粗糙近似算子在模式的可能性和必然性分类中的应用

邓方安[1,2],刘三阳[3],徐 扬[1],杨 磊[3]

(1.西南交通大学电气工程学院,四川成都 610031;2.陕西理工学院数学与计算机科学系,陕西汉中 723001;
3.西安电子科技大学应用数学系,陕西西安 710071)

摘 要: 研究粗糙近似算子关于模式"二分法"的相关性质,描述了利用给定模式把模式空间划分成两组的模式分类的可能性和必然性,并设计了有边界区域的模式分类的可能性和必然性的粗糙神经网络算法.最后,用仿真实验验证了算法的有效性.

关键词: 粗糙近似算子; 模式分类; 可能性和必然性

中图分类号: TP18 **文献标识码:** A **文章编号:** 0372-2112(2004)04-0697-04

Applications of Rough Approximate Operators in the Possibility and Necessity of Pattern Classification

DENG Fang-an[1,2], LIU San-yang[3], XU Yang[1], YANG Lei[3]

(1. *School of Elec. Eng., Southwest Jiaotong University, Chengdu, Sichuan* 610031, *China*;
2. *Department of Mathematics and Computer Science, ShaanXi University of technology, Hanzhong, Shaanxi* 723001, *China*;
3. *Department of Applied Mathematics, Xidian University, Xi'an, Shaanxi* 710071, *China*)

Abstract: In this paper, the properties of rough approximate operator for bisection method of pattern are discussed. The possibility and necessity of pattern classification by dividing pattern space into two decision areas using the given patterns are described. The rough neural networks algorithm of the possibility and necessity of the classification problem with boundary areas is designed. Finally, numerical examples are given to illustrate the efficiency of algorithm.

Key words: rough approximate operator; pattern classification; possibility and necessity

1 引言

粗糙集理论是处理不精确数据和发现新知识的一种新的数据分析工具,它在工业过程控制、医疗诊断、股票分析、水文气象等领域已经受到高度重视并得到了成功的应用.H. Ishibuchi利用神经网络研究了模式的可能性和必然性的分类算法[1].本文基于粗糙集近似算子的性质,研究把模式分成两组的可能性和必然性.

第一部分基于粗糙近似算子给出了模式二分法的相关性质,第二部分叙述了粗糙近似精度与分类器选择的关系.第三部分研究了有边界区域的模式的可能性和必然性粗糙分类算法.最后,用仿真实验验证了算法的有效性.

2 基于粗糙近似的模式"二分法"的相关性质

设 $K=(U,R)$ 为一个信息系统,其中 $U=\{u_1,u_2,\cdots,u_n\}$ 为非空有限集合(论域),R 是 U 上的一族等价关系,对于任意一个子集 $X\subseteq U$,我们用 U/R 表示关系 R 的等价类,并且定义 X 关于 R 的下近似和上近似分别为:

$R_*(X)=\{x\in U\mid[x]_R\subseteq X\}$, $R^*(X)=\{x\in U\mid[x]_R\cap X\neq\phi\}$.

称 $bn_R(X)=R^*(X)-R_*(X)$ 为 X 的 R 边界域, $pos_R(X)=R_*(X)$ 为 X 的 R 正域,它包含根据 R 知识判定肯定属于 X 的 U 中元素组成的集合;称 $neg_R(X)=U-R^*(X)$ 为 X 的 R 负域,它包含根据 R 知识判定肯定不属于 X 的 U 中元素组成的集合.关于粗糙集近似算子的基本性质可参见文献[2].

通常的模式分类问题总可以归结为:借助给定的模式确定把模式空间划分为两个不相交决策域的决策规则问题.下面我们用粗糙集的上、下近似算子来描述把模式分成两组的可能性和必然性区域.

定理1 设 $\Omega=\{x_1,x_2,\cdots,x_n\}$ 是一个模式空间,$x_p=(x_{p1},\cdots,x_{pn})$ 是模式 p 的 n 个可测属性值($p=1,2,\cdots,m$).如果存在 $\Omega_1,\Omega_2\subseteq\Omega$,使得

$$\Omega_1\cup\Omega_2=\Omega,\quad \Omega_1\cap\Omega_2=\phi \tag{1}$$

则 $\forall x_p\in\Omega$, $x_p\in\Omega_1$ 或 $x_p\in\Omega_2$ 是惟一确定的.

收稿日期:2002-04-25;修回日期:2004-01-25
基金项目:陕西省自然科学基金(No.2001SL08).

由粗糙集的上下近似算子的性质可得:

$$\Omega_1^* \cup \Omega_2^* = \Omega \qquad (2)$$
$$\Omega_{1*} \cup \Omega_{2*} = \phi$$

因而满足条件(1)的分类规则构成了 Ω 上的一个等价关系,且 Ω_1,Ω_2 是该等价关系下的精确集,即 Ω_1,Ω_2 将 Ω 的 p 个模式分为两个不相交的子模式空间.

定理2 设 $\Omega=\{x_1,x_2,\cdots,x_n\}$ 是一个模式空间,$x_p=(x_{p1},\cdots,x_{pn})$ 是模式 p 的 n 个可测属性值($p=1,2,\cdots,m$).则有下列结论成立:

(Ⅰ)如果存在 $\Omega_1,\Omega_2\subseteq\Omega$,使得

$$\Omega_1\cup\Omega_2=\Omega, \qquad (3)$$
$$\Omega_1\cap\Omega_2\neq\phi$$

则 $\forall x_p\in\Omega,(p=1,2,\cdots,m)$,此时模式 x_p 有可能属于 Ω_1,也有可能属于 Ω_2,其模式空间的分类可用 Ω_1 和 Ω_2 的上、下近似来描述.

(Ⅱ)如果 $\Omega_1\cup\Omega_2=\Omega$,且 $\Omega_2=\Omega-\Omega_1$,则有:

$$\begin{cases}\Omega_2^*=\Omega-\Omega_{1*}\\ \Omega_{2*}=\Omega-\Omega_1^*\end{cases} \qquad (4)$$

根据上下近似算子的性质:$\Omega_{i*}\subseteq\Omega_i\subseteq\Omega_i^*$,$i=1,2$.我们有:

·模式分类的可能性的描述:$\begin{cases}\Omega_1^*\cup\Omega_2^*=\Omega,\\ \Omega_1^*\cap\Omega_2^*\neq\phi\end{cases}$ (5)

·模式分类的必然性的描述:$\begin{cases}\Omega_{1*}\cup\Omega_{2*}\subseteq\Omega,\\ \Omega_{1*}\cap\Omega_{2*}=\phi\end{cases}$ (6)

由粗糙集的定义容易验证定理1和定理2成立,这里不作证明.

注:①由(4)知:$\Omega_1^*\cap\Omega_{2*}=\Omega_2^*\cap\Omega_{1*}=\phi.=\Omega_2^*-\Omega_{2*}$($\Omega_1$ 和 Ω_2 是关于 R 的边界域);

②在(5)中,$\Omega_1^*\cap\Omega_2^*=\Omega_1^*\cap(\Omega-\Omega_{1*})=\Omega_1^*\cap\Omega-\Omega_1^*\cap\Omega_{1*}=\Omega_1^*-\Omega_{1*}$;

③在(6)中,$\Omega_{1*}\cap\Omega_{2*}=\Omega_{1*}\cap(\Omega-\Omega_1^*)=\Omega_{1*}\cap\Omega-\Omega_{1*}\cap\Omega_2^*=\Omega_{1*}-\Omega_{1*}=\phi$.

这里 Ω_1^*,Ω_2^* 分别表示可能属于子空间 Ω_1 和子空间 Ω_2 的样本的集合;Ω_{1*},Ω_{2*} 分别表示必然属于子空间 Ω_1 和子空间 Ω_2 的样本的集合.

定理3 设 $\Omega=\{x_1,x_2,\cdots,x_n\}$ 是一个模式空间,$x_p=(x_{p1},\cdots,x_{pn})$ 是模式 p 的 n 个可测属性值($p=1,2,\cdots,m$).若存在 $\Omega_1,\Omega_2\subseteq\Omega$,且 $\Omega_1\cup\Omega_2=\Omega$,则有下列结论成立:

(ⅰ)$\Omega_{2*}=\Omega-\Omega_1^*\Rightarrow(\Omega_1\cap\Omega_2)_*=\phi$.

(ⅱ)当 $\Omega_1\cap\Omega_2\neq\phi$ 时,则 $\Omega_{2*}\supseteq\Omega-\Omega_1^*$.

(ⅲ)当 $\Omega_2^*=\Omega-\Omega_{1*},\Omega_{2*}=\Omega-\Omega_1^*$ 时,$bn(\Omega_1\cap\Omega_2)\subseteq bn(\Omega_1)\cap bn(\Omega_2)$.

证明(ⅰ)由 $\Omega_{2*}=\Omega-\Omega_1^*$ 得,$(\Omega_1\cap\Omega_2)_*=\Omega_{1*}\cap\Omega_{2*}=\Omega_{1*}\cap(\Omega-\Omega_1^*)=\Omega_*\cap\Omega-\Omega_{1*}\cap\Omega_1^*=\Omega_{1*}-\Omega_{1*}=\phi$.

(ⅱ)若 $\Omega_1\cup\Omega_2=\Omega$,则当 $\Omega_1\cap\Omega_2\neq\phi$ 时,$\Omega_2\supseteq\Omega-\Omega_1$,从而

$$\Omega_{2*}\supseteq(\Omega-\Omega_1)_*=(\Omega\cap(\overline{\Omega_1}))_*=\Omega_*\cap\overline{\Omega_1^*}=\Omega-\Omega_1^*.\ ()$$

(ⅲ)$bn(\Omega_1\cap\Omega_2)=(\Omega_1\cap\Omega_2)^*-(\Omega_1\cap\Omega_2)_*\subseteq\Omega_1^*\cap\Omega_2^*-\Omega_{1*}\cap\Omega_{2*}$

由[1]知:当 $\Omega_2^*=\Omega-\Omega_{1*}$ 时,$\Omega_1^*\cap\Omega_2^*=\Omega_1^*-\Omega_{1*}$,于是,由 $\Omega_{2*}=\Omega-\Omega_1^*$ 有:

$$\begin{aligned}bn(\Omega_1\cap\Omega_2)&\subseteq\Omega_1^*\cap\Omega_2^*-\Omega_{1*}\cap\Omega_{2*}\\&=\Omega_1^*-\Omega_{1*}-\Omega_{1*}\cap\Omega_{2*}\\&=bn(\Omega_1)-\Omega_{1*}\cap(\Omega-\Omega_1^*)\\&=bn(\Omega_1)-\Omega_{1*}-(\Omega_{1*}\cap\Omega_1^*)=bn(\Omega_1),\end{aligned}$$

同理:$bn(\Omega_1\cap\Omega_2)\subseteq bn(\Omega_2)$.因此 $bn(\Omega_1\cap\Omega_2)\subseteq bn(\Omega_1)\cap bn(\Omega_2)$.

3 粗糙集分类方法在模式空间划分中的应用

在对模式空间划分时,提取分类规则的过程,就是选择适当的判别函数把模式空间分为若干子区域的过程.即选一个函数 $g(x_p)=a_0+a_1x_{p1}+\cdots+x_{pn}$ 为判别函数,实际中多选 $g(x_p)$ 为线性函数,利用超平面 $g(x_p)=0$ 把模式空间分成两个决策域.但也有选判别函数为非线性函数的[3].

本文选一样本空间为$[0,1]\times[0,1]$,其样本点是在样本空间中随机生成的.如图1所示的随机生成的100个样本点,选用 $f(x)=-0.25\sin(2\pi x_1)+x_2-0.5$ 为判别函数,它严格地将样本点分成两类.其中"*"代表 $f(x)\geq0$ 的样本点,它在判别曲线 $f(x)$ 的上方,"△"代表 $f(x)<0$ 的样本点,它在判别曲线 $f(x)$ 的下方.

在实际中这样的判别函数很难准确地选取,这里我们可以用粗糙近似的方法来逼近一个恰当的判别函数.图2中,(a),(b)都是模式空间中样本点的可能性和必然性分类结果示意图,在(a)中,两条正弦曲线之外的样本点分别是 $f(x)=-0.25\sin(2\pi x_1)+x_2-0.5$ 取正值和负值的必然区域,两条正弦曲线中间所夹部分区域是 $f(x)=-0.25\sin(2\pi x_1)+x_2-0.5$ 取正值和负值的边界区域,其边界区域+必然区域=可能区域.类似的,(b)中两条直线 $x_2=0.75$,$x_2=0.25$ 之外的区域分别是 $f(x)=-0.25\sin(2\pi x_1)+x_2-0.5$ 取正值和负值的边界区域,两直线 $x_2=0.75$,和 $x_2=0.25$ 之间所夹部分为边界区域,边界区域+必然区域=可能区域.

从图2看出,分类器的选择对分类结果影响很大,(a)中的边界区域比(b)小,这说明(a)分类精度高于(b).由于(b)中的判别函数与期望目标函数 $f(x)=-0.25\sin(2\pi x_1)+x_2-0.5$ 的曲线差距较大,使得必然区域减小,可能区域增大,而(a)中的情况与此相反.由于(a)中的分类规则更接近期望目标函数 $f(x)=-0.25\sin(2\pi x_1)+x_2-0.5$,所以边界区域小,其误分率低,也就是选择($a$)中的粗糙近似曲线为分类规则时,其粗糙近似精度高;与此相反,(b)中粗糙近似曲线的粗糙近似精度低.

4 确定模式分类的必然域和可能域的算法

从图2(b)中我们容易得到两个分类规则:

(ⅰ)如果 $x_2\geq0.75$,则 $f(x)\geq0$,如果 $x_2\leq0.25$,则 $f(x)\leq0$;

(ⅱ)如果 $x_2\geq0.25$,则 $f(x)$ 有可能大于0,如果 $x_2\leq$

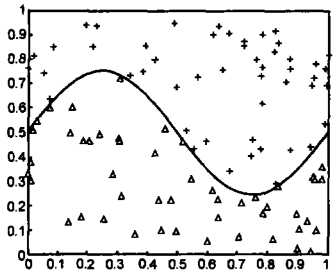

图 1　100个样本点的分类结果

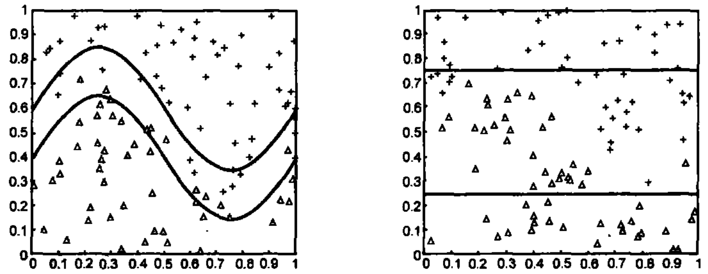

图 2　(a)100个样本点的可能性和必然性分类结果;(b)100个样本点的可能性和必然性分类结果

0.75,则 $f(x)$ 有可能小于 0.

显然规则(ⅰ)是绝对真的,规则(ⅱ)有一定程度的真.为了提高分类的准确率,我们必须改进分类规则,用神经网络的算法逼近理想的分类结果.下面我们用粗糙神经网络算法来确定样本空间中 $f(x)$ 大于 0 和小于 0 的必然性区域和可能性区域.

如果在传统神经网络的自学习算法中,利用粗糙集理论对数据进行预处理,以减少神经网络的输入单元数,这不仅可以简化网络模型和结构,而且能提高和改善神经网络的学习能力.Lingras 把粗糙集理论和传统神经网络算法结合起来形成了粗糙神经网络,它对处理粗糙模式十分有用.在其粗糙模式中的每一个值都由一对上下界值来表示,每一个粗糙神经元都用于储存输入输出上下界的[5,6].

根据粗糙集的上、近似的定义,我们来表示四个域:

(1)如果对象 p 属于组 1,则记 $x_p\in\Omega_1^*$;如果 $x_p\in\Omega_{1*}$,则对象 p 一定属于组 1.

(2)如果对象 p 属于组 2,则记 $x_p\in\Omega_2^*$;如果 $x_p\in\Omega_{2*}$,则对象 p 一定属于组 2.

设理想输出为 $\hat{O}_p=\begin{cases}1 & 如果\ p\in\Omega_1^*\\ 0 & 如果\ p\in\Omega_2^*\end{cases}$,对应输入向量的粗糙神经网络输出 $\mu_1(x)$,

假设划分模式空间 Ω 为 Ω_1 和 Ω_2 两个区域的阀值为 0.5,即分类的边界曲线为 $\mu_1(x_p)=0.5$.则 Ω_1 和 Ω_2 可以表示为 $\Omega_1=\{x\mid\mu_1(x)\geq0.5,x\in\Omega\}$,$\Omega_2=\{x\mid\mu_1(x)<0.5,x\in\Omega\}$.

令算法的迭代次数为 u,对 Ω_1 的罚因子为 1,而对 Ω_2 的罚因子为 $\omega(u)$,其中 $\omega(u)\leq1$.

关于模式 p 属于 Ω_1(或 Ω_2)的可能性和必然性算法如下[1]:

·μ^*——算法:第一步,置 $u=1$;

第二步,$\forall p=1,2,\cdots,m$,基于下列传递函数修正神经网络的加权系数和偏增量,如果模式 p 属于 Ω_1,则

$$E_p=\frac{1}{2}\sum_{p=1}^{m}(O_p-\hat{O}_p)^2;$$

如果模式 p 属于 Ω_2 则

$$E_p=\frac{1}{2}\omega(u)\sum_{p=1}^{m}(O_p-\hat{O}_p)^2$$

第三步,置 $u=u+1$,返回到第二步.

一般情况下选 $\omega(u)=a^u$,$0<a<1$ 或 $\omega(u)=\dfrac{1}{1+(cu)^b}$,$0<a,b$.根据 $\mu_1(x)$ 的定义,模式 p 属于组 1 的程度越高,则 $\mu_1(x_p)$ 越靠近 1,因此算法的迭代终止条件可设为:$1-\mu_1^*(x_p)\leq\varepsilon$,其中 ε 是一个充分小的正数.由于 Ω_2 的必然域也就是 Ω_1^* 补,故由 $\Omega_1^*=\{x\mid\mu_1^*(x)\geq0.5\}$,$x\in\Omega$ 得,$\Omega_{2*}=\{x\mid\mu_1^*(x)<0.5,x\in\Omega\}$.

类似于 μ^*—算法,可以得到刻画必然属于 Ω_1 的 μ_{1*}—算法,因篇幅所限,这里不再叙述.

如图 3 中"*"代表第一组中的模式,"+"代表第二组中的模式.在区域 $x_1\geq0.9$ 或 $x_2\geq0.9$ 的所有模式属于第一组,满足 $(x_1,x_2)\leq(0.6,0.6)$ 的区域内的模式属于第二组,上述两区域间所夹的区域也就是两组间的边界域.

根据 μ^*—算法和 μ_*—算法,可得在分类仿真结果.在图 4 中,(a)和(b)的可能域与必然域分别为

$\Omega_1^*=\{(x_1,x_2)\mid0.7\leq x_1,0.7\leq x_2\}$

$\Omega_{1*}=\{(x_1,x_2)\mid0.9\leq x_1,0.9\leq x_2\}$;

$\Omega_2^*=\{(x_1,x_2)\mid0.8\geq x_1,0.8\geq x_2\}$

$\Omega_{2*}=\{(x_1,x_2)\mid0.6\geq x_1,0.6\geq x_2\}$;

这与我们的直觉一致.

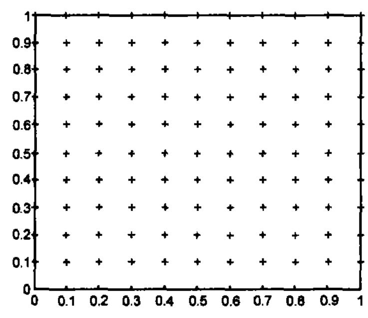

图 3

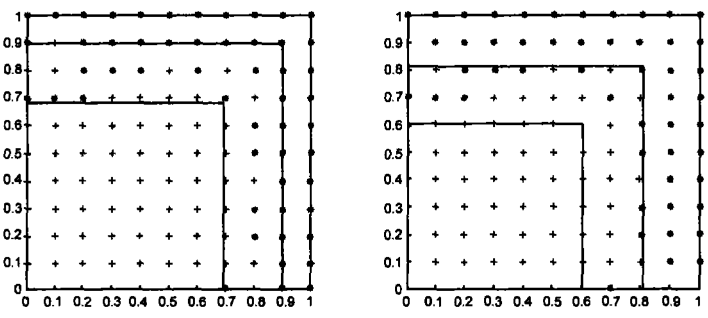

图 4

参考文献：

[1] H Ishibuchi, R Fujioka, H Tanaka. Possibility and necessity pattern classification using neural networks [J]. Fuzzy Sets and Systems, 1992, 48: 331 - 340.

[2] Pawlak Z. Rough set [J]. International Journal of Information and Computer Science, 1982, 11: 341 - 356.

[3] 边肇祺，张学工．模式识别 [M]．北京：清华大学出版社，2000年．

[4] H Ishibuchi, K Nozaki, H Tanaka. Distributed representation of fuzzy rules and its application to pattern classification [J]. Fuzzy Sets and Systems, 1992, 52: 21 - 32.

[5] P J Lingras. Rough neural networks [A]. Proceedings of Sixth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based System [C]. Granada, Spain, 1996: 1445 - 1450.

[6] P J Lingras. Comparison of neofuzzy and rough neural networks [J]. Information Sciences, 1998, 110(3-4): 207 - 215.

[7] 梅晓丹，孙圣和．神经网络的禁止搜索训练算法研究 [J]．电子学报，2001，29(12A)：1908 - 1911．

[8] 邓方安．粗糙集理论及其在多目标优化中的应用 [D]．西安电子科技大学，西安，2002

作者简介：

邓方安 男，1963 年生于陕西宁强县，陕西理工学院副教授，博士，主要研究方向为粗糙集理论及其在多目标优化中的应用，granular 计算，半群理论等．E-mail：dengfangans@263.net.

刘三阳 男，1959 年生于陕西临潼，西安电子科技大学教授，博士生导师，主要从事最优化和网络算法研究等．

徐 扬 男，1956 年生于河南获嘉县，西交通大学应用数学系教授，博士生导师，主要从事模糊逻辑和智能控制研究等．