机制, 在很大程度上影响了分类精度. 基于结构风险最小化理论的 SVM (support vector machine, 支持向量机) 学习算法是少数可以成功解决第一个问题的学习算法, 但是, 经典的 SVM 学习算法并不直接支持增量式的学习. 因此, SVM 分类算法中的增量学习研究具有重要的理论意义和实用价值.

本文首先对 SVM 理论中 SV (support vector, 支持向量) 集的特点进行了深入分析; 然后给出了一种简单的 SVM 增量学习算法, 并在此基础上作出改进, 提出了一种基于遗忘因子 α 的 SVM 增量学习算法—— α -ISVM; 最后给出实验结果与分析.

* 收稿日期: 2000-04-04; 修改日期: 2000-07-10

基金项目: 国家自然科学基金资助项目(69903006, 60073030); 江苏省“九五”科技重点攻关资助项目(BE96017)

作者简介: 萧嵘(1972-), 男, 江苏苏州人, 博士, 主要研究领域为机器学习、数据挖掘技术; 王继成(1973-), 男, 江苏镇江人, 博士, 讲师, 主要研究领域为信息检索与挖掘; 孙正兴(1964-), 男, 江苏张家港人, 博士, 副教授, 主要研究领域为多媒体辅助工程、数据挖掘、数字化图书馆; 张福炎(1939-), 男, 江苏常州人, 教授, 博士生导师, 主要研究领域为多媒体技术、数字化图书馆.

1 理论准备

在讨论具有增量学习能力的 SVM 分类算法之前,我们先简单地回顾一下 SVM 理论,并对 SV 集的特点进行分析,这构成了本文的理论基础.

1.1 SVM 理论概述

支持向量机的理论最初来自于对数据二值分类问题的处理.其机理可以简单地描述为:寻找一个满足分类要求的最优分割超平面^[3],使其在保证分类精度的同时最大化超平面两侧的空白区域(如图 1 所示).这使得 SVM 分类器的结果不仅在训练集上得到优化,而且在整个样本集上的风险也拥有上界,这就是 SVM 的结构风险最小化的思想.在线性不可分的问题上,Vapnik^[4]等人成功地引入了核空间理论,将低维的输入空间数据通过非线性映射函数映射到高维属性空间,从而把分类问题转化到高维属性空间进行.可以证明,如果选用适当的映射函数,大多数输入空间线性不可分问题在属性空间可以转化为线性可分问题.这一问题的解决,使得 SVM 分类器正式成为通用的分类器之一.

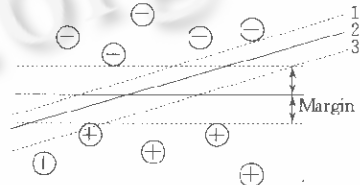


Fig. 1 The data set in R^2 , shattered by hyperplane
图1 数据点集的超平面分割

1.2 SV 的概念与特点分析

最优分割超平面的计算问题可以描述为一个条件极值问题,Vapnik 通过求解下列 Lagrange 函数的鞍点获得其最优解:

$$L_p = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum_i \alpha_i \{y_i(\Phi(x_i) \cdot w + b) - 1 + \xi_i\} - \sum_i \mu_i \xi_i \quad (1)$$

其中 w 和 b 分别为属性空间最优超平面的法向量和阈值, ξ_i 和 μ_i 分别为非负的 Lagrange 乘数, C 为非负的误差控制参数, $\Phi(x)$ 为输入空间向属性空间映射的函数.

由 Karush-Kuhn-Tucker 定理可知,最优解满足以下条件:

$$w = \sum_i \alpha_i y_i \Phi(x_i), \quad (2)$$

$$0 \leq \alpha_i \leq C, \sum_i \alpha_i y_i = 0. \quad (3)$$

一般情况下,在式(2) w 的展开式中,大多数系数 α_i 为零值,并不影响分类的结果.而对 w 的确定有贡献的仅仅是非零值的 α_i 所对应的 x_i ,这就是 SV (support vector,支持向量).因此,SV 集充分描述了整个训练数据集数据的特征,对 SV 集的划分等价于对整个数据集的分割.在大多数情况下,训练集中 SV 的数量只占训练样本集的很少一部分,即 $N_{sv} \ll N_{train}$.因此可以使用 SV 集取代训练样本集进行分类学习,使得在不影响分类精度的同时极大地减少训练时间.

本文进一步将 SV 集中的向量划分为两种类型:一种类型是对应于 $\alpha_i = C$ 的 SV 向量,这类向量也被称为 BSV (boundary support vector,边界支持向量),代表了所有不能被正确分类的样本向量;另一种类型的 SV,其对应的 α_i 值满足 $0 < \alpha_i < C$,这类向量代表了训练集中大部分样本的分类特征,并与 BSV 共同决定最终的结果分类器的形式.

综上所述,我们认为,SV 集与训练样本集的分类等价性在增量学习算法的研究中具有重要的意义.

2 具有增量学习功能的 SVM 分类算法

在增量学习的情况下,由于增量训练集的引入,打破了 SV 集和整个训练样本集的等价关系,从而有必要寻找新的 SV 集.因此,本文在上述讨论的基础上,给出了一种简单的 SVM 增量算法,并通过对内样本的淘汰分析提出了基于遗忘因子 α 的 SVM 增量学习算法—— α -ISVM.

2.1 简单的 SVM 增量学习算法

为简化问题的复杂度,可以简单地把问题描述为:

- 前提条件:存在历史数据集 A ,增量样本集 B ,并且假定保证对于两个数据集有 $A \cap B = \emptyset$; Γ^1 和 A_{sv}^1 分别为数据集 A 上的初始 SVM 分类器以及对应的 SV 集;
- 问题的目标:在初始分类器 Γ^1 和对应 SV 集 A_{sv}^1 的基础上,寻找基于样本集合 $A \cup B$ 上 SVM 分类器 Γ 和对应的 SV 集 A_{sv} .

结合上一节的讨论,我们对这个问题作进一步的分析.根据分类器 Γ^1 对样本集 B 分类的正确性,可以将 B 集划分为测试错误集 B_{err} 和测试正确集 B_{ok} ,此处假定 $B_{err} \neq \emptyset$.根据上一节的定义,由于 BSV 代表了所有不能被正确分类的样本向量,因此,集合 B_{err} 应该以相当大概率包含问题求解集的 A_{sv} 中的元素,也就是说,一般有 $B_{err} \cap A_{sv} \neq \emptyset$.将集合 A_{sv}^1 和 B_{err} 的并集 A^1 作为新的训练集,得到一个新的分类器 Γ^2 和 SV 集 A_{sv}^2 .根据上面的分析,一般有 $\|A_{sv}^1 \cap A_{sv}\| \leq \|A_{sv}^2 \cap A_{sv}\|$,这表明对于测试集 $A \cup B$,一般有, A_{sv}^2 比 A_{sv}^1 更接近所要求解的 A_{sv} .可以证明,随着训练过程中训练样本集逐步积累,上述命题成立的概率将迅速向 1 逼近,这为下面的迭代算法提供了理论依据.

将集合 A 中除去 SV 集的剩余样本与集合 B_{ok} 合并在一起,为新生成的分类器 Γ^2 构建新的“增量样本集” B^1 .重复上面的工作,通过不断的迭代,得到新的分类器 $\Gamma^3, \Gamma^4, \dots$ 显然,当迭代至 $B_{err}^i = \emptyset$ 时,分类器 Γ^i 与 Γ 近似等价.

前面的分析建立在两个隐含前提下.第 1 个前提是假定 $B_{err} \neq \emptyset$.当 $B_{err} = \emptyset$ 时,证明原分类器覆盖了新增样本空间,因此不需要进行增量学习.第 2 个假设是分类器 Γ^1 对其训练数据有较高的精度,显然,这是可以接受的.我们在分析的基础上提出了简单的 SVM 增量学习算法 SISVM (simple incremental SVM).

算法

- (1) 在初始训练集上训练 SVM 初始分类器 Γ^1 ;
- (2) 根据分类器 Γ^1 对样本集 B 分类的正确性,可以将 B 集划分为测试错误集 B_{err} 和测试正确集 B_{ok} ;
- (3) 如果 $B_{err} = \emptyset$,则迭代结束;

如果 $B_{err} \neq \emptyset$,则将集合 A_{sv}^1 和 B_{err} 的并集 A^1 作为新的训练集,得到一个新的分类器 Γ^2 和 SV 集 A_{sv}^2 ,并将集合 A 中除去 SV 集的剩余样本与集合 B_{ok} 合并在一起,为新生成的分类器 Γ^2 构建新的“增量样本集” B^1 ;

- (4) 令 $B = B^1$,重复上面(2)、(3)两个步骤的工作;
- (5) 通过多次迭代,算法收敛于分类器 Γ^i .令 $\Gamma = \Gamma^i$, Γ 为所求解的结果.

2.2 改进的 SVM 增量学习算法

与普通的 SVM 学习算法相比,简单 SVM 增量算法 SISVM 的改进主要体现在训练速度上.但是, SISVM 算法仍然必须保存全部的历史数据,并且在初始样本不足的情况下,后继的增量训练将出现“振荡”现象,影响了训练的收敛速度.因此,在很多情况下 SISVM 算法依旧不能满足要求.如有选择地淘汰一部分对分类没有多大贡献的数据,从而有效地降低对空间存储的要求并进一步加快训练的收敛速度,则成为改进算法所研究的重点.

考察实验(如图 2 所示),在初始训练集 $X^0 = \{A, B, C, D, E, F, G, H, I, J, K, L, M, N, O\}$ 和加入增量集 $X^1 = \{P, Q, R, S, T, U, V, W\}$ 后的样本集 $X^0 \cup X^1$ 上分别使用 SVM 分类算法学习,得到初始训练集的 SV 集 $X_{sv}^0 = \{H, I, J, K, L, M, N, O\}$ 和样本集 $X^0 \cup X^1$ 的 SV 集 $X_{sv}^1 = X_{sv}^0 \cup \{E, R, U, V, W\}$. 比较图中的实验结果可以发现,样本集 $\{A, B, C, D, F, G\}$ 远离分类平面 ϕ_{hp} , 而入选新 SV 集的样本点 E 则相距分类平面较近. 显然,两次训练中前者对分类产生的影响远不如后者. 如果用上节提出的 SISVM 算法进行多轮训练实验,这一点将表现得更为明显. 为此,本文提出了改进的 SVM 增量学习算法—— α -ISVM.

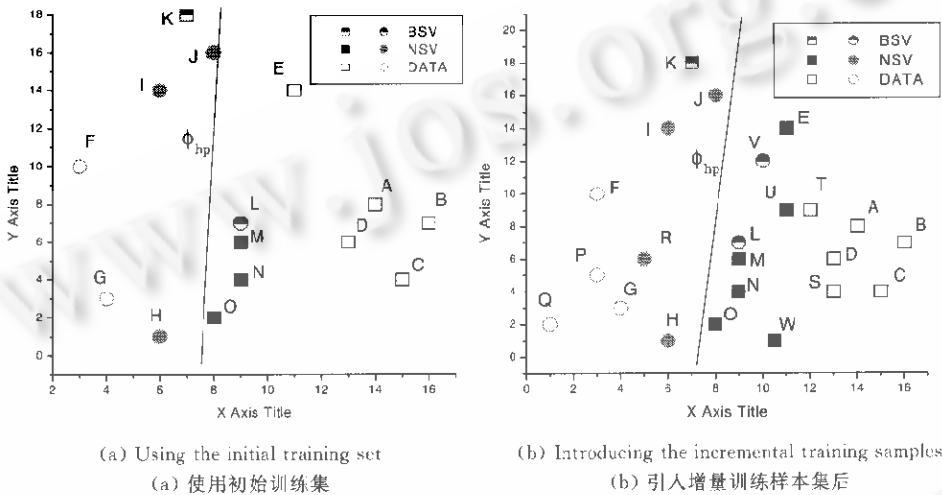


Fig. 2 The influence of the incremental training samples

图 2 增量样本对 SVM 分类结果的影响

在进一步分析之前,我们首先定义下面 3 个概念:(1) 样本集中从未入选过任何一轮训练的 SV 集的样本称为类的内样本,通常这一部分样本占有相当大的比例;(2) 那些始终在各轮 SV 集出现的样本称为类的边界样本;(3) 在 SV 集中抖动出现的样本被定义为类的准边界样本. 不难看出,不同的样本将根据其不同空间分布特性对最终分类学习器作出不同的贡献. 从直观的角度来看,边界样本刻画了样本空间的主要的分类知识,而准边界样本是对边界样本分类知识的补充,它们对于积累知识有至关重要的作用. 而内样本代表的那部分知识则完全是可以通过其他两类样本来表示的. 基于上面的分析我们对 SISVM 算法作以下两方面的改进:首先是按一定比率逐步淘汰类的内样本,减少历史样本的存储和训练样本集的规模;其次是按一定规则选取较优的准边界样本,加速 SV 集搜索的收敛速度.

2.3 SVM 增量学习改进算法—— α -ISVM

首先,引进权值 Ψ_i 刻画训练集中样本点 x_i 的空间分布特性,并定义:当 $\Psi_i \geq C_b$ 时, x_i 为边界样本点;当 $C_b > \Psi_i \geq C_s$ 时, x_i 为准边界样本点;当 $C_{inside} > \Psi_i \geq 0$ 时, x_i 为类的内样本;当 $C_s > \Psi_i \geq C_{inside}$ 时, x_i 为处于权值调定中的样本点;当 $\Psi_i = C_{init}$ 时, x_i 为新增样本,且 $C_b > C_s > C_{init} > C_{inside} > 0$.

改进的增量算法—— α -ISVM 包含以下两个过程:

(1) 初始化过程

使用初始分类样本集 X^0 训练 SVM 分类器 \mathcal{D}^0 , 得到 SV 集 X_{sv}^0 和非 SV 集 X_{nsv}^0 , 并按权值调整

规则 1 为两个集合的每个样本赋予权值.

权值调整规则 1(赋值规则):

(i) 集合 X_{sv}^0 中, BSV 样本赋予权值 C_b , 其余样本赋予权值 C_i ;

(ii) 集合 X_{nsv}^0 中的每个样本赋予相同权值 C_{init} .

(2) 增量训练过程

首先分别定义新增加的样本集 X^{i+1} 、上轮训练结果集 R^i 和分类器 Γ^i 为该轮训练的测试集 V_0^{i+1} 、初始训练集 T_0^{i+1} 和初始分类器 Γ_0^{i+1} , 并为集合 X^{i+1} 中的每个样本赋予相同权值 C_{init} , 然后进行若干次训练及调整, 寻找该轮增量训练的最优分类器 Γ^{i+1} , 这称为搜索训练. 具体搜索步骤如下:

• 使用分类器 Γ_0^{i+1} 测试样本集 V_0^{i+1} , 根据测试结果将 V_0^{i+1} 分为两类: 分类错误集 E_0^{i+1} 和分类正确集 O_0^{i+1} . 按权值调整规则 2 为两个集合中每个样本调整权值. 选取集合 T_0^{i+1} 的 SV 集 $SV(T_0^{i+1})$ 和集合 E_0^{i+1} 中的所有样本, 并在集合 T_0^{i+1} 的非 SV 集 $NSV(T_0^{i+1})$ 和集合 O_0^{i+1} 中按权值大小和挑选率 β 随机选取若干个样本组成新的训练集 T_1^{i+1} , 剩余的样本组成测试集 V_1^{i+1} .

权值调整规则 2:

(i) 集合 O_0^{i+1} 中的样本 x_s 的权值 $\Psi_s = \max(0, \Psi_s - d_{ok})$;

(ii) 集合 E_0^{i+1} 中的样本 x_i 的权值 $\Psi_i = \Psi_i + d_{er}$, 其中 $(d_{crit} > 0, d_{ok} > 0)$.

• 在集合 T_1^{i+1} 上使用标准 SVM 分类算法进行训练, 得到分类器 Γ_1^{i+1} , 并按权值调整规则 3 为集中每一个样本 x_i 调整其权值 Ψ_i .

权值调整规则 3:

(i) 在 SV 集 $SV(T_1^{i+1})$ 上, 令 $\Psi_s = \max(C_s, \gamma\Psi_s)$, 其中 $\gamma > 1$;

(ii) 在非 SV 集 $NSV(T_1^{i+1})$ 上, 令 $\Psi_s = \Psi_s - d_{ok}$.

• 重复上述两个步骤, 得到一系列分类器 $\Gamma_0^{i+1}, \Gamma_1^{i+1}, \dots$, 直到某分类器 Γ_j^{i+1} 为止, 其对相应的测试集 V_j^{i+1} 分类精度为 100%. 这时有分类器 Γ_j^{i+1} 与 Γ^{i+1} 近似等价.

• 在集合 V_j^{i+1} 的内样本集中, 根据遗忘因子 α 按样本权值大小淘汰其中部分元素, 将剩余的样本和训练集 T_j^{i+1} 共同构成该轮训练的结果集 R^{i+1} .

2.4 SVM 增量学习算法的性能分析

SVM 分类器的计算复杂度受到训练集的规模 L 、输入空间的维度 d_i 和支持向量的数目 N_{sv} 及其 SV 分布的影响. 实验证明, 在大多数情况下, $N_{sv}/L \ll 1$, 并且 SV 集中大多数 SV 不是 BSV, 这时有分类器的计算复杂度 $O(N_{sv}^3 + LN_{sv}^2 + d_i LN_{sv})^{[5]}$.

使用规模为 L_{new} 的增量样本集进行增量训练, 得到 SV 集的规模为 N'_{sv} . 一般有 $N'_{sv}/N_{sv} \approx 1$, 因此简单的增量算法的计算复杂度为 $O(N_{sv}^3 + d_i N_{sv}^2)$, 远小于重新训练的计算复杂度 $O(N_{sv}^3 + (L + L_{new})N_{sv}^2 + d_i(L + L_{new})N_{sv})$.

引入样本分布特点的增量改进算法 α -ISVM, 并通过缩减历史样本集的规模和加速 SV 集搜索的收敛速度, 进一步提高了训练速度. 但是另一方面, 遗忘因子 α 也直接影响了分类精度, 当 $\alpha = 0$ 时, 所有的内样本均被保留, 本算法得到的分类器和在现存所有训练样本集的并集上训练出的 SVM 分类器近似等价, 从而保证了前者的分类精度和过量匹配控制性能不弱于后者; 而当 $\alpha > 0$ 时, 部分内样本在训练中被淘汰, 并在增量学习过程中被逐步“遗忘”, 对后继分类器的精度有可能产生影响, 但由于边界样本集和准边界样本集的存在, 理论和实验都证明, 在大多数情况下这一影响是很微弱的.

3 实验结果和分析

在以上研究的基础上,本文使用 α -ISVM 算法在一个文本库上进行增量学习实验(其中运行的软件环境是 Matlab5.3/VC6.0/WINNT,硬件环境是 PIII333/256MB 内存),实现对电子文本的自动分类.实验数据库中共计 1 493 个文档(其中每个文档用 10^4 维的词条向量表示),通过手工标定为两个类别,然后选择其中 317 个样本做测试,其中 1 176 个样本做训练.

实验设计如下:选用 1 176 中的一半作为初始训练,其余的数据随机分为 5 组,对不同的遗忘因子 α 做增量训练测试(其中实验 A 对应 $\alpha=0.9$,实验 B 对应 $\alpha=0.5$,实验 C 对应 $\alpha=0$),实验 D 代表的是非增量训练测试.实验结果见表 1 和表 2(其中 SVM 训练核函数采用一阶多项式函数).

Table 1 Comparison of classification results between incremental learning and iterative learning

表 1 增量学习结果和重复学习结果的比较

Training set ^①	Samples in incremental set ^②	Experiment ^③ A		Experiment B		Experiment C		Experiment D	
		Time ^④ (s)	Precision ^⑤ (%)	Time (s)	Precision (%)	Time (s)	Precision (%)	Time (s)	Precision (%)
First training ^⑥	588	106.4	91.3	106.4	91.3	106.4	91.3	106.4	91.3
Incremental ^⑦ 1	246	77.1	92.7	76.9	92.7	80.4	92.7	247	92.7
Incremental 2	79	83.7	92.9	87.4	92.9	89.2	93.1	379.5	93.1
Incremental 3	187	75.4	93.6	92.5	93.7	95.1	93.9	491.4	93.9
Incremental 4	42	88.3	93.8	90.5	93.8	97.4	94.3	582.7	94.3
Incremental 5	34	96.5	93.5	98.3	93.7	109.3	94.2	613.8	94.2

①训练集,②增量集样本数,③实验,④时间,⑤精度,⑥初始训练,⑦增量.

Table 2 Number of samples in the result set of each training phase with different discarding factor

表 2 不同遗忘因子下的各轮训练的结果集样本数

Sifting factor $\alpha^{\text{①}}$	First training ^②	Incremental ^③ 1	Incremental 2	Incremental 3	Incremental 4	Incremental 5
0 (Experiment ^④ C)	588	834	913	1 100	1 142	1 176
0.5 (Experiment B)	588	753	532	586	591	589
0.9 (Experiment A)	588	689	231	283	285	279

①遗忘因子,②初始训练,③增量,④实验.

从上面的实验可以看出, α 值的提高可以显著地减少历史数据的存储,同时也加速了分类训练的进行.通过分析 α 值对分类精度的影响可以发现,当 $\alpha=0$ 时分类精度不受影响,随着 α 值的增大,分类精度稍有降低.与训练 D 相比,实验 A 和 B 以稍低的精度换取了很高的训练速度,并减少了存储空间占用,而实验 C 则在精度不受影响的情况下使训练速度有了很大的提高.

4 结束语

本文通过对 SVM 分类算法中 SV 集属性的深入分析,对简单 SVM 的增量学习算法进行了研究,并在此基础上,提出了基于遗忘因子 α 的增量 SVM 学习算法 α -ISVM.算法充分使用了 SV 集属性,并通过权值的调整在多次训练中逐步积累起的关于样本空间分类特性的知识,使得对样本有选择的遗忘成为可能,同时也进一步提高了分类速度.实验结果表明,这一算法在保证分类精度的同时,能有效地提高训练速度并降低存储空间的占用.进一步的研究工作包括 SVM 增量学习中噪声样本淘汰算法、多值和兼值算法的设计等.

References:

[1] Ratsby, J. Incremental learning with sample queries. IEEE Transactions on Pattern Analysis and Machine Intelligence,

1998,20(8):883~888.

- [2] Yamauchi, K., Yamaguchi, N., Ishii, N. Incremental learning methods with retrieving of interfered patterns. *IEEE Transactions on Neural Networks*, 1999,10(11):1351~1365.
- [3] Koji, Tsuda. Optimal hyperplane classifier with adaptive norm; Koji Tsuda. TR-99-9, Ibaraki, Japan; Electrotechnical Laboratory, 1999. 2~3.
- [4] Vapnik, V. *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995. 5~13.
- [5] Christopher, J., Burges, C. A tutorial on support vector machines for pattern recognition. *Knowledge Discovery and Data Mining*, 1998,2(2):235~244.

An Incremental SVM Learning Algorithm α -ISVM'

XIAO Rong, WANG Ji-cheng, SUN Zheng-xing, ZHANG Fu-yan

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China);

(Institute of Multimedia Computing Technology, Nanjing University, Nanjing 210093, China)

E-mail: cloud@graphics.nju.edu.cn

http://www.nju.edu.cn

Abstract: The classification algorithm based on SVM (support vector machine) attracts more attention from researchers due to its perfect theoretical properties and good empirical results. In this paper, the properties of SV set are analyzed thoroughly, and a new learning method is introduced to extend the SVM Classification algorithm to incremental learning area. After that, a new improved incremental SVM learning algorithm is proposed, which is based on a sifting factor. This algorithm accumulates distribution knowledge of the training sample while the incremental training is proceeded, and thus makes it possible to discard samples optimally. The theoretical analysis and experimental results show that this algorithm could not only improve the training speed, but also reduce storage cost.

Key words: SVM (support vector machine); classification; pattern recognition; incremental learning; machine learning

* Received April 4, 2000; accepted July 10, 2000

Supported by the National Natural Science Foundation of China under Grant Nos. 69903006, 60073030; the Sci-Tech Project of the 'Ninth Five-Year-Plan' of Jiangsu Province of China under Grant No. BE96017