

多类支持向量机方法的研究现状与分析

赵春晖,陈万海,郭春燕

(哈尔滨工程大学 信息与通信工程学院,黑龙江 哈尔滨 150001)

摘要:支持向量机(SVM)是建立在统计学理论基础上的小样本机器学习方法,最初应用于解决两类分类问题。然而在解决实际问题中遇到的多为多分类问题,如何有效的将其推广到多类分类问题是一个正在研究的问题。该文对现有的多类支持向量机方法从组合多个两类分类器、层次结构、一次性优化问题和纠错编码等4个角度进行了综合归纳和分析,详细介绍了每种方法的代表性算法,并比较其优劣。

关键词:多类支持向量机;两类分类器;层次结构;一次性优化;纠错编码

中图分类号:TP391 **文献标识码:**A **文章编号:**1673-4785(2007)02-0011-07

Research and analysis of methods for multiclass support vector machines

ZHAO Chun-hui, CHEN Wan-hai, GUO Chun-yan

(College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: The SVM is a limited sample learning method which was developed from statistical theory, and originally designed for binary classification. However, many practical problems are multi-classification ones. How to effectively extend binary classification to multi-classification is an ongoing research issue. This paper generalizes and analyzes multiclass support vector machines from four angles: combination of several binary classifiers, hierarchical structures, one-off optimization and error correcting codes. Several representative algorithms for various methods are introduced in detail and their advantages and disadvantages are compared.

Key words: multiclass support vector machine; binary classifier; hierarchical structure; one-off optimization; error correcting code

1992~1995年,Vapnik在统计学习理论的基础上发展出了一种新的模式识别方法—支持向量机(support vector machine, SVM)^[1],它采用结构风险最小化原则代替了传统机器学习方法中的经验风险最小化原则,在解决小样本、非线性及高维模式识别问题中表现出许多特有的优势。支持向量机的基本思想可概括为:寻找一个最优分类超平面,使得训练样本中的2类样本点能被无错误的分开,并且要使2类的分类间隔最大;而对线性不可分问题,通过核函数将低维输入空间的数据映射到高维空间,从而将原低维空间的线性不可分问题转化为高维空间

上的线性可分问题,然后在这个新空间中求取最优分类面。

最初支持向量机是用于解决2类分类问题,不能直接用于多类分类。而实际应用中遇到的多为多分类问题,如何有效地将其推广到多类分类问题仍是当前支持向量机研究的重要内容之一,文中把支持向量机应用到多类分类问题中的算法统称为多类支持向量机。目前人们对多类支持向量机的研究还有待于完善,但也取得了一定的成就。至今已经有几种卓有成效的方法将SVM推广到多类分类问题^[2-5]。下面根据多类支持向量机的不同实现方法把它分成4大类来详细介绍,并比较其优劣,以便读者在今后的研究中尽量避免这些方法所存在的各种弊端,发展出更多有效的多类支持向量机算法。

收稿日期:2006-11-14.

基金项目:国家自然科学基金资助项目(60672034);高校博士点基金资助项目(20060217021);黑龙江省自然科学基金资助项目(ZJ G20606-01)。

1 用多个两类分类器实现多类分类

1.1 1-a-r SVM

1-a-r SVM (one-against-rest)^{[1][6]} 算法是用支持向量机解决多类分类问题的最早的方法,对于 k ($k \geq 2$) 类 SVM 分类问题,构造 k 个两类分类器,自然地,将 k 分类问题转化为 k 个两类 SVM 分类问题.第 i 个 SVM 分类问题是把第 i 类作为一类,其余 $k-1$ 类视为另一类.设有 n 个训练数据为 $(x_1, y_1), \dots, (x_n, y_n)$, 其中 $x_i \in R^D, y_i \in \{1, \dots, k\}$ 为 x_i 的类别标号, $i=1, 2, \dots, n$. 第 i 个 SVM 需要解决下面的最优化问题:

$$\begin{aligned} \min_{w^i, b^i} \quad & \frac{1}{2} (w^i)^T w^i + C \sum_{j=1}^n \xi_j (w^i)^T, \\ (w^i)^T \phi(x_i) + b^i \quad & \geq 1 - \xi_j, \text{ if } y_j = i, \\ (w^i)^T \phi(x_i) + b^i \quad & \leq -1 + \xi_j, \text{ if } y_j \neq i, \\ \xi_j \quad & \geq 0, j = 1, \dots, n. \end{aligned} \tag{1}$$

求解式(1)得到 k 个决策函数:

$$(w^1)^T \phi(x) + b^1,$$

...

$$(w^k)^T \phi(x) + b^k.$$

测试时 x 属于决策函数输出值最大的那一类:

$$\text{class of } x = \arg \max_{i=1, \dots, k} ((w^i)^T \phi(x) + b^i). \tag{2}$$

这种方法简单、有效,训练时间较短,可用于大规模数据.但其缺点是当类别数较大时,某一类的训练样本将大大少于其他类的训练样本的总和,这种训练样本的不均衡将对精度产生影响,且存在误分、拒分区域.

1.2 1-a-1 SVM

1-a-1 SVM 算法 (one-against-one) 是在每 2 类之间训练一个分类器,因此对于一个 k 类问题,训练阶段共构造 C_k^2 个两类分类器,每个分类器是取任意 2 个类别的数据进行训练.对于第 i 类和第 j 类之间的训练,需要解决下面的两类分类问题:

$$\begin{aligned} \min_{w^{ij}, b^{ij}} \quad & \frac{1}{2} (w^{ij})^T w^{ij} + C \sum_t \xi_t (w^{ij})^T, \\ (w^{ij})^T \phi(x_t) + b^{ij} \quad & \geq 1 - \xi_t, \text{ if } y_t = i; \\ (w^{ij})^T \phi(x_t) + b^{ij} \quad & \leq -1 + \xi_t, \text{ if } y_t = j. \\ \xi_t \quad & \geq 0 \end{aligned} \tag{3}$$

在测试阶段有不同的方法确定样本属于哪一类,最常用的一种方法是“最大投票法”,即每个两类分类器都对样本的类别进行判断,采用投票机制为其相应的类别投上一票,最后得票最多的类即是该未知样本的所属类.

对未知样本 x 进行分类时,判断符号函数

$$f(x) = \text{sign}((w^j)^T \phi(x) + b^j), \tag{4}$$

若 x 属于第 i 类则第 i 类的票数加 1,反之第 j 类加 1. x 属于最后票数最多的那一类.

由于每个 SVM 只考虑 2 类样本,故单个 SVM 容易训练;另外,虽然它的复杂度以类数按平方增长,但就分类速度来说,其并不比传统的 1-a-r SVM 方法慢;而且分类精度也比 1-a-r SVM 高.

但是这种方法也存在缺点:分类器的数目随类数急剧增加,导致决策时速度降低,在测试阶段,采用最大投票法时,如果 2 个类别的票数一致时则产生不可分情况,便存在误分、拒分区域.

2 用层次型两类分类器实现多类分类

这种组合形式也是建立在两类分类器的基础上的,与上面介绍的方法的不同之处在于,这种方法的测试阶段都是采用层次结构即树形结构,然而在每个节点上仍用两类 SVM 进行分类,因此文中称之为用层次型两类分类器实现多类分类.

2.1 有向无环图 SVM

有向无环图 SVM (directed acyclic graph SVM, DAG-SVM)^[1,6-7] 是 Platt^[10] 针对 1-a-1 SVM 存在的误分拒分现象提出的一种新算法.这种方法在训练阶段采用 1-a-1 SVM 的任意两两组合训练方式,同样也需要构造 C_k^2 个子分类器,但是在分类决策过程中,DAG 将所用子分类器构造成有向无环图如图 1 所示.

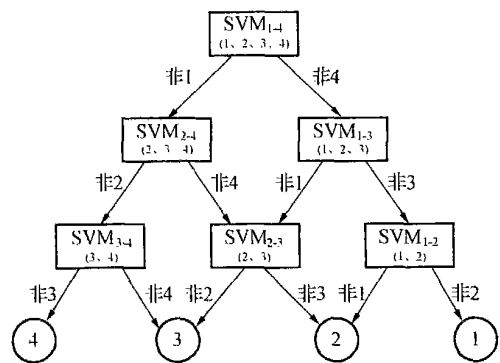


图 1 DAG-SVM 4 分类的树权结构

Fig. 1 Tree structure of DAG-SVM four-classification

它包括 C_k^2 个内部节点和 k 个叶子节点,其中每个内部节点都是一个两类分类的子分类器,并与下一层的 2 个内部节点相连.当对未知样本分类时,首先从根节点开始,根据根节点的分类结果用下一层中的左节点或者右节点继续分类,直到达到底层某个叶节点为止,该叶子节点就是未知样本所属的类

别。

DAG-SVM 简单易行,每个未知样本只需使用 $k-1$ 个决策函数即可得出结果,较“1-a-1 SVM”方法提高了测试速度,而且不存在拒分区域。

这种方法有 2 个主要的缺点:一个是存在自上而下的“误差累积”现象。这是层次结构的固有弊端,故 DAG-SVM 也逃脱不掉,即如果在某个结点上发生分类错误,则会把分类错误延续到该结点的后续结点上。若采用该方法处理一个 20 类的问题,每个未知样本所属的真实类别需要和其他类别经过 19 次两类测试才能得到最后结果。假设每个两类分类器的错分概率为 1%,那么该未知样本累积的错分的概率为 $1 - 0.99^{19} = 17.38\%$,由此可以看出累积误差随着类别数目的增加而增加。而且分类错误在越靠近根结点的地方发生,误差的累积效应越明显,分类性能就越差,尤其在根结点上发生分类错误,将严重影响分类性能。另一个缺点是该方法的最后输出结果对节点的排列顺序的依赖性很大,用一个二维的 4 类分类问题的图例来说明一下这个问题,如图 2 所示。

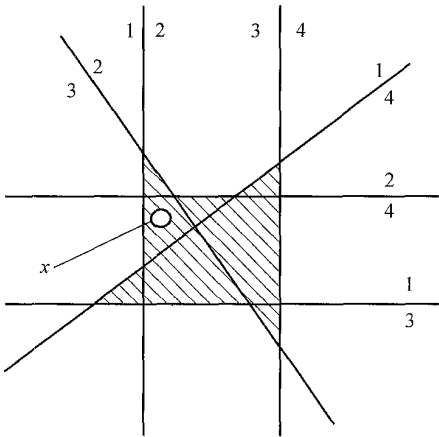


图 2 DAG 输出取决于不同节点序列

Fig.2 DAG output depends on different node sequence

图 2 中‘1’,‘2’,‘3’和‘4’是类别符号,图中阴影部分的数据最后的输出类别取决于 DAG 不同的节点序列,节点的排序不同,未知样本的输出结果也会不一样。以图中样本点 x 为例,当节点的序列为 1-2-3-4 时,首先 1vs4 分类超平面会把类 4 给否决掉,因为样本点 x 不在类 4 的那一边;然后序列变成 1-2-3,1vs3 超平面把类 3 从序列中排除掉;最后序列变为 1-2,1vs2 的超平面会把样本点 x 的类别定为第 2 类。

改变一下节点的排列顺序,则会得到另外一种结果。若序列为 2-1-3-4,首先 2vs4 的分类超平

面便会把第 2 类排除掉;然后序列更新为 1-3-4,1vs4 的超平面则把第 4 类从序列中去掉;最后序列变为 1-3,1vs3 的超平面把样本点 x 归属于第 1 类。这个例子说明 DAG 中不同的节点顺序会导致部分样本的分类结果不同,从而降低了分类精度。

2.2 自适应 DAG-SVM

针对 DAG 存在的误差累积效应,文献[2]提出了一种自适应 DAG-SVM (adaptive directed acyclic graph SVM, ADAG-SVM) 方法。该方法是在 DAG 的基础上改进的,是倒三角结构的 DAG。

对于 k 类分类问题,在训练阶段仍采用 1-a-1 SVM 的任意两两组合训练方式,构造 C_k^2 个子分类器。测试阶段,ADAG 包括 $k-1$ 个内部节点,其中每个内部节点都是一个两类分类的子分类器,节点的排列是一个倒三角的结构,在最顶层有 $k/2$ 个节点,第 2 层有 $k/2^2$ 个节点,以此类推,最底层就是最后的输出结果即叶节点,其结构如图 3 所示。

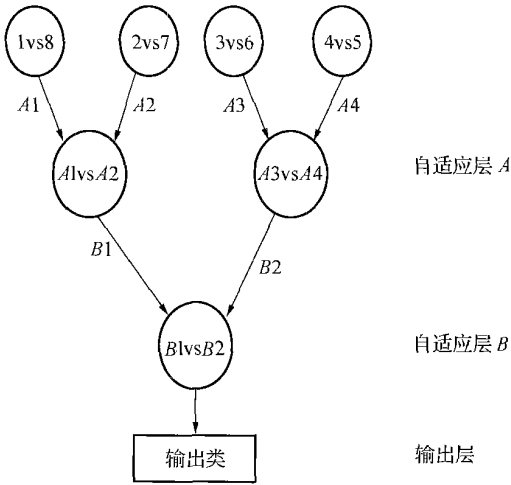


图 3 8 类 ADAG 分类结构图

Fig.3 Framework of eight sorts of ADAG classification

用 ADAG 进行分类,顶层有 $k/2$ 个两类分类器,每个分类器输出的结果参与下一层的分类,在第 2 层分类时,分类器减少到 $k/2^2$ 个,以此类推,每循环一层分类器将会减少一半,直到达到最底层。和 DAG 算法一样,未知样本也需要计算 $k-1$ 个决策函数得到最后的结果。但是未知样本的真实类别只需要和其他类别计算 $\log_2 k$ 次或更少,和 DAG 的 $k-1$ 次相比减小了很多,这样就可以在很大程度上减少了误差的累积。

2.3 重新排序 ADAG-SVM

针对 DAG 的节点序列的依赖性,文献[3]在文献[2]的基础上提出了一种重新排序的 ADAG (re-ordering adaptive directed acyclic graph SVM,

RADAG-SVM)方法,该方法是根据权系数向量为未知样本寻找最佳的节点序列,以克服 DAG 中节点序列产生的分类误差。

支持向量机分类原理是寻找最优分类面,不但能将两类无错误的分开,而且要使两类的分类间隔最大,分类间隔就是两类样本中离分类面最近的两个样本之间的距离即 $2/w$ 。也就是说权系数向量 w 越小,分类间隔越大,两类样本之间的可分性就越大。因此 w 的值在一定程度上会影响数据的分类精度,文献[3]把这一现象应用到分类节点的选择上以解决 DAG 中节点排序不同而产生分类误差的问题。

对于 k 类分类问题,在训练阶段仍采用 1-a-1 SVM 的任意两两组合训练方式,构造 C_k^2 个两类分类器。测试阶段和 ADAG 一样包括 $k-1$ 个内部节点,每个节点都会否定一个类别,不同之处在于节点序列的初始化以及每一层节点的排列顺序。

首先,利用权向量最小值的最佳匹配算法把个类别按最佳排列顺序初始化为一个列表;第2步,顶层分类的每个节点的类别组合是根据初始化列表进行排列的,列表中的第一个类别和最后一个类别组合成第一个决策节点,第2个类别和列表中倒数第2个类别组合成第2个决策节点,以此类推;第3步,顶层分类器的输出结果排除了一部分类别,剩下的类别将参与第2层分类。在第2层分类之前,仍需要利用权向量最小值的最佳匹配算法把这些类别重新排序。不同的测试样本在这一层会得到不同的序列,这依赖于上一层分类的结果。重复第2步和第3步直到最后一层只剩下一个类别。RADAG 分类结构如图4所示。

权向量最小值的最佳匹配算法是根据不同的 w 值选择最佳的两类构造子分类器,以减少错分样本提高分类精度。 w 值越小两类样本之间的可分性就越大。如图5(a)所示,每两个类别构成的分类器都有存在一个 w 值。权向量最小值的最佳匹配算法就是选择一个权向量和最小的一个子集来构造一系列的两类分类器如图5(b)。

2.4 二叉树多级 SVM

针对 1-a-r SVM 和 1-a-1 SVM 中存在的无法识别的区域,文献[4]提出了二叉树多级 SVM (binary tree multistage support vector machine, BT-MSVM)^[4,8]。根据其层次结构的形态不同该方法又可分为2种情况:1)从顶层开始,每一个包含多个类别的节点上的分类器只将一个类别与其他类别分开,这样的结构称之为“偏态树”;2)从顶层开始,每

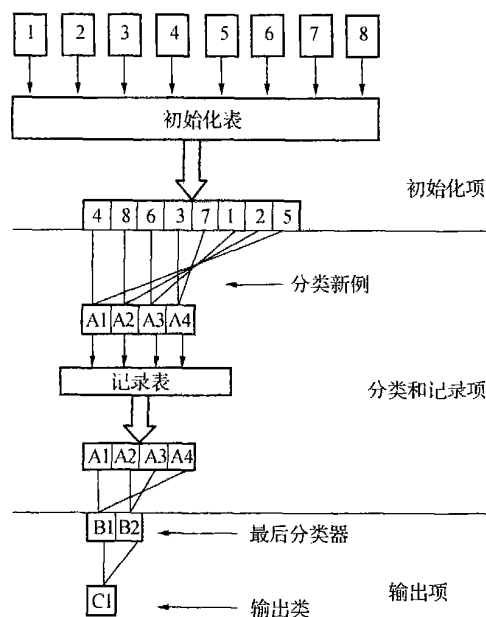


图4 8类 RADAG 分类结构图

Fig. 4 Framework of eight sorts of RADAG classification

一个包含多个类别的节点上的分类器都将其中类别均分为2类,这样的结构称为“正态树”。下面主要介绍一下偏态树的分类原理。

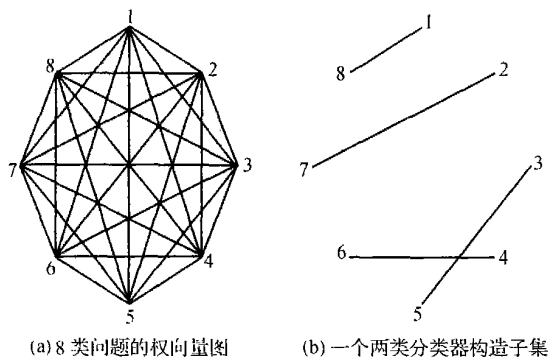


图5 8类问题的权向量图和其中一个两类分类器构造子集
Fig. 5 Weight vector of eight sorts of problems and Structuring subset of a two-classification classifier

给定一个由 n 个学习样本组成的 k 类分类问题,学习样本为 (x_i, y_i) , 其中 $x_i \in R^D, i=1, 2, \dots, n$, $y_i \in \{1, \dots, k\}$ 为 x_i 的类别标号,设 m 第类学习样本数目为 n_m 。该算法的核心思想是在训练阶段构造 $k-1$ 个两类 SVM, 其中第 m 个 ($m=1, 2, \dots, k-1$) 两类 SVM 的学习样本仅有分类标号 $y_i = m$ 的样本子集组成,且将原第 m 类的样本的分类标号改为 $+1$, 其余 $k-m$ 类的样本的分类标号改为 -1 , 即

$$y_i = \begin{cases} +1, & (y_i = m), \\ -1, & (y_i > m). \end{cases} \quad (5)$$

1) 训练阶段:

第 1 步:令 $m = 0$, 选择合适的核函数 $K(x, x_i)$ 及有关参数;

第 2 步:令 $m = m + 1$;

第 3 步:构造第 m 个两类 SVM 的学习样本集合 $L_m: (x_i^m, y_i^m), x_i \in R^D, i = 1, 2, \dots, n^m, y_i \in \{+1, -1\}, n^m = \sum_{j=m}^k n_j$, 针对学习样本 L_m 求解两类分类问题;

第 4 步:建立 L_m 的最优决策超平面:

$$f^m(x) = \operatorname{sgn}\left(\sum_{i=1}^n y_i^m K(x_i^m \cdot x) + b^m\right). \quad (6)$$

第 5 步:若 $m = k - 1$, 则训练结束; 否则转至第 2 步继续训练.

2) 测试阶段

根据式(6) 计算未知样本在每个两类 SVM 中 $f^m(x) (m = 1, 2, \dots, k - 1)$ 的决策输出值, 并按式(7) 做出分类决策:

$$f(x) = \begin{cases} m, & \text{若 } f^1(x) = f^2(x) = \dots = f^{m-1}(x) = -1 \text{ 且 } f^m(x) = +1, \\ k, & \text{若 } f^1(x) = f^2(x) = \dots = f^{k-1}(x) = -1. \end{cases} \quad (7)$$

BTM-SVM 的分类结构图如图 6 所示.

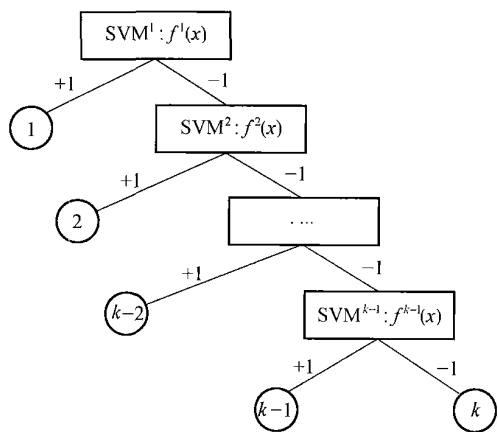


图 6 类二叉树多级 SVM 分类结构

Fig. 6 Structure of k -sort two-branch multiclass SVM classification

3 用一个最优化问题一次性实现多类分类

这种方法和上述的方法的不同之处在于它是采用将多个分类面的参数求解合并到一个最优化问题中,通过求解该最优化问题“一次性”的实现多类分

类^[1,6]. 该方法类似于 1-a-r SVM, 针对 k 类问题, 需要构造 k 个两类分类函数, 其中第 m 个函数 $w_m^T \phi(x) + b$ 是用来把第 m 类与其他类别区分开的. 这 k 个分类函数用一个最优化问题来解决, 公式描述如下:

$$\min_{w, b} \frac{1}{2} \sum_{m=1}^k w_m^T w_m + C \sum_{i=1}^l \sum_{y_i}^m w_{y_i}^T \phi(x_i) + b_{y_i} - w_m^T \phi(x_i) + b_m + 2 - \frac{m}{i}, \quad m = 0, i = 1, \dots, l, m \in \{1, \dots, k\} \setminus y_i. \quad (8)$$

求解式(8) 得到的决策函数为

$$\arg \max_{m=1, 2, \dots, k} (w_m^T \phi(x) + b_m). \quad (9)$$

该方法尽管看起来简洁, 但是在最优化问题的求解过程中变量过多, 选择的目标函数过于复杂, 从而导致它的计算复杂度高, 分类精度上也不占优势. 当训练样本数目非常大的时候, 这一问题更加突出. 因此, 这一类方法在实际应用中并不常用.

4 纠错编码 SVM

该方法是对类别进行二进制编码将多类问题转化为多个两类问题. 采用具有纠错能力的编码对类别进行编码, 并将 SVM 作为码位分类器, 这种多类分类方法称之为纠错编码支持向量机 (error correcting codes, ECC-SVM)^[7].

对于 k 类分类问题, 给每一个类别赋予长度为 L 的二进制编码, 形成一个 k 行 L 列的码本 (如表 1). 对于其中第 $i \in \{1, 2, \dots, L\}$ 列, 将码字为“0”的所有类别作为一类, 码字为“1”的类别作为另一类, 因此在每个码位上对应着一个两类分类问题. 这样就将 k 类分类问题转化为 L 个两类分类问题. 当对一个未知样本分类时, L 个 SVM 分类器的分类结果构成一个编码 s , 然后计算码本中 k 个标准编码与 s 之间的汉明距离, 距离最小者所代表的类别即为该样本所属的类别.

Dietterich 和 Bakiri^[7] 给出了一个好的纠错码应该满足的条件:

- 1) 行 $r_i (i = 1, \dots, k)$ 与行 $r_j (j = 1, \dots, k; j \neq i)$ 之间不相关;
- 2) 列 $c_i (i = 1, \dots, L)$ 与列 $c_j (j = 1, \dots, L; j \neq i)$ 及其补 \bar{c}_j 之间不相关;
- 3) 没有全为“0”或全为“1”的列.

根据上述条件对于一个 k 类分类问题, 纠错码长度 L 的取值范围是 $\log_2 k, 2^{k-1} - 1$. 当 $L = \log_2 k$ 时, 编码失去纠错功能.

表1 8类分类问题的 $L=5$ 的纠错编码
Table 1 Error-correcting codes of eight sorts of problems for $L=5$

类别 编号	码 字				
	1	2	3	4	5
1	0	0	0	1	1
2	0	0	1	0	1
3	0	1	0	0	1
4	0	1	1	0	0
5	1	0	0	1	0
6	1	1	0	1	1
7	1	1	1	0	1
8	1	1	1	1	0

纠错编码支持向量机在分类过程中所需分类器的个数等于纠错编码的位数 L , 当类别大的时候仅需要较少的分类器, 但是如何根据具体问题确定码本、选择排列顺序以达到最优的分类性能依然有待研究。

5 结束语

该文总结了现有多类支持向量机的4种主要结构形式, 并详细介绍了其代表性算法, 对各种方法进行了性能优劣的比较。用多个两类分类器实现多类分类是最早发展起来的多类支持向量机, 其代表算法是 $1-a-r$ SVM 和 $1-a-1$ SVM。和 $1-a-r$ SVM 相比, $1-a-1$ SVM 不仅提高了训练速度而且改善了误分、拒分区域范围, 但在采用投票机制进行决策时仍存在一些无法分类的样本。当类别数目过大时, 这种方法的训练速度和分类速度都会大幅度的降低, 因此比较适用于小类别的分类研究。

经过进一步研究, 人们提出了几种层次结构的 SVM, 与第一种方法不同之处在于, 这种方法的测试阶段采用树形结构, 在每个节点上仍用两类 SVM 进行分类, 其代表算法有 DAG-SVM 和 BTM-SVM 等。这种方法简单易行, 在分类速度和精度上也有一定的改善, 对于一般规模的多分类问题, 是一种有效的方法。用一个最优化问题实现多分类问题的计算复杂度比较高, 因此一般实际应用中都不采用。纠错编码 SVM 在如何选择码本及排列顺序以达到最优的分类性能上还有待于进一步的理论研究。

随着研究的深入, 提出了模糊支持向量机, 它是针对 SVM 多类问题中存在不可分区域的现象提出的。当一些样本不能被确切地定义为属于某一类时, 文中所述的多类方法便会把此类样本硬性分配给某一类别, 而模糊支持向量机则是引入了模糊隶属度

函数, 提出了将多种分类方法结合到一起组成混合分类器的思想。由于实际应用中需要解决大量的多类别的分类问题, 如何有效地运用方法解决多类分类问题将会受到越来越多的重视。该文总结了现有的主要的多类支持向量机及其优缺点, 希望在这方面起到一点承前启后的作用, 以便读者更好地学习、掌握和运用多类支持向量机技术。

参考文献:

- [1] HSU C W, LIN C J. A comparison of methods for multi-class support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2): 415 - 425.
- [2] KIJ SIK KUL B, USSIVA KUL N. Multiclass support vector machines using adaptive directed acyclic graph [A]. Proceedings of the 2002 International Joint Conference on Neural Networks [C]. Honolulu, HI, USA, 2002, 1(5): 980 - 985.
- [3] PHETKAEW T, KIJ SIK KUL B, RIVEPIBOON W. Reordering adaptive directed acyclic graphs: an improved algorithm for multiclass support vector machines [A]. Proceedings of the 2003 International Joint Conference on Neural Networks[C]. Portland, OR, USA, 2003.
- [4] TIAN X, DENG F Q. An improved multi-class SVM algorithm and its application to the credit scoring model[A]. Proceedings of the fifth World Congress on Intelligent Control and Automation[C]. Hangzhou, China, 2004.
- [5] ANGUITA D, RIDELLA S, STERPID D. A new method for multiclass support vector machines[A]. Proceedings of the 2004 International Joint Conference on Neural Networks[C]. Budapest, Hungary, 2004.
- [6] 黄 勇, 郑春颖, 宋忠虎. 多类支持向量机算法综述[J]. 计算技术与自动化, 2005, 24(4): 61 - 63.
HUANG Yong, ZENG Chunyin, SONG Zhonghu. Multiclass support vector machines algorithm summarization [J]. Computing Technology and Automation, 2005, 24(4): 61 - 63.
- [7] 刘志刚, 李德仁, 秦前清, 史文中. 支持向量机在多类分类问题中的推广[J]. 计算机工程与应用, 2004(7): 10 - 13.
LIN Zhigang, LI Deren, QIN Qianqing, SHI Wenzhong. An analytical overview of methods for multi-category support vector machines[J]. Computer Engineering and Applications, 2004(7): 10 - 13.
- [8] TAKAHASHI F. Decision tree-based multi-class support vector machines[A]. Proceedings of the 9th International Conference on Neural Information Processing [C]. Orchid Country Club, Singapore, 2002.
- [9] PLATT J, CRISTIANINI N, SHAWE-TAYLOR J. Large margin DAGs for multiclass classification [A]. Proceedings of Neural Information Processing Systems [C]. [s.l.], Cambridge, 2000.

- [10] PAUL G. Cox, Reza Adhami. Multi-class support vector machine classifier applied to hyper-spectral data [A]. Proceedings of the Thirty-Fourth Southeastern Symposium on System Theory [C]. Huntsville, Alabama, 2002.
- [11] FRANC V, HLAVAC V. Multi-class support vector machine [A]. Proceedings of the 16th International Conference on Pattern Recognition [C]. Quebec City, Canada, 2002.
- [12] LIN Chunfu, WANG Shengde. Fuzzy support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2): 464 - 471.
- [13] LI Kunlun, HUANG Houkuan, TIAN Shengfeng. A novel multi-class SVM classifier based on DDAG[A]. Proceedings of the 2002 International Conference on Machine Learning and Cybernetics[C]. Beijing, China, 2002.
- [14] ARENAS-GARCIA J, PEREZ-CRUZ F. Multi-class support vector machines: a new approach[A]. Proceedings of the 2003 IEEE International Conference on Acoustics, Speech and Signal Processing [C]. Hong Kong, China, 2003.
- [15] LIU Yi, ZHENG Y F. One-against-all multi-Class SVM classification using reliability measures [A]. Proceedings of the 2005 International Joint Conference on Neural Networks[C]. Montreal, Canada, 2005.
- [16] LU Baoliang, WANG Kaian, UTIYAMA M. A part-versus-part method for massively parallel training of support vector machines[A]. Proceedings of the 2004 International Joint Conference on Neural Networks[C]. Budapest, Hungary, 2004.
- [17] XIN Dong, WU Zhaohui, PAN Yunhe. A new multi-class support vector machines[A]. IEEE International Conference on Systems, Man and Cybernetics[C]. Tucson, USA, 2001.
- [18] 董学锋, 石繁槐. FSVM 在有限集脱机手写体汉字识别中的应用[J]. 计算机工程, 2003, 29(3): 109 - 111. TONG Xuefeng, SHI Fanhuai. Application Of FSVM for limited-set off-line handwritten chinese characters recognition[J]. Computer Engineering, 2003, 29(3): 109 - 111.
- [19] 李昆仑, 黄厚宽, 田盛丰. 一种基于有向无环图的多类 SVM 分类器[J]. 模式识别与人工智能, 2003, 16(2): 164 - 168. LI Kunlun, HUANG Houkuan, TIAN Shengfeng. A novel multi-class SVM classifier based on DDAG[J]. Pattern Recognition and Artificial Intelligence, 2003, 16(2): 164 - 168.
- [20] 安金龙, 王正欧, 马振平. 一种新的支持向量机多类分类方法[J]. 信息与控制, 2004, 33(3): 262 - 267. AN Jinlong, WANG Zhengou, MA Zhenping. A new SVM multiclass classification method[J]. Information and Control, 2004, 33(3): 262 - 267.
- [21] 唐发明, 王仲东, 陈绵云. 支持向量机多类分类算法研究[J]. 控制与决策, 2005, 20(7): 746 - 749. TANG Faming, WANG Zhongdong, CHEN Mianyun. On multiclass classification methods for support vector machines [J]. Control and Decision, 2005, 20(7): 746 - 749.
- [22] 郑勇涛, 刘玉树. 支持向量机多分类问题研究[J]. 计算机工程与应用, 2005, 41(23): 190 - 192. ZHENG Yongtao, LIU Yushu. An analysis of multi-class support vector machines[J]. Computer Engineering and Applications, 2005, 41(23): 190 - 192.
- [23] 祁亨年. 支持向量机及其应用研究综述[J]. 计算机工程, 2004, 30(10): 6 - 9. QI Hengnian. Support vector machines and application research overview[J]. Computer Engineering, 2005, 41(23): 190 - 192.
- [24] 宋晓宁, 束鑫. 一种改进的模糊支持向量机的人脸识别方法[J]. 微机发展, 2005, 15(3): 23 - 25. SONG Xiaoning, SHU Xin. An improved fuzzy support vector machine and its application to face recognition[J]. Microcomputer Development, 2005, 15(3): 23 - 25.
- [25] 刘江华, 程君实, 陈佳品. 支持向量机训练算法综述[J]. 信息与控制, 2002, 31(1): 45 - 50. LIU JIanghua, CHENG Junshi, CHEN Jiapin. Support vector machine training algorithm: a review [J]. Information and Control, 2002, 31(1): 45 - 50.
- [26] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32 - 42. ZHANG Xuegong. Introduction to statistical learning theory and support vector machines[J]. Acta Automatica Sinica, 2000, 26(1): 32 - 42.

作者简介:



赵春晖,男,1965年生,博士,教授,博士生导师,主要研究方向为智能信息处理技术、图像处理。获省部级科技进步奖5项,发表论文180多篇,出版著作3部。
E-mail: zhaochunhui@hrbeu.edu.cn.



陈万海,男,1963年生,副教授,信号与信息处理专业博士研究生,主要研究方向为超光谱遥感图像处理技术,发表论文8篇。



郭春燕,女,1984年生,硕士研究生,主要研究方向为超光谱遥感图像处理技术、目标分类。