

基于超曲面的分类算法研究进展

何 清^{1,2}, 史忠植^{1,2}

(1. 中国科学院 计算技术研究所, 北京 100080; 2. 中国科学院 研究生院, 北京 100039)

摘 要:综述了基于超曲面的分类算法,该算法通过区域合并计算获得多个超平面组成的双侧闭曲面作为分类超曲面对空间进行划分. 分类超曲面可以有效地解决在有限连通区域分布很复杂的非线性数据多类分类问题,分析了算法准确率与极小样本集的关系,总结了已有成就和最新进展,指出了基于超曲面的分类算法进一步发展的方向.

关键词:超曲面;分类算法;机器学习

中图分类号: TP301 **文献标识码:** A **文章编号:** 1673-4785(2007)06-0001-07

Research advances in classification algorithm based on hyper-surface

HE Qing^{1,2}, SHI Zhong-zhi^{1,2}

(1. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China; 2. Graduate University, Chinese Academy of Sciences, Beijing 100039, China)

Abstract :In this paper , a classification method based on Hyper Surface (“ HSC ”for short) is introduced. In this method , the space is partitioned through classification hyper-surface which are double-sided closed surfaces consisting of several hyper-surfaces by merging the connected regions. HSC can efficiently solve the nonlinear multi-class classification problems , in which the sample data distributions are very complicated within the finite connected regions. The relationship between accuracy and the minimal consistent subset is analyzed. Finally , the existing achievements and the latest progresses in this subject are summarized , and the future research directions are pointed out.

Key words :hyper-surface ; classification algorithm ; machine learning

分类算法研究是机器学习的核心研究内容,分类能力是人类智能的最显著特征之一. 机器学习在最近 30 多年取得了很大进展. 1971 年,前苏联数学家 Vapnik 与 Chervonenkis 提出了一种基于 VC 维度的对空间划分的理论^[1]. 1984 年 Valiant 提出了可学习理论 (probability approximately correct, PAC),并将可学习性与计算复杂性联系在一起^[2]. 在 Valiant 学习理论中,有 2 种学习复杂性测度. 一是样本复杂性,这是随机实例的数目,用以产生具有高的概率和小的误差;二是计算复杂性,定义为最坏情况下给定数目的样本产生假设所要求的计算时间. 这 2 种复杂性在对分类算法的研究中起着重要作用. 1986 年 Blumer 等人证明了 VC 维度与 Val-

iant 的“可学习理论”(PAC)之间的联系^[3]. 关于 PAC 的研究派生出被称为“计算学习理论 (computational learning theory, COLT)”的学派,现已定期召开这方面的国际会议^[4]. 1995 年 Vapnik 出版了“统计学习理论”一书 (the nature of statistical learning theory),在理论上,这是继 Duda 等人在 20 世纪 60 年代奠定统计模式识别理论之后^[5-6],对统计模式识别最为完整的研究. 这个理论的基础之一是 VC 维度 (Vapnik-Chervonenkis dimension). 关于统计学习理论最新综述包括在文献[7]中. 事实上,对 PAC 的研究一直是理论性的、存在性的, Vapnik 的这个研究却是构造性的,并将感知机的研究包括在其中,他将这种模型称为支持向量机 (support vector machine, SVM). 在 SVM 的研究中我国学者作出了大量工作^[8-11]. 与此同时,在基于覆盖的分类学习算法方面,我国学者在最近 10 年

收稿日期:2007-01-10.
基金项目:国家自然科学基金资助项目 (60435010, 60675010); 国家重点基础研究发展计划资助项目 (2006AA01Z128); 北京市自然科学基金资助项目 (4052025).

相继提出了一些很有价值的分类学习方法. 张铃、张钹教授给出了 M-P 神经元的几何意义, 通过球面投影变换将神经网络的最优设计问题转化为某种最优覆盖问题^[12]. 张铃、张钹教授给出了邻域覆盖算法和交叉覆盖算法, 以及改进的函数覆盖算法和核覆盖算法. 王守觉院士提出了仿生模式识别(拓扑模式识别)^[17-18], 它是基于“认识”事物而不是基于“区分”事物为目的, 与传统以“最佳划分”为目标的统计模式识别相比, 它更接近于人类“认识”事物的特性, 故称为“仿生模式识别”. 徐宗本教授最近提出了一种基于视觉的分类方法 (visual classification algorithm, VCA)^[21], 这种方法是针对目前很多分类方法主要是通过发现数据内在的结构来分类, 很少或没有注意到从模拟人的感觉和感知来进行分类而提出的. 何清等提出了基于超曲面的分类学习算法 (hyper surface classification, HSC)^[23], 这也是一种覆盖分类算法. 本文综述了基于超曲面的分类算法, 对算法进行了详细分析和阐述, 并指出了基于超曲面的分类算法进一步发展的方向.

1 基于超曲面的分类学习算法

基于超曲面的分类学习算法基本思想是从几何学和拓扑学出发的, 该算法基于 Jordan 曲线定理, 根据围绕数的奇偶进行分类判断, 该算法不需要考虑使用何种核函数, 而直接地解决非线性分类问题. 下面给出 Jordan 曲线定理和分类判别定理.

Jordan 曲线定理 设 $X \subset R^3$ 是闭子集, X 同胚于球面 S^2 , 那么它的余集 R^3/S 有 2 个连通分支, 一个是有界的, 另一个是无界的, X 中任何一点的任何邻域与这 2 个连通分支均相交.

上述定理可推广到高维空间.

定理 1 高维空间的 Jordan 曲线定理.

若 $X \subset S^n$ 同胚于球面 S^m , 那么 $m < n$ 否则 $X = S^n$. 若 $m < n$, 余集的同调群为

$$H_k(S^n \setminus X) \cong \begin{cases} Z \oplus Z, m = n - 1 \text{ 且 } k = 0, \\ Z, m < n - 1 \text{ 且 } k = 0, \\ 0, \text{其他.} \end{cases}$$

当 $m = n - 1$ 时 $S^n \setminus X$ 由 2 个连通分支组成, 当 $m < n - 1$ 时, 只有 1 个连通分支.

Jordan 曲线定理表明任何由 $n - 1$ 维球面经连续变形得到的双侧闭曲面都把 n 维空间分成 2 个区域——一个外部和一个内部, 这种曲面可用于分类, 称之为分类超曲面. 给定一个点 x , 判断它在分类超曲面 X 的内部, 还是在外部的方法是: $x \in X$ 的内部 \Leftarrow 自 x 引出的射线与 X 的相交数 (即 X 关于 x 的

围绕数) 为奇数, $x \in X$ 的外部 \Leftarrow 自 x 引出的射线与 X 的相交数为偶数. 如图 1 所示. 现在问题的关键在于如何获得分类超曲面.

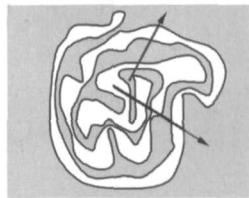


图1 分类判别图示

Fig. 1 Classification theorem illustration

下面的算法给出了构造和使用分类超曲面的基本过程.

假定 M 个训练样本: $K_{\text{train}} = \{x(i) \mid i = 1, 2, \dots, M\}$, 不失一般性, 假设样本点已预先标记好的类别 $t_i \in \{1, 2\}$; 训练空间 D 为一封闭区域, 且其边界满足 Jordan 曲线定理的基本条件.

1) 设前 m 个训练样本落在区域 D 内, $m \in \{1, 2, \dots, M\}$.

2) 将区域 D 划分成 n 个小区域, 设 $D = D_1 \cup D_2 \cup \dots \cup D_n$, $n \geq m$. 且满足以下条件:

D_j ($j = 1, 2, \dots, n$) 的边界 H_j 满足 Jordan 曲线定理的基本条件;

$x(i) \in D_j$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, 不妨设 $i = j$, 即每个小区域至多含一个样本点.

3) 设 D_j ($j = 1, 2, \dots, m$) 的边界 H_j 分别由 k_j 个超平面片组成, 可以将每片超平面表示为 H_p^j , $i = 1, 2, \dots, k_j$; 则可将 H_j 表示为含类别分量和超平面片的集合,

$$H_j = \{t_j, H_p^j \mid p = 1, 2, \dots, k_j; t_j = 1, 2\}.$$

式中: t_j 为 D_j 所含样本点 $x(j)$ 的类别.

4) 若 D_i 与 D_j ($i \neq j$) 相邻, 且 $t_i = t_j$, 则合并边界 H_i 和 H_j (不妨设 H_p^i 与 H_p^j 重合, 即 $H_p^i = H_p^j$), 将重合超平面片消去, 则获得封闭区域 A_L , 其边界为

$$B_L = H_i \cup H_j =$$

$$\{t_L, H_a^i, H_b^j \mid a = 1, \dots, k_i, \text{ 且 } a \neq p; \\ b = 1, \dots, k_j, \text{ 且 } b \neq q; t_L = t_i\}.$$

继续合并相邻同类区域, 最终获得一组由若干超平面片组成的封闭超曲面——分类超曲面, 记为

$$SHS_L^i = \{t_i, HP_l^i \mid i = \{1, 2, \dots, r\};$$

$$l = 1, 2, \dots, k_i; t_i = 1, 2\}, r = m.$$

以链的形式存储 SHS_L^i 的各超平面片 HP_l^i , $l = 1, 2, \dots, k_i$.

5) 输入新样本点 $X = x'(m + 1)$, 按照以下方法

进行判别(设 T 为样本点分类所得类别):选择适当的由待定点 x 出发的射线 f_x , 设 f_x 与 SHS_a^1 的相交点数为 C_a^1 , 分别计算 f_x 与 $\{\text{SHS}_a^1 \mid a=1, \dots, r_1\}$ 和 $\{\text{SHS}_b^2 \mid b=1, \dots, r_2\}$ ($r_1 + r_2 = r$) 的相交点数之和,

记为 $\sum_{a=1}^{r_1} C_a^1$ 和 $\sum_{b=1}^{r_2} C_b^2$, 则有

$$\sum_{a=1}^{r_1} C_a^1 \text{ 为奇数} \Leftrightarrow T = 1$$

或 $\sum_{b=1}^{r_2} C_b^2 \text{ 为奇数} \Leftrightarrow T = 2$.

实际上,求 C_i^j 的过程,就是求 f_x 与 HP_i^j 相交数量的过程.

6) 若不能判别 x 的类别,就对 x 所在的小区域边界进行标定,不妨设 $x^i(m+1) \in D_{m+1}$, 则 D_{m+1} 的边界可表示为

$$H_{m+1} = \{t_{m+1}, H_p^{m+1} \mid p = 1, 2, \dots, k_{m+1}\},$$

之后转入 4), 继续合并相邻同类区域.

以上给出了基于分类超曲面的分类判别方法的基本算法,即通过区域合并计算获得多个超平面组成的双侧闭曲面作为分类超曲面对空间进行划分,也就是在样本点周围形成一个封闭区域,该区域由多个分类超平面片围成,并使得该区域覆盖某一类尽可能多的样本点,同时不覆盖异类样本点.

2 算法特点与已有研究成果

HSC 算法有 2 个关键步骤,一是局部化策略,另一个是用围绕数判断类别.该算法中,判别样本所属类别,不需与所有分类边界链表作相交操作后再判断,而只需满足:由样本点所引射线与一完整分类边界链表相交点数为奇数即可,这样可提高判别速度.这种方法得到的分类超曲面是由若干个封闭闭曲面构成,而曲面的局部是由低维平面片构成.每个闭曲面内部是一类样本,这样对闭曲面可以进行类别标记.样本类别可以是多个,所以这种方法对于多类问题的解决是很方便的,因为多个分类器可以在一次训练过程中产生,避免了 2 类分类器转化为多类分类器的技术处理.对二维和三维双螺旋及 UCI 中的数据分类实验结果说明,分类超曲面可以有效地解决在有限区域分布很复杂的海量 (10^7) 的非线性数据多类分类问题,计算速度较高,同时对计算机资源要求很低,而传统的 SVM 不具备这种优点.另外小样本训练大样本测试结果表明,基于分类超曲面的分类法的泛化能力较好.该方法是对直接解决非线性分类问题的一种尝试,此方法的一个前提是同类样本点应具有在有限个连通分支分布的特

点,但与连通分支的形状无关,在现实中的数据分布大都满足条件之一^[25].

HSC 算法发展至今,已相继解决了二维 2 类分类^[23]、二维多类分类^[24]、二维一般连通区域分类^[25]、三维多类^[26]、高维换维分类问题^[27]、高维多类集成分类问题^[28-45].

关于 HSC 算法的最新进展包括:利用 HSC 算法构造了极小样本集,提出了基于覆盖的极小样本集的概念使之不仅可以代表整个模型,而且可以反映出整个训练集的拓扑结构^[46],并研究了 BAGGING 和 ADABOOST 算法使用不同的训练样本集时对 HSC 分类器性能的影响,试验结果表明,他们对分类精度的提高是受极小样本集制约的^[47].另外从视觉认知角度研究了超曲面在和数据理解中的作用^[48].采用 Agent (智能主体) 技术用于多个 HSC 分类器的合成,使得 HSC 算法适合在分布式环境中进行数据挖掘.这种合成的特点是不把样本集划分为若干小样本集的横向划分,而是对分布在不同地点的不同属性的样本集就地作属性集纵向划分后的合成^[49].

3 基于超曲面的覆盖分类学习算法的研究方向

基于超曲面的分类学习算法 HSC 作为一种新的算法,有很多问题亟待研究,这既包括算法优化,又包括理论分析,还包括应用中遇到的现实问题.

首先是优化高维数据分类算法问题.从理论上讲,这种方法可以推广到高维,因为 Jordan 定理在任意有限维空间都是成立的.但是高维空间中的实现存在以下挑战性问题:一是高维空间的单位方体的合并计算复杂度随着维数增加而提高,另一方面高维超曲面的存储开销大.但是这并不意味着 HSC 不能处理高维数据,借助数据预处理和集成学习技术,对于高维数据处理提出并实现了 2 种解决办法.这 2 种高维处理方法与 HSC 算法特点紧密结合,这就是基于样本数据重排的换维分类学习算法和基于集成学习思想^[30]的分维分类学习算法.有关集成学习周志华教授做了出色的工作^[31-32].HSC 的集成特点是通过分维获得子分类器,不是通过划分样本集获得子分类器^[28].基于样本数据重排的换维分类学习^[27],将涉及到维排序和维组合问题,这些策略具有多样性,他们如何影响分类器性能,如何找到最优的策略是待研究的问题.由于分类器集成方法是基于分维的,那么维的排序策略、划分策略及权重策略就值得研究.

其次是初始的划分尺度与泛化能力的关系问题. 基于超曲面的分类方法与传统方法相比, 如与 Parzen 窗分类器相比^[5], 由于 HSC 采用了局部化策略, 克服了 Parzen 窗在样本分布不均衡情况下, 若窗宽度较小所导致的分类区域过于零散, 分类曲面复杂, 推广性差, 以及窗宽度较大时, 由于分类区域融合过度所造成的分类误差大的问题. 在 HSC 中由于采用局部化策略是对存在异类数据分布的同一单元区域进行, 因而这种方法是基于对数据分布的感知来工作. Parzen 窗分类器的窗宽度是可以通过实验逐步择优的, 但是一旦选定某个值, 在整个分类过程中就不再变化, 这对于分布不均衡的样本分类有明显的缺陷. 但是初始的划分尺度, 对 HSC 的分类精度是有影响的, 研究有关这种影响的估计以及提高精度的策略是重要的. 划分尺度与极小样本集之间的关系. 在一定意义上来说同仿生模式识别类似, HSC 将模式识别问题看成是模式的认识, 而不是分类划分, 不是模式分类. 因而, 其数学模型与传统模式识别的“最优分类”界面的概念大不相同. 划分尺度越小 HSC 所得到的模型和训练样本的拓扑结构的匹配程度就越高, 误识率就越低, 但这也将导致误识率的提高、泛化能力的下降. 但是无论采用多大的划分尺度最终都会得到一个一致的分类模型, 区别在于其细化程度的不同. 还要研究不同划分尺度产生的极小样本集间有何不同, 并比较它们之间的性质.

再次, 最优的 HSC 覆盖问题. 张铃、张钊教授给出了 M-P 神经元的几何意义, 通过球面投影变换将神经网络的最优设计问题转化为某种最优覆盖问题. 他们把神经元与几何上样本的球形邻域对应起来. 按照这种观点, HSC 可以看成神经元是由分类超曲面构成的神经网络, 分类超曲面的个数就是神经元的个数. 所不同的是 HSC 的分类是靠样本围绕数来计算得出的, 而神经网络是通过修正权值后加权计算获得. 如何找到超曲面个数少分类性能好的 HSC 分类器是一个重要的问题, 这也是一个最优覆盖问题.

还有, HSC 的抽取规则问题. 以往人们觉得神经网络学习算法的分类过程不可理解, 难以解释. 这样在 Gallant 提出了一个简单的算法来解释连接主义专家系统所做的推理^[33]. 该算法通过产生规则来解释神经网络如何为某个给定案例得出结论. 其基本思想就是从当前已知的信息集中选择一个能有效地产生该结论的最小信息集合, 也就是说, 不管其他未知输入分量的取值为多少, 只要满足该最小信息

集合的取值要求就可以得出结论. Gallant 的这篇论文开创了神经网络规则抽取这一领域, 成为该领域被引用最多的文献之一. 在 Gallant 之后, 陆续有一些研究者对神经网络规则抽取进行了研究^[34-40]. 1995 年, Andrews 等人^[43]为从神经网络抽取的规则提出了一个评价体系, 并提出了规则抽取算法的分类体系. 前者为不同规则抽取算法的比较提供了标准, 并对新算法的设计具有指导作用, 后者使得对规则抽取算法的系统化分析成为可能. 这 2 个体系为神经网络规则抽取这一领域的进一步发展奠定了基础, 因此, Andrews 等人的这篇论文^[43]被认为是该领域的一个里程碑. HSC 分类超曲面以链表方式表达, 已经实现了分类超曲面的可视化, 但可视化并不意味着数据可理解, 这些链表包含了分类信息, 这些信息能否像神经网络规则抽取那样, 从中抽出分类规则是值得研究的问题. 从分类超曲面中得到的分类规则就是可以学习、可以理解、可以传播的知识了.

另外, HSC 在数据理解中的作用问题. 理解数据, 即获得数据集合的不同简洁程度表示, 已成为机器学习研究的另一个重要研究方向. 解决这个问题的途径不能沿袭传统检验有效性的方法, 即以检验个别事例为基础, 而需要寻找必要的数学理论. 数据理解包括人对数据的理解和机器对数据理解. 人对数据的理解, 可以理解为借用符号机器学习的约简与可解释的特性, 将一本使用数据语言书写的书翻译为人可理解的表示形式, 从而丰富人的知识. 这就是数据挖掘的主要任务之一. 计算机的数据理解就是传统意义下的机器学习, 分类超曲面可以看作数据分布的一个包络, 这是对数据理解的一个方面. 另一个方面的理解是数据分布的主曲线, 主曲线相当于数据分布的骨架, 这两者结合将会得到对数据更全面的理解.

另外, 要研究 HSC 学习算法的计算复杂性, 包括时间复杂性、空间复杂性、样本复杂性以及模型对新问题的求解能力, 或称为泛化能力. 在数学上来看, 学习理论就是通过计算有限的随机样本获得数据中包含的知识^[44-45]. 函数海量数据与学习精确度(泛化能力)以及数据性质的多样性(领域依赖)等要求, 需要考虑使用更多更复杂的数学理论, 如函数逼近论和宽度理论来揭示已有或正在发展的理论与方法所存在的问题, 及其对问题的适应性.

最后, 极小样本集的性质以及与 PAC 样本复杂度的关系问题亟待研究. 此外 PAC 的样本复杂度理论给出了有多少随机抽取训练样例才足以可能近似正确(PAC)地学习到任意目标概念. 通过 HSC 得

到极小样本集的方法给出了选择的训练样本,为了使在其上学习得到的模型有很好的性能应满足的性质(包含某一极小样本集). 将从包含极小样本集这个目标出发,在一定的限制条件下得到一个新的样本集所包含实例个数的边界,并与 PAC 的样本复杂度理论进行比较.

3 结束语

总之,在基于覆盖的分类学习算法方面,基于覆盖思想在最近 10 年相继提出过一些很有价值的分类学习方法和理论分析,但是,大量的理论问题尚未解决. HSC 学习算法作为一种覆盖分类算法,其性能提高以及计算复杂性、泛化能力等理论问题值得深入研究.

参考文献:

- [1] VAPNIK V. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1995.
- [2] VALIANT L G. A theory of the learnable[J]. Communications of the ACM, 1984, 27(11): 1134 - 1142.
- [3] BLUMER A, EHRENFEUCHT A, HASSLER D, et al. Classifying learnable geometric concepts with the Vapnik-Chervonenkis dimension[A]. Proceedings of the 19th Annual ACM Symposium on Theory of Computing [C]. Berkeley, US, 1986.
- [4] DANA A. Computational learning theory: survey and selected bibliography[A]. Proceedings of the Twenty-fourth Annual ACM Symposium on Theory of Computing[C]. [S.L.], 1992.
- [5] DUDA R O, HART P E. Pattern classification and scene analysis [M]. New York: John Wiley & Sons, 1973.
- [6] DUDA R O, HART P E, STOCK D G. 模式分类[M]. 北京:机械工业出版社, 2003.
- [7] MENDELSON S. A few notes on statistical learning theory [R]. Advanced Lectures on Machine Learning, LNAI 2600, 2003.
- [8] 张文生, 丁辉, 王珏. 基于邻域原理计算海量数据支持向量的研究[J]. 软件学报, 2001, 12(5): 711 - 720.
ZHANG Wensheng, DING Hui, WANG Jue. Study on computing the support vectors of massive data based on neighborhood principle [J]. Journal of Software, 2001, 12(5): 711 - 720.
- [9] TAO Q, WANG J. Kernel projection algorithm for large-scale SVM problems[J]. Journal of Computer Science Technology, 2002, 17(5): 556 - 564.
- [10] TAO Qing, WU Gaowei, WANG Jue. A generalized $\frac{3}{K}$ algorithm for learning-SVM classifiers[J]. Pattern Recognition Letters, 2004, 25(10): 1165 - 1171.
- [11] 许建华, 张学工, 李衍达. 一种基于核函数的非线性感知器算法[J]. 计算机学报, 2002, 25(7): 689 - 695.
XU Jianhua, ZHANG Xuegong, LI Yanda. A nonlinear perceptron algorithm based on kernel functions[J]. Chinese Journal of Computers, 2002, 25(7): 689 - 695.
- [12] ZHANG Ling, ZHANG Bo. A geometrical representation of McCulloch-Pitts neural model and its applications [J]. IEEE Transactions on Neural Networks, 1999, 10(4): 925 - 929.
- [13] 吴涛, 张铃, 张燕平. 机器学习中的核覆盖算法[J]. 计算机学报, 2005, 28(8): 1295 - 1301.
WU Tao, ZHANG Ling, ZHANG Yanping. Kernel covering algorithm for machine learning [J]. Chinese Journal of Computers, 2005, 28(8): 1295 - 1301.
- [14] 陶品, 张钹, 叶榛. 构造型神经网络双交叉覆盖增量学习算法[J]. 软件学报, 2003, 14(2): 194 - 201.
TAO Pin, ZHANG Bo, YE Zhen. An incremental bi-covering learning algorithm for constructive neural network[J]. Journal of Software, 2003, 14(2): 194 - 201.
- [15] 张铃, 张钹, 殷海风. 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, YIN Haifeng. An alternative covering design algorithm of multi-layer neural networks[J]. Journal of Software, 1999, 10(7): 737 - 742.
- [16] 周志华, 曹存根. 神经网络及其应用[M]. 北京:清华大学出版社, 2004.
- [17] 王守觉. 仿生模式识别(拓扑模式识别)——一种模式识别新模型的理论与应用[J]. 电子学报, 2002, 30(10): 1417 - 1420.
WANG Shoujue. Bionic (topological) pattern recognition—a new model of pattern recognition theory and its applications [J]. Acta Electronica Sinica, 2002, 30(10): 1417 - 1420.
- [18] WANG S J, QU Y F, LI W J. Face recognition: biomimetic pattern recognition vs traditional recognition[J]. Acta Electronica Sinica, 2004, 32(7): 1057 - 1061.
- [19] CAO W M, HAO F, WANG S J. The application of DBF neural networks for object recognition[J]. Information Sciences-Informatics and Computer Science: An International Journal, 2004, 160(1-4): 153 - 160.
- [20] 王守觉, 王柏南. 人工神经网络的多维空间几何分析及其理论[J]. 电子学报, 2002, 30(1): 1 - 4.
WANG Shoujue, WANG Bainan. Analysis and theory of high-dimension space geometry for artificial neural networks[J]. Acta Electronica Sinica, 2002, 30(1): 1 - 4.

- [21] XU Zongben, MENG Deyu, JING Wenfeng. A new approach for classification: visual simulation point of view [A]. Proceedings of the ISNN 2005 [C]. Springer-Verlag, 2005.
- [22] 张涌, 朱洪. 一类弱集合覆盖问题的近似算法[J]. 计算机学报, 2005, 28(9): 1497 - 1500.
ZHANG Yong, ZHU Hong. Approximation algorithms for the problems of weak set cover[J]. Chinese Journal of Computers, 2005, 28(9): 1497 - 1500.
- [23] HE Qing, SHI Zhongzhi, REN Li'an. The classification method based on hyper surface [A]. Proceedings of the IEEE International Joint Conference on Neural Networks [C]. Hawaii, 2002.
- [24] HE Qing, SHI Zhongzhi, REN Li'an. The multi-class classification method in large database based on hyper surface [A]. Proceedings of the International Conference on Machine Learning and Application [C]. Las Vegas, 2002.
- [25] 何清, 任力安, 史忠植. 基于超曲面的海量数据直接分类法[J]. 计算机学报, 2003, 26(2): 206 - 211.
HE Qing, REN Li'an, SHI Zhongzhi. The large data direct classifying method based on hyper surface [J]. Chinese Journal of Computers, 2003, 26(2): 206 - 211.
- [26] HE Qing, SHI Zhongzhi, REN Li'an, et al. A novel classification method based on hypersurface [J]. International Journal of Mathematical and Computer Modeling, 2003, 38: 395 - 407.
- [27] HE Qing, ZHAO Xiurong, SHI Zhongzhi. Classification based on dimension transposition for high dimension data [J]. Soft Computing, 2007, 11(4): 329 - 334.
- [28] ZHAO Xiurong, HE Qing, SHI Zhongzhi. Hyper surface classifier ensemble for high dimensional data sets [A]. Proceedings of the 3rd International Symposium on Neural Networks (ISNN 2006) [C]. Berlin: Springer-Verlag, 2006.
- [29] DASARATHY B V. Minimal consistent set (MCS) identification for optimal nearest neighbor decision systems design [J]. IEEE Trans Syst, Man, Cybern, 1994, 24(3): 511 - 517.
- [30] HANSEN L K, SALAMON P. Neural network ensembles [J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1990, 12(10): 993 - 1001.
- [31] ZHOU Zhihua, WU Jianxin, JIANG Yuan, et al. Genetic algorithm based selective neural network ensemble [A]. Proceedings of the 17th International Joint Conference on Artificial Intelligence [C]. Seattle, US, 2001.
- [32] ZHOU Z H, WU J, TANG W. Ensembling neural networks: many could be better than all [J]. Artificial Intelligence, 2002, 137(1-2): 239 - 263.
- [33] GALLANT S I. Connectionist expert systems [J]. Communications of the ACM, 1988, 31(2): 152 - 169.
- [34] SETIONO R, LIU H. Understanding neural networks via rule extraction [A]. Proceedings of the 14th International Joint Conference on Artificial Intelligence [C]. Montreal, Canada, 1995.
- [35] WU X. Knowledge acquisition from databases [M]. Norwood, 1995.
- [36] ALEXANDER J A, MOZER M C. Template-based procedures for neural network interpretation [J]. Neural Networks, 1999, 12(3): 479 - 498.
- [37] 周志华, 曹存根. 神经网络及其应用 [M]. 北京: 清华大学出版社, 2004.
- [38] ZHOU Z H. Rule extraction: using neural networks or for neural networks [J]. Journal of Computer Science and Technology, 2004, 19(2): 24 - 253.
- [39] ZHOU Z H, JIANG Y, CHEN S F. A general neural framework for classification rule mining [J]. International Journal of Computers, Systems and Signals, 2000, 1(2): 154 - 168.
- [40] 周志华, 陈世福. 神经网络规则抽取 [J]. 计算机研究与发展, 2002, 39(4): 398 - 405.
ZHOU Zhihua, CHEN Shifu. Rule extraction from neural networks [J]. Journal of Computers Research and Development, 2002, 39(4): 398 - 405.
- [41] 周志华, 陈世福. 神经网络集成 [J]. 计算机学报, 2002, 25(1): 1 - 8.
ZHOU Zhihua, CHEN Shifu. Neural network ensemble [J]. Chinese Journal of Computers, 2002, 25(1): 1 - 8.
- [42] ZHOU Z H, JIANG Y, CHEN S F. Extracting symbolic rules from trained neural networks [J]. AI Communications, 2003, 16(1): 3 - 15.
- [43] ANDREWS R, DIEDERICH J, TICKLE A B. Survey and critique of techniques for extracting rules from trained artificial neural networks [J]. Knowledge-Based Systems, 1995, 8(6): 373 - 389.
- [44] JOHN S T, ROBERT C, WILLIAMSON P. Generalization performance of classifiers in terms of observed covering numbers [A]. Proceedings of the EuroCOLT99 [C]. Berlin: Springer-Verlag, 1999.
- [45] 王珏, 石纯一. 机器学习研究 [J]. 广西师范大学学报 (自然科学版), 2003, 21(2): 1 - 15.
WANG Jue, SHI Chunyi. Investigations on machine learning [J]. Journal of Guangxi Normal University (Natural Science Edition), 2003, 21(2): 1 - 15.
- [46] HE Qing, ZHAO Xiurong, SHI Zhongzhi. Sampling based on minimal consistent subset for hyper surface classification [A]. Proceedings of the Sixth International Conference on Machine Learning and Cybernetics

[C]. Hong Kong, 2007.

[47] HE Qing, ZHUANG Fuzhen, ZHAO Xiurong, et al. Enhanced algorithm performance for classification based on hyper surface using bagging and adaboost [A]. Proceedings of the Sixth International Conference on Machine Learning and Cybernetics [C]. Hong Kong, 2007.

[48] HE Qing, ZHAO Xiurong, SHI Zhongzhi. A cognitive data visualization method based on hyped surface [A]. Proc 6th IEEE Int Conf on Cognitive Informatics (IC-CIO07) [C]. [S.l.], 2007.

[49] HE Qing, ZHAO Xiurong, LUO Ping, et al. Combination methodologies of multi-Agent hyper surface classifiers: design and implementation issues [A]. LNAI 4476 [C]. Berlin: Springer-Verlag, 2007.

作者简介:



何清, 1965年生, 教授, 博士生导师, 主要研究方向为人工智能、数据挖掘、机器学习、模糊集理论, 发表学术论文 60 多篇。

E-mail: heq @ics.ict.ac.cn.



史忠植, 1944年生, 研究员, 博士生导师, 主要研究方向为人工智能、多主体系统、数据挖掘、机器学习、知识工程等。1979年、1998年、2001年均获中国科学院科技进步二等奖, 1994年获中国科学院科技进步特等奖, 2002年获国家科技进步二等奖, 发表学术论文 400 多篇, 出版专著 5 部。

E-mail: shizz @ics.ict.ac.cn.

《可拓工程》简介

Introduction to Extensional Engineering

可拓工程的基本思想是用形式化的方法处理各领域中的矛盾问题, 研究如何化不相容为相容、化对立为共存、化不是为是。

把理论应用于各个实际领域的关键在于方法。为了更多的学者能运用可拓学的基本理论去处理所在领域的矛盾问题, 我们总结了多年来的研究工作, 从可拓学的基本原理出发, 完善和发展了可拓方法, 它是可拓论应用于实际的桥梁。把可拓方法与各实际领域相结合, 去解决其中的矛盾问题的方法, 统称为可拓工程方法。本书系统地阐述了可拓工程的理论基础、方法体系和应用领域, 并给出可拓工程方法的应用案例。

为便于读者学习, 本书第一章简要介绍可拓学的概况, 包括可拓学的研究概况与发展历程、可拓学的理论框架、可拓学的方法论体系、研究可拓学的科学意义和可拓工程研究现状; 第二章介绍可拓工程的理论基础——可拓论, 包括基元的概念、拓展分析原理、共轭分析原理、可拓变换、复合元、可拓集、关联函数和可拓逻辑简介; 第三章介绍可拓工程的方法基础——可拓方法, 包括拓展分析方法、可拓变换方法、共轭分析与共轭变换方法、可拓集方法、优度评价方法和可拓思维模式; 第四章介绍矛盾问题的求解方法, 包括矛盾问题的界定及其可拓模型、矛盾问题的运算、拓展与变换、不相容问题的求解方法——可拓策略生成方法、对立问题的求解方法——转换桥方法和矛盾问题智能化处理的初步研究; 第五章介绍可拓工程方法与技术, 包括可拓信息-知识-策略的形式化体系、可拓策略生成系统的实用技术、可拓数据挖掘方法、可拓营销方法、可拓策划方法、可拓设计方法、可拓控制与可拓检测方法简介和可拓方法在识别、搜索、诊断中的应用。各部分内容都以若干案例来帮助读者理解可拓学的基本方法和可能的应用。我们期望高等院校和科研单位的教学科研人员能将方法与自己的研究领域相结合, 提出更多适合于各专业的可拓工程方法。

本书是作者承接和参加的国家自然科学基金项目和广东省自然科学基金项目的有关研究成果的总结, 作者冀求以此拙作作为引玉之砖, 以使更多领域的学者利用可拓工程方法去处理所在领域的矛盾问题, 同时, 也希望这本书能成为可拓学通向应用之路的桥梁。本书理论与应用相结合, 分析透彻, 可操作性强。为方便不同知识背景和不同层次读者的学习, 各部分内容都配备一些通俗易懂的案例, 适合高等院校从事管理科学与工程、智能科学、信息科学、计算机、设计等领域的师生、工程技术人员和管理决策人员阅读, 特别适合作为高等院校相关专业本科、硕士、博士生的选修课教材。

本书是“可拓学丛书”之一, 由杨春燕研究员和蔡文研究员共同撰写, 全书约 35 万字, 科学出版社于 2007 年 7 月出版。