机器学习中的多侧面递进算法 MIDA

张燕平[1,2] ,张　铃[1,2] ,吴　涛[2]
(1. 安徽大学计算智能与信号处理实验室,安徽合肥 230039;2. 安徽大学人工智能研究所,安徽合肥 230039)

摘　要: 对高维海量数据,为解决准确率与泛化能力之间的矛盾,提出机器学习中的多侧面递进算法 MIDA (Multi-side Increase by Degrees Algorithm),该算法将样本集分成几个部分,对各部分分别选择一组适应它们的特征子集. 这种分而治之的方法,在保证一定的精度的前提下,符合人类对复杂问题的求解分重点,多方面考虑的方式,可有效地识别复杂问题的分类,提高泛化能力,降低了计算的复杂性. 本文利用覆盖算法给出具体的多侧面递进算法,并给出实验结果,实验结果表明新的方法是有效的.

关键词: 机器学习; 覆盖算法; 多侧面递进; 特征选择
中图分类号: TP393　**文献标识码**: A　**文章编号**: 0372-2112 (2005) 02-0327-05

A Multi-Side Increase by Degrees Algorithm at Machine Learning

ZHANG Yan-ping[1,2] ,ZHANG Ling[1,2] ,WU Tao[2]
(1. *Key Lab of Intelligent computing & Signal Processing at Anhui University, Ministry of Education, Hefei, Anhui* 230039, *China*;
2. *Institute of Artificial intelligence, Anhui University, Hefei, Anhui* 230039, *China*)

Abstract: The conflict between validity and extensibility can be solved by using a multi-side increase by degrees algorithm (shortened form MIDA) at machine learning in a data set with a feature space of high dimensionality and with large amount of samples that belong to many different classes. In the algorithm, the sample set is divided into several sample subsets step by step by double-point. These feature subsets to match each sample set may be extracted at the same time. The method of different treatment to each sample subset is similas to that facing difficult problems people consider and seek a answer from different emphases and multi sides. The algorithm can classify the difficult problems effectually and raise the extensibility and reduce the complexity in condition of established accuracy. The multi-side increase by degrees algorithm bases on a covering algorithm at machine learning. MIDA is used to classify a date set from Shanghai's stock, and the result is satisfied.

Key words: machine learning; covering algorithm; multi-side increase by degrees; feature extraction

1　引言

人类在解决复杂问题时,通常不是一次性地考虑问题的全部细节,而是先把问题分解或简化,忽略其中的部分细节,然后从简化的较抽象层次开始,层层分析研究,实现从局部到全体的解决问题的方法. 例如,对机器零部件,人们习惯用主视图、俯视图、侧视图来分析,若三视图还不能详细给出部件特征,则可对特殊部分进一步说明.

用数学语言描述就是:若元素 x 的属性函数是多维的,有 n 个属性函数分量 $f_1, f_2, \cdots, f_n$,若暂不考虑其中 i 个属性 $f_1, f_2, \cdots, f_i$,将 $f_{i+1}, f_{i+2}, \cdots, f_n$ 属性作为分析研究对象. 同样地对海量数据,人们首先想到的是:是否可以对数据进行某种划分,分成若干小块(每小块的规模可以处理),然后将其合并起来,得到整个对象的特性. 但随即就引出一个问题:如何分? 按什么原则进行划分? 其附加的计算量如何等等.

这方面已有不少的研究,统计学中的主成份分析,特征向量的计算等[1,2],在知识挖掘中利用粗糙集方法的(属性函数)的约简[3],或是定义人为的标准按此标准对各特征进行选择[4]等. 这些方法的一个特点就是希望找到一组特征,它对全体数据都合适,即利用这组特征能把所讨论的数据进行合乎要求的分类(指分类问题).

由于覆盖算法[5,6]对测试样本可存在"拒认状态",利用这一特点,我们提出一个分层划分特征向量的办法,即多侧面递进的学习算法及多侧面递进的学习变形算法,此法充分考虑了上述的问题,能在保证一定精度的前提下,较好的解决上述问题,有效地提高识别率、降低计算量,使算法的泛化能力大大地提高,为处理海量数据问题,提供一个新的途径.

收稿日期:2003-08-30;修回日期:2004-12-22
基金项目:973 计划(No. 2004CB318108);国家自然科学基金(No. 60475017, No. 60135010, No. 60175018);安徽省自然科学基金;教育部博士点基金(No. 20040357002)

2 多侧面递进算法

2.1 覆盖算法简介

张铃教授于 1997 年就给出了 M-P 神经元模型的几何意义[3],指出用三层神经网络构造分类器,等价于求出一组领域,这组领域能将不同类的点分隔开来,并进一步给出覆盖设计算法[5~8].

定义 1 覆盖 C 是指 n 维欧氏空间中的某一球形领域(以 a 为中心,以 r 为半径的开超球).

定义 2 给定样本集 K 分为 s 类,即 $K=(K^1,\cdots,K^i,\cdots,K^S)$, $i=1,2,\cdots,s$. 若覆盖集 $C=\{C^1,\cdots,C^j,\cdots,C^P\}$, $j=1,2,\cdots,p$. 满足:每个 C^j 只与一类 K^i 相交,而不与其他类的 K 相交,且 C^j 的并覆盖整个 K,则称 C 是 K 的划分覆盖集.

定义 3 设覆盖集 $C=\{C^1,\cdots,C^p\}$ 是样本集 K 的划分覆盖集,称 C 是无冗余的,若从 C 中任删去一覆盖 C^j 则 $C/\{C^j\}$ 就不再是 K 的划分覆盖.

该算法的主要思路是:先求一个领域 C^1,它只覆盖一类中的点,而不覆盖其他类的点;对余下的点求二类覆盖领域 C^2,它只覆盖二类中的点而不覆盖其他类的点;…如此交叉进行覆盖,直到样本集中的点均被领域覆盖了为止.

覆盖算法的实质就是用求出的覆盖领域作为三层网络的隐含层,输入层为测试集,输出层为测试集的分类结果. 以问题的方式解释. 设给定一输入集 $K=\{x^1,x^2,\cdots,x^k\}$ (K 是 n 维欧氏空间的点集),设 K 分为 s 个子集 $K^1=\{x^1,x^2,\cdots,x^{m(1)}\}$, $\cdots$, $K^s=\{x^{m(s-1)+1},x^{m(s-2)+2},\cdots,x^k\}$,通过三层网络后,属于 K^i 的点的输出均为"y^i",其中 $y^i=(0,\cdots,1,0,\cdots,0)$(即其第 i 个分量为 1,其余分量为 0 的向量). $i=1,2,\cdots\cdots,s$.

第 2 层组件层,取 p(p 为覆盖领域的个数)个组件 $A^1,A^2,\cdots,A^P$,其中 A^i 为对应于 C^i 的神经元,$i=1,2,\cdots,p$. 其功能函数可认为是特征函数.

第 3 层输出层,取一个输入 x^j,其与 A^i(设以 a^i 为中心,r^i 为半径)神经元的权和阈值($W^i=(\omega^j)$, $\theta=(\theta_i)$),可按下面的公式求得:

$$\theta_i=r_i,\ W=(a^i),\ \theta=(\theta_i),\ r(x)=\langle a^i,x\rangle$$

若 $r(x)>=r_i$,则对应的 y^j 的第 i 个分量为 1,否则为 0. 这样的网络就构成分类器,将 K 分为 s 个子集.

覆盖算法完全真实地反映了样本的分布,因此,可迅速地、构造性地得到对于训练数据百分之百正确分类的神经网络,而不必像传统的 BP 算法那样反复训练还不一定获得好结果. 由于输入向量的类别完全由它被哪个领域覆盖所确定,因此没有被任何覆盖领域覆盖的训练数据,就产生了"拒认状态",称之为拒识.

在模式识别的分类器设计阶段,学习机器需要寻找样本之间的相似性,并根据训练样本集合上的相似性去归纳(推测)更一般的相似规则,从而获得具有预测能力的分类器,这实际上是一种不完全归纳过程[9]. 为此,过于精确和细致的描述对象,必使其的泛化能力低,而过于粗略和笼统的描述,虽泛化能力增加,但识别的精确性必然降低.

对覆盖算法而言,在统计学习理论的框架下,覆盖算法的对训练样本百分之百正确识别的特性可认为是经验风险趋于零,此时提高学习机器的泛化能力的最有效的方法就是简化学习机器的复杂度(减小 VC 维)[4],为此,我们提出降维、减少覆盖领域的总数的改进的覆盖算法-多侧面递进的学习算法.

2.2 基本思想

我们先以个简单的例子来说明算法的基本思想. 设样本集 $K=\{\{1,2,3,4,5,6\},\{7,8,9,10,11,12,13\}\}$,即分为两类,每个样本 x_i 是一 $n=4$ 维向量,具体如下:

$x1=(2,2,1,2)$, $x2=(4,3,2,1)$, $x3=(3,4,4,1)$, $x4=(11,3,2,2)$, $x5=(2,8,2,3)$, $x6=(4,9,2,1)$, $x7=(9,3,2,7)$, $x8=(10,2,2,1)$, $x9=(11,7,2,5)$, $x10=(2,9,6,2)$, $x11=(8,7,2,3)$, $x12=(2,8,7,5)$, $x13=(9,5,1,3)$

设 K 在一 2 维平面($x1$, $x2$)上的投影 $F1$ 如图 1 所示. 注:小圆点和小方点相连表示两类点的重合,图中点 5(小圆点)与点 12(小方点)在这个平面上的投影相重合.

划分算法的思路是:

现对图 1,进行分类求解. 第一轮,用覆盖算法对图 1 求覆盖,得覆盖领域八个,其中,有五个覆盖领域的半径较小,且覆盖领域中所盖住的点数很少;还有一对点(12,5)(简称点对,即:某一覆盖领域最靠近的两个异类点)重合,无法用覆盖领域将其分开,也称为拒识向量. 对半径过小的覆盖领域,从特征选择的角度分析,就是对这些向量,目前给出的特征虽能将它与异类区别开来,但其区别的能力较弱,故泛化能力低,缺乏联想能力. 对异类点重合,说明目前给出的特征对这些向量是不合适的. 因此,对半径过小的覆盖领域的向量和拒识向量,还需要补充特征再学习. 本例中取 $\varepsilon=2$,对覆盖领域的半径 $R<\varepsilon=2$ 的向量,再补充特征,继续学习.

由图 1 知,被这些覆盖领域盖住的点集为 $K2=\{\{4,5,6\},\{7,8,10,12\}\}$,对每个覆盖取一点对,$K2$ 对应的点对为:

$P2=\{(4,8),(5,12),(6,12),(7,4),(8,4),(10,5),(12,5)\}$.

对 $P2$ 统计点对各分量差的绝对值之和. 如 $P2$ 点对中 $x1$ 分量差的绝对值之和 $=6$, x_2 分量差的绝对值之和 $=4$,所以从 $F1$ 中删除绝对值最小的 f 个分量(一般:$f\leqq(F1/2)$),如在例中,删去属性 x_2. 再对 $F/F1$ 中的属性统计 $P2$ 点对对应的分量差值的绝对值之和,取绝对值之和最大的 f 个属性,加入 $F1$. 如在例中,加入属性 x_3 得 $F2=\{x_1,x_3\}$.

第二轮,求 $K2$ 关于属性 $F2$ 的覆盖,得四个覆盖,见图 2,仍取 $\varepsilon=2$,得 4、7 为圆心的两覆盖半径 <2. 得 $K3=\{(4),$

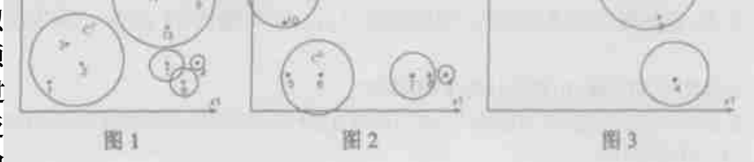

图 1 图 2 图 3

(8)},$P3=\{(4,8)\}$. $P3$ 的 x_1,x_3 分量的绝对值分别为 1,0,故删去 x_3 分量,加入 x_4 分量得 $F3=\{x_1,x_4\}$.

第三轮求 $K3=\{(4),(7,8)\}$,关于 $F3=\{x_1,x_4\}$的覆盖. 见图 3. 求覆盖,仍取 $\varepsilon=2$,得两个覆盖半径均大于 2. 结束.

这样我们将样本划分成三组,{(1,2,3),(9,11,13)};{(5,6),(10,12)};{(4),(7,8)}六个覆盖 C^1,…,C^6,它们分别对应于属性组为:$\{x_1,x_2\}$;$\{x_1,x_3\}$;$\{x_1,x_4\}$. 这样,在一定的精度要求(覆盖的半径 $>\varepsilon$)下,我们正确地分类了样本,并降低了计算复杂性.

下面我们给出覆盖算法的复杂性估计问题. 设样本集的规模为 n,维数为 m,用覆盖算法进行求解其计算量为 n^2m. 若每个覆盖,平均覆盖 k 个样本,则计算量为 (n^2m/k). 而同样情况用 SVM 算法(求二次规划),则其计算量的阶为 $O(n^{3.5}m)$[10],故覆盖算法比 SVM 算法的计算量要小得多. 针对多侧面算法,我们设 d_i 代表每一侧面时的维数,对每个 $d_i(i=1,\cdots,s)$,覆盖算法的计算量为:

$n_i^2d_i/k_i$,其中,$n_1=n$,$\sum_{i=2}^{s}n_i=n$,$d_i<m$,一般 d_i 小于 $m/3$,$k_i\geq k$

所以,$\sum_{i=1}^{s}n_i^2d_i/k_i\leq n^2m/3k_1+\sum_{i=2}^{s}n_i^2d_i/3k_i$

又因为:$\sum_{i=2}^{s}n_i^2<n^2$.

所以能保证 $\sum_{i=1}^{s}n^2d_i/k_i$ 一定小于 (n^2m/k),且有,分的侧面越多,越能降低计算的复杂度.

对实际的数据进行测试的的结果也表明上述结论. 例如当样本的规模为 10000 *60 时,每个覆盖平均覆盖个数为 10,则:

覆盖算法:$10000^2*60/10=60*10^7$

SVM:$10000^{3.5}*60=60*10^{14}$

多侧面递进算法其计算量与 SVM 算法相比要快 7 到 8 个数量级,计算复杂度的降低是非常可观的.

2.3　多侧面递进算法

现将算法总结如下:

算法一:

对给定样本集 K(分成两类),和属性集 F.

①取 $K1=K$,取属性子集 $F1$.

②将样本集 K 投影到 $F1$ 所张的子空间上.

③对不同类别的点集用覆盖算法进行求解.

④设求到的覆盖集为 $C1$.

⑤给定 $\varepsilon1>0$,将覆盖半径 $<\varepsilon1$ 的覆盖删掉,记被覆盖的点集为 $K2$. 计算出 $K2$ 对应的点对,得点对集合 $P2$.

⑥统计 $P2$ 中点对的各属性差值绝对值之和,从 $F1$ 中删去绝对值之和小于 k 的最小的 s 个属性. 得 $F1$.

⑦对 $F/F1$ 中的各属性,统计 $P2$ 中点对的属性差值的绝对值之和,取绝对值之和最大的 s 个属性加入 $F1$ 集中,所得的集合记为 $F2$.

⑧$K1\leftarrow K2$,$F1\leftarrow F2$. 回第 2 步.

⑨$Ki=\phi$或小于某个 n 值,停止.

多侧面递进的学习算法中的 $F1$ 的选取在整个算法中起到了基石的作用,如何选取 $F1$ 可按如下的步骤进行:

①对给定样本集 K 用覆盖算法求出所有分类的覆盖领域.

②每个覆盖取一点对(取覆盖领域圆点与覆盖领域外最近距离的不同类点,构成点对).

③统计点对各分量差的绝对值之和,选取 N 个分量差的绝对值最大的分量,即形成 $F1$.

④N 值一般小于样本集 K 维数的 1/3.

当然,若有了一定的先验知识,则可按已知的经验来选取 $F1$ 和确定 N 值.

2.4　多侧面递进算法的分析

多侧面递进算法的特点是:

所谓多侧面算法,就是将样本集分成几个子集,将各子集的样本分别投影到子空间(侧面)上,在子空间上各样本更容易被识别,以提高识别的精度和降低复杂性. 与以往的算法不同在于以往的算法是选出一个子空间,在其上要求所有的样本都能很好被识别(一刀切). 我们是对不同的子集选取不同的子空间(“一国两制”),不同的子集用不同的属性进行判别. 具体可归纳为:

①将复杂的高维海量数据按确定的算法划分,分成若干子集(每个子集的规模可以处理),然后将其合并起来. 这符合人类对复杂问题的处理方式,即对复杂难解的问题,首先按从主要到次要的不同的角度进行分析,得出其基本特征,然后再综合分析. 该方法不仅有效的降低对象的维数,从而降低计算复杂性(如上例中分成三个子集后,对每一子集至需在二维空间中进行分类,而原问题要在 4 维空间中进行分类),而且多角度匹配复杂问题,使算法的泛化能力大大提高.

②利用覆盖算法中的 ’拒认状态 ”,可自动分成子集. 不必用约束条件来描述子集的划分,只要利用落在“拒认状态 ”,就是要另行分为子集,使得样本自动进行子集分解.

③属性分量的增减过程,就是对局部样本的属性投影过程,当参考数 $\varepsilon1$、k 给定后,整个算法是自动完成的. 我们是利用属性分量对样本区分类“能力 ”,自动进行提取的.

④算法与初始 $F1$ 集合的取法有关,而 $F1$ 可通过先覆盖,再由点对求出.

⑤当 $F1$、$\varepsilon1$、k 取定后,上述算法是自动进行的,其结果是确定的.

多侧面递进的学习算法的网络结构如图 4 所示.

由图 4 可看出,多侧面递进的学习算法将覆盖算法中的第二层隐元层按属性投影划分成 N 层,即将原覆盖算法的水平方向的一层隐元划分成纵向的 N 层,其输出也由一层隐元的全匹配输出,转化为属性不同的分层输出,由于要匹配的属性维数的降低,故泛化能力必大大提高,另一方面,将样本集属性分成几个部分,对各部分分别给出适应它们的属性子集,这种分而治之的方法可有效防止过学习的情况,对识别率也会大大地改善.

本算法给出了一种自动构造多层前向神经网络的方法,并证明比三层前向神经网络的识别率更高.

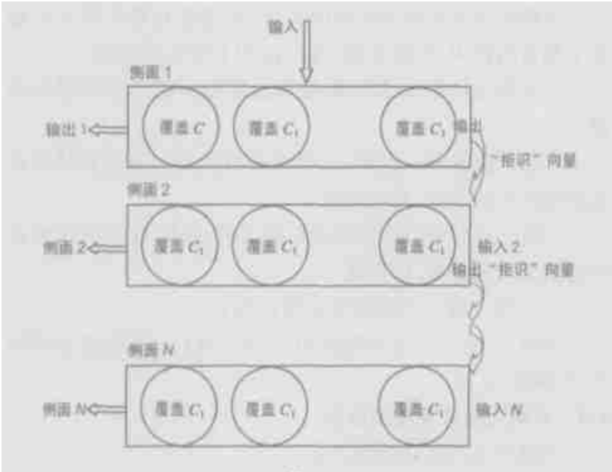

图 4

2.5 算法的变形

在算法一的基础上还可以进行各种变形,下面提供一种变形:

给定参考数 ε_1、ε_2、k. 样本集 K(分成两类),属性集 F.

算法二:

①取属性 F 的子集 $F1$.

②将样本集 K 投影到 $F1$ 所张的子空间上.

③对不同类别的点集用覆盖算法进行求解.

④设求到的覆盖集为 $C1$.

⑤给定 $0<\varepsilon_1<\varepsilon_2$,将覆盖半径 $<\varepsilon_1$ 的覆盖删掉,记被其所覆盖的点集为 $K2$,并计算 $K2$ 对应的点对集合 $P2$.

⑥统计 $P2$ 中点对的属性差绝对值之和,从 $F1$ 中删去和值最小的 s 个属性,得 $F1'$.

⑦对 $F/F1$ 中的各属性,统计 $P2$ 中点对的属性差值绝对值之和,取和值最大的 s 个属性加入 $F1$ 集中,所得的集合记为 $F2$.

⑧令被覆盖半径 $<\varepsilon_2$ 的覆盖盖住的样本集记为 $K2$.

⑨$K1\leftarrow K2$, $F1\leftarrow F2$. 回第 2 步.

⑩$Ki=\phi$或小于某个 n 值,停止.

算法二与算法一不同之处在于,算法一中每个样本最后只被一个覆盖盖住,而算法二中有些样本可能被几个覆盖盖住,如一样本在第一轮中被一半径在 $(\varepsilon_1,\varepsilon_2)$ 之间的覆盖盖住,那么,它在第二轮还要被另外覆盖盖住,故这样的样本可能被几个覆盖盖住,则这些样本可由盖住它的覆盖进行投票来决定它属于哪一类,也可以用某种加权的办法,来决定 S 的所属的类别. 算法二允许样本被几个覆盖盖住,然后投票来决定它属于哪一类,类似与群体决策,从不同侧面对问题进行分析,结果不由某组决定,而是依据某种决策规则进行,故正确率得到改善,但拒识数较多,从而识别率有所下降.

3 应用与测试

我们以 1990 年 12 月 25 日至 2001 年 7 月 16 日的 2600 个交易日的数据转化一组 42 维、共 2000 个样本的训练集(42 * 2000),对 2001 年 7 月 17 日至 2002 年 5 月 17 日的 196 交易日的走势预测为测试集(42 * 196)的实例来说明.

算法	覆盖	多侧面 1	多侧面 2
训练集	一个 42 *2000	五个 10 *2000	五个 10 *2000
测试集	一个 42 *196	五个 10 *196	五个 10 *196
正确	90	156	149
错误	20	31	22
拒识	86	9	25
正确率	90/196 = 46 %	156/196 = 80 %	149/196 = 76 %
识别率	110/196 = 56 %	187/196 = 95 %	171/196 = 87 %

表面上看,算法 1 的测试结果最好,其正确率和识别率均是最大的,仔细分析算法 1 和算法 2 的数据,若将拒识的不考虑,则算法 1 的正确率为:156/(156 + 31) = 83 %,算法 2 的正确率为:149/(149 + 22) = 87 %,也即在算法的识别向量中,算法 2 的正确率要比算法 1 的高,但其识别率要比算法 1 低.

对算法中"拒识"的测试点 X_j,仍用 $F1$(主分量)的覆盖领域,按就近原则以下式确定采用哪个相近的覆盖领域的测试结果.

$$\mathrm{Min}[(X_j-a^i)/d(i)] \tag{1}$$

(设临近的覆盖领域以 a^i 为中心, $d(i)$ 为半径)

算法 1 和算法 2 采用就近原则后,结果为:

算法 1　　正确:162　　错误:34

算法 2　　正确:166　　错误:30

多侧面算法利用覆盖算法中存在"拒认状态"的特点,提出一个合理的属性投影办法,即多侧面递进的学习算法及多侧面递进的学习算法的变形算法,此法能在保证一定精度的前提下,有效地提高识别率、降低计算量,使算法的泛化能力大大地提高,应用结果证实了这点,多侧面递进的学习算法为处理海量数据问题,提供了一个新的途径.

多侧面算法还可有效减少隐元的个数,实验中我们取 ε 的值近似为总覆盖数的 2/5(一般 ε 的值都是由实际覆盖后方可决定,事先预估难以实现). 如主分量组 $F1$ 产生的覆盖共 1193,接近 2/5 的覆盖半径都小于 2,故我们取 $\varepsilon=2$,即删除覆盖半径小于 2 的覆盖,共删去 556 个覆盖领域,识别效果没有变化. 就是说用 $F1$(主特征组)进行测试,其结果为:

测试集	领域个数	正确	错误	拒识
10 *196	1193	93	14	89
10 *196	637	93	14	89

隐元个数在识别效果没有改变的前提下能有效减少,不仅可提高识别速度,而且使网络结构简化[11],因此,多侧面算法对实际问题的求解效果十分理想.

4 结论

本文利用覆盖算法中存在"拒认状态"的特点,提出一个合理划分特征向量的办法,即多侧面递进方法及多侧面递进方法的变形算法,此法能在保证一定精度的前提下,有效地提高识别率、降低计算量,使算法的泛化能力大大地提高,应用结果证实了这点,多侧面递进方法为处理海量数据问题,提供了一个新的途径.

分而治之,对复杂的问题合理划分,获取多侧面的特征是有效的解决方法,对此算法还有许多问题需要我们继续研究,如多侧面实质是一种投影运算,而投影运算实际为商集的运

算[10] ,当取属性子集 $F1$ 时,由 $F1$ 划分的等价类,就是这个投影对应的商集,也即将对象的信息粒度适当的粗化进行分析,得到分类结果.各个不同的投影给出了对象的不同方面的描述,综合各侧面的特征信息,最终获得对象的完整的分类结果.关于多侧面与商空间的关系和多侧面算法的改进,我们仍在继续.

此外,按分类的类别去分别构造合适的向量空间,再综合各子空间的分类结果,如决策森林的构造[12] ,最终获得对象的完整的分类结果也是我们今后的研究方向之一.

参考文献:

[1] Vapnik V N. Statistical Learning Theory[M]. INC. New York:John Wiley & Sons,1998.

[2] 边肇祺,张学工. 模式识别[M]. 北京:清华大学出版社,2001.
Bian Zhao-Qi,Zhang Xue-Gong. Pattern recognition[M]. Beijing:Tsinghua University Press,2001(in Chinese).

[3] Z Pawlak,Rough Sets:Theoretical Aspects of Reasoning about Data [M]. Kluwer Academic Publishers,Dordrecht,1991.

[4] 张燕平. 机器学习中特征提取的新方法-重复覆盖算法[J]. 安徽大学学报,2002,26(2):9 - 13.
Zhang Yan-Ping. A repeated cover algorithm of achieving characteristic rule[J]. Journal of Anhui University,2002,26(2):9 - 13(in Chinese).

[5] Zhang Ling,Zhang Bo. A geometrical representation of McCulloch-Pitts neural model and its applications[J]. IEEE Trans on Neural Networks,1999(ISSN 1045 - 9227),10(4):925 - 929.

[6] 张铃,张钹. 多层前向网络的交叉覆盖设计算法[J]. 软件学报,1999,10(7):737 - 742.
Zhang Ling,Zhang Bo. An alternative covering design algorithm of multi-layer neural networks[J]. Journal of Software,1999,10(7):737 - 742(in Chinese).

[7] Zhang Ling,Zhang Bo. Relational between support vector set and kernel functions in SVM[J]. Journal of Computer Science & Technology,2002,17(5):549 - 555.

[8] 张铃,张钹. 神经网络的规划学习算法[J]. 计算机学报,1994,17(9):669 - 675.
Zhang Ling,Zhang Bo. Programming based learning algorithm of neural networks[J]. Chinese J Computers,1994,17(9):669 - 675(in Chinese).

[9] 张铃,张钹. 人工神经网络理论及应用[M]. 杭州:浙江科学技术出版社,1996.
Zhang Ling,Zhang Bo. Theory and Applications of Artificial Neural Networks[M]. Hangzhou:Zhejiang Science & Technology Press,1996 (in Chinese).

[10] Panos M Pardalos,et al. Algorithms for the solution of quadratic knapsack problems[J]. Linear Algebra and its Applications,1991,152:69 - 82.

[11] 张钹,张铃. 问题求解的理论及应用[M]. 北京:清华大学出版社,1990.
Bo Zhang,Ling Zhang. Theory and Application of Problem Solving [M]. Beijing:Tsinghua University Publisher. 1990(in Chinese).

[12] T K HO. The random subspace method for constructing decision forests [J]. IEEE Trans Patt Analy & Machine Intelligence,1998,20(8):832 - 844.

作者简介:

张燕平 女,1962 出生于安徽巢湖,博士,教授,主要研究领域为人工神经网络、机器学习、人工智能在金融工程中的应用. E-mail:ZhangYP@mail.hf.ah.cn.

张　铃 男,1937 年出生于福建福清,教授,博士生导师,主要研究方向:人工智能理论、机器学习理论和方法、智能计算技术、神经网络技术等.