

文章编号: 1001-0920(2009)01-0007-06

基于遗传算法的 SVM 多分类决策树优化算法研究

连可, 陈世杰, 周建明, 龙兵, 王厚军

(电子科技大学 自动化工程学院, 成都 610054)

摘要: 设计一种基于遗传算法(GA)的支持向量机(SVM)多分类决策树优化算法, 以克服因传统 SVM 多分类决策树结构固定, 单个 SVM 节点在树中位置随意而引起“误差积累”现象严重的缺陷. 采用了 SVM 分类间隔作为 GA 适应度函数. 利用 GA 在每一决策节点自动选择最优或近优的分类决策, 最终自适应地实现了对决策树的优化. 仿真实验表明, 与传统方法相比, 所提出的方法可使“误差积累”现象明显降低, 分类质量大大提高.

关键词: 支持向量机; 遗传算法; 决策树; 误差积累

中图分类号: TP39 文献标识码: A

Study on GA-based SVM multi-class classification decision-tree optimization algorithm

LIAN Ke, CHEN Shi-jie, ZHOU Jian-ming, LONG Bing, WANG Hou-jun

(College of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China. Correspondent: LIAN Ke, E-mail: lian_k@163.com)

Abstract: For the fixed tree configuration of traditional support vector machine (SVM) multi-class classification decision-tree algorithms and the random positions of their decision nodes, the error accumulate is very severity. Therefore, we present a GA-based SVM multi-class classification decision-tree optimization algorithm. We adopt the “margin” of SVM as adaption function to design GA. Then, GA is used to create optimal or suboptimal decision-tree automatically, which makes the margin between two classes maximal at every decision node. Experiment results show that the error accumulation phenomena is weaken obviously and classification quality is advanced greatly.

Key words: Support vector machine; Genetic algorithm; Decision tree; Error accumulation

1 引言

支持向量机(SVM)是一种新的基于统计学习理论的机器学习方法. 自诞生以来, 由于其突出的小样本学习的能力和好的泛化性能, 已在不同的领域得到了广泛的应用. 但是, 传统的 SVM 只能处理二分类问题, 而在许多应用领域中(如医学诊断、设备故障诊断等), 待识别的样本往往包含多种模式, 因此, 研究利用 SVM 处理多类别分类问题具有重要的现实意义, 也是目前 SVM 研究的热点问题之一.

目前已提出的 SVM 多分类方法大致可分为两类: 一次性求解法和分解重构法. 一次性求解法是由 Weston 等人^[1]在 1998 年最早提出的, 它在所有训练样本上求解一个大型二次规划问题, 同时将所有

样本分开. 由于该方法变量个数多, 计算复杂度高, 实用性并不强. 而分解重构法是一种比一次性求解法更适用于实际应用的方法^[2], 其主要思想是将多类分类问题转化为多个两类分类问题, 并采用某种策略将多个两类分类器组合起来以实现多分类. 其中, one-against-rest (1-a-r)^[3] 和 one-against-one (1-a-1)^[4] 是较早提出的两种经典的 SVM 多分类法. 1-a-1 采用“一对一”的方法训练基本 SVM, 采用投票法组合策略进行多分类; 1-a-r 则采用“一对其余”的方法训练基本 SVM, 同时采用“最大输出”法实现多分类. 这两种方法的优点是分类精度较高. 缺点是: 1) 其泛化误差无界; 2) 对于 K 分类问题必须分别训练 $K(K-1)/2$ 个和 K 个基本 SVM, 且分类必须遍历所有训练的 SVM, 训练复杂, 分类效率低. 对

收稿日期: 2007-10-19; 修回日期: 2008-01-15.

基金项目: 国防基础科研项目(A1420061264).

作者简介: 连可(1978—), 男, 重庆璧山人, 博士生, 从事故障诊断、模式识别的研究; 王厚军(1961—), 男, 北京人, 教授, 博士生导师, 从事信号信息处理、电子测试等研究.

此, Decision-tree-based multiclass SVM (DT-SVM)^[5]和 Directed acyclic graph SVM (DAG-SVM)^[6]采用决策树结构进行了改进,使得训练和分类效率大大提高.但是由于两种方法均采用固定树结构,而且包括根节点在内各个决策节点的选择具有随意性^[7-9],其单个子分类器在决策树中的位置不同,分类结果可能不同,从而使得分类性能具有不确定性,往往会产生严重的“误差积累”现象.

将遗传算法(GA)用于SVM分类识别中SVM参数优化和样本的特征选择已有学者进行了研究^[10-12].但利用GA进行SVM多分类决策树结构优化的工作还鲜有尝试.本文提出一种基于遗传算法的SVM多分类决策树优化方法,利用遗传算法在决策树的每个决策节点进行两分类决策优化,以确定单个两分类器的位置,最终自动生成最优(或近优)决策树.由于GA优化的引入,使得决策树的结构具有自适应性,在保证训练和分类效率的同时,克服多分类决策树中“误差积累”影响.实验结果表明,与DT-SVM和DAG-SVM方法相比,新方法可使“误差积累”大大减小,分类质量得以提高.

2 支持向量机

支持向量机是一种建立在统计学习理论上的机器学习方法,它追求结构化风险最小而不是经验风险最小,具有很强的推广能力.SVM是从线性可分的二分类问题发展而来的,其基本思想是寻找两类样本的最优分类面,使得两类样本的分类间隔(margin)最大.以图1所示为例.图中:实心点和空心点分别代表两类样本; H 为分类线, H_1 和 H_2 分别为各类中离分类线最近的样本且平行于分类线的直线,它们之间的距离称为分类间隔.所谓最优分类线就是要求分类线不但能将两类正确分开,而且使分类间隔最大.

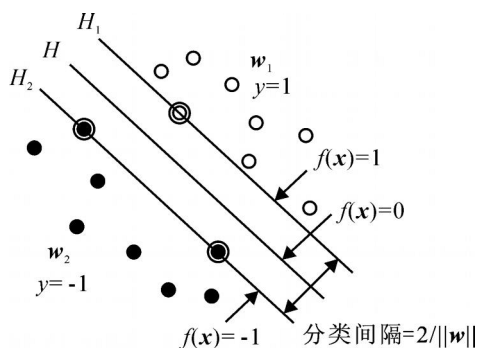


图1 SVM二分类问题

不失一般性,设训练样本为 (x_i, y_i) , $i = 1, 2, \dots, l$, $x \in R^n$, $y \in \{+1, -1\}$, l 为样本数, n 为输入维数.在线性可分的情况下,将两类样本完全分开的超平面为

$$w \cdot x + b = 0. \quad (1)$$

使分类间隔最大的超平面即为最优分类面.为此,要求解如下二次规划问题:

$$\begin{aligned} \min \quad & w^2/2, \\ \text{s. t.} \quad & y_i(w \cdot x_i + b) - 1 \leq 0, \\ & i = 1, 2, \dots, l. \end{aligned} \quad (2)$$

当训练样本集线性不可分时,需引入非负松弛变量 λ_i , $i = 1, 2, \dots, l$.求解最优分类面问题为

$$\begin{aligned} \min \quad & w^2/2 + C \sum_{i=1}^l \lambda_i, \\ \text{s. t.} \quad & y_i(w \cdot x_i + b) - 1 + \lambda_i \leq 0, \\ & \lambda_i \geq 0, i = 1, 2, \dots, l. \end{aligned} \quad (3)$$

式中 C 为惩罚参数, C 越大表示对错误分类的惩罚越大.通过Lagrange乘子法求解上述优化问题,可得最优决策函数为

$$f(x) = \text{sgn} \left[\sum_{i=1}^l y_i \lambda_i (x \cdot x_i) + b \right], \quad (4)$$

式中 λ_i 为Lagrange系数.在对输入测试样本 x 进行测试时,由式(4)确定 x 的所属类别.根据K-T条件,上述优化问题的解必须满足

$$\lambda_i (y_i (w \cdot x + b) - 1) = 0. \quad (5)$$

因此,对于多数样本 λ_i 将为零,而取值不为零的 λ_i 所对应的样本即为支持向量,它们通常只是全体样本中的很少一部分.这样,仅需要少量支持向量即可完成正确的样本分类.

当样本集为非线性时,可将样本 x 映射到某个高维空间 H ,并在 H 中使用线性分类器.根据Mercer条件,采用不同内积函数 $K(x_i \cdot x_j)$ 便可实现非线性的线性分类. $K(x_i \cdot x_j)$ 称为核函数,此时相应的最优决策函数变为

$$f(x) = \text{sgn} \left[\sum_{i=1}^l y_i \lambda_i K(x \cdot x_i) + b \right]. \quad (6)$$

3 传统的DAGSVM和DT-SVM

为了将两分类SVM推广到解决多类分类问题,同时保证训练和分类效率,Platt和Fumitake分别提出了基于有向无环图(DAG)和基于决策树(DT)的SVM多分类方法.

DAG-SVM的分类思路如图2(a)所示.它是有向无环图将多个两分类SVM组合起来.在训练阶段,每个SVM采用“1-a-1”的方法训练;而在决策阶段,使用从根节点开始的有向无环图.给定一个测试样本,从根节点开始根据分类器的输出值决定走其左侧还是右侧,如此直到叶子节点为止,得到样本所属的类值.该方法训练时共需训练 $K(K-1)/2$ 个两分类SVM分类器,测试时仅需遍历其中的 $K-1$ 个SVM.

DT-SVM 的思路与 DAG-SVM 类似,只不过它采用二叉树的组合策略,在每个决策点用“1-a-r”的方法将一类和其他类分割开,直至所有的节点都成为叶子节点,如图 2(b) 所示.这种方法所需训练的 SVM 数为 $K-1$ 个,分类所需遍历的 SVM 平均个数为 $(K-1)(K+2)/2K$ 个.

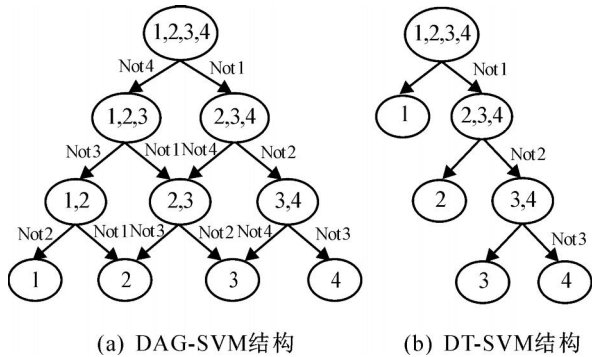


图 2 DAG-SVM 及 DT-SVM 结构示意图

上述两种方法的优点是其决策速度显著比“1-a-1”和“1-a-r”快.但共同的缺点是只要分类类别数固定,其决策树结构就是固定的,不能根据具体的分类问题作出自适应的调整;另外各个两分类 SVM 在决策树/图中的位置不同,其分类性能往往也会不同,越接近根节点的位置出现错分,其“误差积累”现象越严重.但传统的 DAG-SVM 方法和 DT-SVM 方法均没有考虑如何最优地安排各个两类 SVM 的位置问题,即没有考虑每个决策点上的决策优化问题.

4 基于遗传算法的决策树生成

如前所述,不同的决策树结构会对分类精度产生不同的影响,而且这种影响有可能产生“误差积累”现象,即如果在某个节点上发生分类错误,则错误会沿树结构向后续节点扩散传播,最终导致分类结果与实际情况相去甚远的现象.理论上讲,对于 K 类问题,所有可能构造的严格二叉树的数目为^[13]

$$N_k = \sum_{i=1}^{K-1} (2i-1), \quad (7)$$

其中 $K \geq 3$. 因此,如何构造具有最优(或近优)分类性能的决策树便成为 SVM 决策树组合多分类策略的关键.本节采用遗传算法,以类间分类间隔最大为准则,在每个决策节点将多类训练样本划分为两类进行训练,使两个子类间的可分性尽可能强,以构造合理的树结构,最终生成最优(或近优)的决策树.

4.1 遗传算法的设计

应用 GA 在每个决策节点将原始训练样本以最优(或近优)原则划分为两类问题,必须考虑如下问题:1) 参数编码;2) 初始种群的产生;3) 适应度函数的确定;4) 遗传操作.

4.1.1 实值编码策略

为了提高搜索效率,本文采用实值编码的策略实现原始训练样本集类别的编码,编码形式为: {类别编号 / 标识码}. 整个染色体分为两个基因段,其中“类别编号”基因段编码为当前节点类别号的一个随机排列,其长度为当前节点的类别数 L ;“标识码”基因段编码为 1 到 L 之间的一个随机整数,用以标注当前节点的分类策略.例如,决策树根节点的某一染色体编码为 {1, 2, ..., K/N }, 其中: $K \geq 3$ 为原始训练样本集类别总数,随机整数 $N \in [1, K]$ 表示在染色体第 N 位将训练样本分开.

对于根节点以下的子节点,根据其父节点的划分结果,将父节点染色体“类别编号”基因段中本节点不包含的类别所对应的基因剔除后,重新随机排列,并随机产生新的“标识码”,从而形成新的染色体.

4.1.2 初始种群的产生

采用实值编码,解空间与染色体空间重合.考虑到种群数目过大不仅增加了 GA 运算时间,而且会使种群形态过于分散,从而使算法收敛困难,所以本文选择种群规模的大小为训练样本的 30%. 在解空间中随机产生初始种群,并使其均匀分布于解空间.

4.1.3 适应度函数的确定

对决策树而言,越上层节点的分类性能对整个分类模型的推广性能影响越大.因此,在生成决策树的过程中,应使最容易分割的类最早分割出来,即在决策树的上层节点处分割,以减小“误差积累”的影响.根据前面的分析,SVM 的分类间隔可以作为两类间的可分离测度.因此,本文选取 SVM 分类间隔作为 GA 适应度函数,以使每个决策节点类间可分离度尽可能大.

根据 SVM 理论,GA 适应度函数设计为

$$\text{margin} = 2 / w^*, \quad (8)$$

其中

$$w^* = \sum_{i=1}^n y_i x_i, \quad (9)$$

y_i 为支持向量对应的 Lagrange 乘子.

4.1.4 遗传操作算子

遗传操作是实现寻优的关键,通常包括选择、交叉和变异操作算子.为使染色体完整地包含当前决策节点训练样本所有的类别种类,避免一条染色体上“类别编号”基因段的基因出现重复(该情况会使划分出现交叠),本文设计的遗传操作算子如下.

1) 选择.选择算子的作用是从上一代的遗传结果中以一定概率繁殖适应度较大的个体进入下一代的遗传操作.若个体 i 的适应度函数为 $\text{fit}(i)$,种群规模为 pop_size ,则选中 a_i 为下一代个体的概率为

$$P(a_i) = \text{fit}(a_i) / \sum_{j=1}^{\text{pop_size}} \text{fit}(a_j). \quad (10)$$

显然适应度高的个体,繁殖下一代的数目较多;而适应度较小的个体,繁殖的数目较小,甚至被淘汰。

2) 交叉. 交叉操作是指两个相互配对的染色体按某种方式相互交换其部分基因. 本文先采用随机的方式将种群中的个体两两配对, 然后在每对个体中随机设置一个交叉点, 进行单点交叉运算. 特别地, 为防止划分出现基因交叠, 在交叉操作后必须对“类别编号”基因段进行完备性检测: 若交叉操作后“类别编号”基因段仍保持当前决策节点完备的基因, 则接受本次交叉操作; 否则放弃本次交叉操作, 返回原染色体。

3) 变异. 为了在遗传操作初期取得较大的变异算子以维持种群的多样性, 防止出现早熟现象, 在算法已接近最优解邻域时, 减小变异算子, 以确保其局部搜索能力. 本文采用如下自适应变异概率:

$$P_m = \exp(-1.5 \times 0.5t) / \text{pop_size} \times \sqrt{L+1}. \quad (11)$$

其中: t 是进化代数, $L+1$ 是染色体长度。

同样, 为防止“类别编号”基因段上的基因出现重复, 当“类别编号”基因段上某一基因座上的基因发生变异时, 变异后的基因编码所对应的基因座上的基因要相应地变换为变异基因座上的原基因编码. 例如, 8 分类问题的某一原始染色体“类别编号”基因段编码如图 3(a) 所示, 若第 2 基因座上的基因发生变异, 变异后的基因编码为 7, 则最终变异操作得到的新染色体“类别编号”基因段编码如图 3(b) 所示。

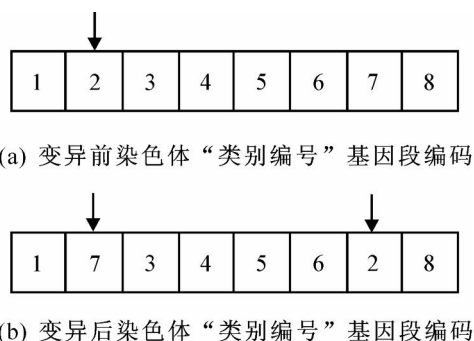


图3 变异操作示意图

4.2 决策树生成算法

基于遗传算法的 SVM 决策树生成算法可描述如下:

Step1: 将全部训练样本集所属类别按实值编码策略进行编码, 并在根节点调用 GA 得到最优或近优的染色体, 根据染色体的“标识码”值, 将“类别编号”基因所对应的原始训练样本类划分为两类训练基本 SVM.

Step2: 判断各子节点是否只包含一类样本, 若是转 Step4; 反之转 Step3.

Step3: 对包含两类以上样本的子节点, 按照编码策略形成新的染色体, 并调用 GA 得到子节点最优或近优的染色体, 根据染色体的“标识码”值, 将“类别编号”基因所对应的样本类划分为两类训练基本 SVM. 转 Step2.

Step4: 结束循环, 生成最优(或近优)决策树。

与 DAG-SVM 和 DT-SVM 不同, 上述算法在每一决策节点采取“一对一”、“一对多”还是“多对多”的训练策略由算法自己决定, 直到所有的节点被划分为叶子节点. 因为在每一决策节点引入了 GA 进行决策优化, 对于不同的分类问题将生成不同的决策树结构, 所以决策树的生成具有自适应性。

需要说明的是, GA 是一种启发式搜索算法, 因此并不能保证本文方法在任何情况下生成的决策树一定是最优的, 有可能是近优的. 另外, 由于 GA 的引入, 本文方法需训练的 SVM 数量会有所增加, 相应的训练时间也会有所延长. 但对于具体的分类问题, 决策树结构的优化和训练过程通常是离线进行的, 以增加离线优化时间为代价换取 SVM 决策树在线的最优(或近优)分类性能是可行且合理的。

5 仿真实验

本节通过实验对 DAG-SVM, DT-SVM 及本文提出的基于遗传算法决策树优化的 SVM (GADT-

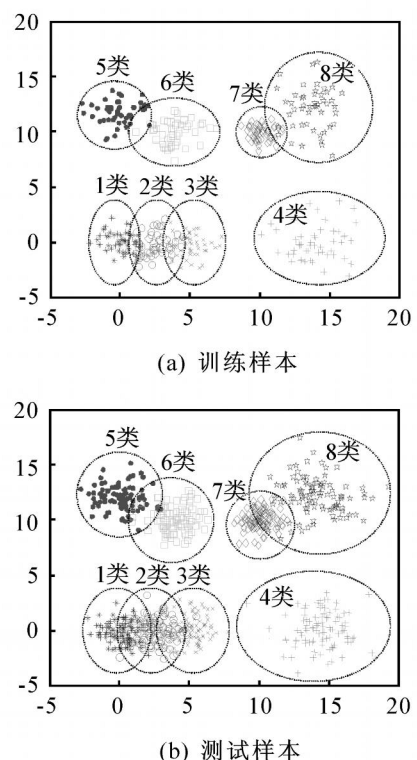


图4 样本分布图

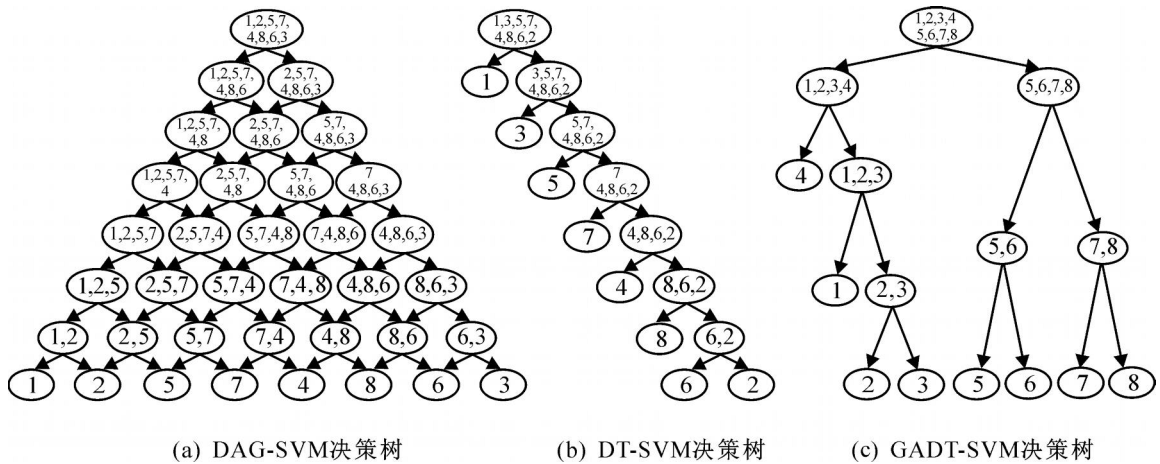


图 5 DAG-SVM,DT-SVM,GADT-SVM 决策树结构图

SVM) 多分类方法进行对比. 实验中随机产生 8 类分别服从二维正态分布 $N(0,0,1,1,0)$, $N(2.5,0,1,1,0)$, $N(5,0,1,1,0)$, $N(14,0,2.5,2.5,0)$, $N(0,12,1.2,1.2,0)$, $N(4,10,1.2,1.2,0)$, $N(10,10,0.6,0.6,0)$ 和 $N(14,12,2.5,2.5,0)$ 的样本进行训练和测试. 其中取每类 50 组样本用于训练,100 组样本用于测试, 训练和测试样本的分布分别如图 4(a),4(b) 所示.

为了比较 3 种方法抗“误差积累”的能力,在实验设计中将 8 类样本构成 4 个模糊组. 其中:1 ~ 3 类样本服从标准二维正态分布,类间间隔很小,交叠严重,构成一个模糊组;第 4 类样本均值为 (14,0), 方差为 (2.5,2.5), 样本散布较大,单独构成一个模糊组;第 5 类和第 6 类样本的方差均为 (1.2,1.2), 构成第 3 个模糊组;第 7 类和第 8 类样本分别服从方差为 (0.6,0.6), (2.5,2.5) 的正态分布,形成第 4 个模糊组. 从整体来看,第 1,2 模糊组和第 3,4 模糊组又可构成规模更大的两个模糊组. 为便于比较,所有的基本两分类 SVM 均采用径向基核函数

$$K(x, x_i) = \exp\{-\|x - x_i\|^2 / (2\sigma^2)\}, \quad \sigma = 5;$$

惩罚参数取 $C=2$.

DAG-SVM 和 DT-SVM 中的决策节点可以任意选择. 实验中采用的 DAG-SVM 和 DT-SVM 的结构如图 5(a),5(b) 所示;图 5(c) 是经本文算法优化后生成的决策树结构. 由图 5 可见,本文算法很好地将 8 类样本按其所属模糊组进行了划分.

3 种方法中,DAG-SVM 共需训练 28 个基本两分类 SVM;DT-SVM 仅需训练 7 个基本两分类 SVM;GADT-SVM 经优化后的决策树也仅包含 7 个基本两分类 SVM. 3 种方法对测试样本的分类结果如表 1~表 3 所示.

表中每一行对应一类样本,每一列显示样本被划分为某一类别的次数. 例如,表 1 第 1 行显示第 1

表1 DAG-SVM 分类结果

	1	2	3	4	5	6	7	8
1	93	7						
2	13	73						14
3			96					4
4				99				1
5					100			
6				6		94		
7			4				96	
8		11						89
分类正确率92.5%								

表2 DT-SVM 分类结果

	1	2	3	4	5	6	7	8
1	93							7
2		93						7
3			96					4
4				99			1	
5		1			99			
6				8		92		
7			4				96	
8	13	15						72
分类正确率92.5%								

表3 GADT-SVM 分类结果

	1	2	3	4	5	6	7	8
1	93	7						
2	13	73	14					
3			11	89				
4				100				
5					96	4		
6					4	96		
7							99	1
8							6	94
分类正确率92.5%								

类样本被正确分类为类别 1 的次数是 93 次,被误分为类别 2 的次数是 7 次.表中错分样本数用黑体显示,阴影区代表样本所属的模糊组.

一般而言,算法的分类质量可以从两方面进行评价:一是分类的正确率,描述算法正确分类的性能;二是误差的散布,反映算法对分类错误的控制能力.由表可见,虽然 3 种方法对测试样本的分类正确率均达到了 92.5%,但是分类质量却大不相同.由于没有进行决策优化,受“误差积累”的影响,DA G-SVM 和 DT-SVM 的分类误差散布很大,往往远远偏离原样本所属模糊组(脱离阴影区);而本文提出的 GADT-SVM 由于在每一节点进行了决策优化,可以根据具体分类问题自适应地选择最优(或次优)的决策树结构,从而大大减小了“误差积累”的影响.由表 3 可见,其分类误差很好地控制在相应的模糊组内部(错分样本均在阴影区内部),分类质量明显优于 DA G-SVM 和 DT-SVM. GADT-SVM 的这种特性在诸如医学诊断、设备故障诊断等不仅追求分类正确率,而且对分类误差也有很严格要求的应用场合具有重要意义.

6 结 论

本文提出了一种基于 GA 的 SVM 多分类决策树优化算法.针对传统的 DA G-SVM 和 DT-SVM 决策树结构固定,“误差积累”现象严重等不足,利用 GA 在每一决策节点选择最优的分类决策,最终自适应地生成最优或近优决策树.文中针对 SVM 多分类问题的特点,设计了 GA 的编码策略、适应度函数及遗传算子;在此基础上,提出了最优(近优)SVM 决策树的生成算法;最后进行了仿真实验.实验结果表明,由于采用了优化的决策树结构,本文方法可大大减小“误差积累”的影响,分类质量明显优于 DT-SVM 和 DA G-SVM 方法.

参考文献(References)

- [1] Weston J, Watkins C. Modeling multi-class support vector machines[R]. London: University of London, 1998: 1-10.
- [2] Hsu C W, Lin C J. A comparison of methods for multi-class support vector machine[J]. IEEE Trans on Neural Networks, 2002, 13(2): 415-425.
- [3] Krebel U. Pairwise classification and support vector machines[M]. MA: MIT Press, 1999: 255-268.
- [4] Bottou L, Cortes C, Denker J, et al. Comparison of classifier methods: A case study in handwriting digit recognition [C]. Proc of Int Conf on Pattern Recognition. Jerusalem, 1994: 77-87.
- [5] Fumitake T, Shigeo A. Decision tree-based multiclass support vector machines[C]. Proc of the 9th Int Conf on Neural Information Processing. Singapore, 2002: 1418-1422.
- [6] Platt J C, Cristianini N, Shawe Taylor J. Large margin DAGs for multiclass classification [J]. Advances in Neural Information Processing Systems, 2000, 12(3): 547-553.
- [7] Takahashi F, Abe S. Decision-tree-based multiclass support vector machines[C]. Proc of the 9th Int Conf on Neural Information Processing. Singapore, 2002, (3): 1418-1422.
- [8] Sungmoon C, Sang H O, Soor-Young L. Support vector machines with binary tree architecture for multi-class classification [J]. Neural Information Processing — Letters and Reviews, 2004, 2(3): 47-51.
- [9] 马笑啸,黄席樾,柴毅.基于 SVM 的二叉树多类分类算法及其在故障诊断中的应用[J].控制与决策,2003,18(3): 272-276.
(Ma X X, Huang X Y, Chai Y. 2PTMC classification algorithm based on support vector machines and its application to fault diagnosis[J]. Control and Decision, 2003, 18(3): 272-276.)
- [10] 郑春红,焦季成,郑贵文.基于 GA 的遥感图像目标 SVM 自动识别[J].控制与决策,2005,20(11): 1212-1215.
(Zheng C H, Jiao L C, Zheng G W. Genetic algorithm-based SVM for automatic target classification of remote sensing images[J]. Control and Decision, 2005, 20(11): 1212-1215.)
- [11] 杨宏辉,孙进才,袁骏.基于支持向量机和遗传算法的水下目标特征选择算法[J].西北工业大学学报,2005,23(4): 512-515.
(Yang H H, Sun J C, Yuan J. A new method for feature selection for underwater acoustic targets[J]. J of Northwestern Polytechnical University, 2005, 23(4): 512-515.)
- [12] 李杰,楚恒,朱维乐,等.基于支持向量机和遗传算法的纹理识别[J].四川大学学报,2005,37(4): 104-108.
(Li J, Chu H, Zhu W L, et al. Texture recognition using support vector machines and genetic algorithm [J]. J of Sichuan University, 2005, 37(4): 104-108.)
- [13] 孟媛媛,刘希玉.一种新的基于二叉树的 SVM 多类分类方法[J].计算机应用,2005,25(11): 2653-2657.
(Meng Y Y, Liu X Y. A new SVM multiclass classification based on binary tree [J]. Computer Applications, 2005, 25(11): 2653-2657.)