第 38 卷 第 12 期
2004 年 12 月

西 安 交 通 大 学 学 报
JOURNAL OF XI'AN JIAOTONG UNIVERSITY

Vol. 38 №12
Dec. 2004

一种基于粗糙集的粗糙神经网络构造方法

何 明, 冯博琴, 马兆丰, 傅向华
(西安交通大学计算机科学与技术系, 710049, 西安)

摘要: 针对神经网络中各神经元和权不能赋予明确的物理意义,提出了一种基于粗糙集的神经网络模型. 该模型利用粗糙集理论数值分析的能力,首先从给定的数据集中抽取出规则,然后根据这些规则构造神经网络隐含层的神经元个数,从而确定粗糙神经网络的初始拓扑结构. 同时,将输入映射到输出子空间, 并在这个子空间上用神经网络进行逼近,由此得到一种可理解性好、收敛速度快的神经网络模型. 实验结果表明,该模型能够较好地处理神经网络拓扑结构、训练样本的大小、样本质量等对神经网络的精度及泛化能力有直接影响的问题,在大大缩短训练时间的同时,它的预测精度可达 9614 %,较同条件下径向基函数神经网络模型的精度高 316 %.

关键词: 粗糙集;神经网络;粗糙集数据分析;粗糙神经元
中图分类号: TP18 **文献标识码**: A **文章编号**: 0253 - 987X(2004)12 - 1240 - 03

Approach to Construct a Rough Neural Networks Based on Rough Set

He Ming, Feng Boqin, Ma Zhaofeng, Fu Xianghua
(Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Aiming at the problem that clear physical meanings can't be given by nerve cells and weights of neu2ral networks, a neural network model based on rough set was proposed, in which rules were extracted from giv2en training data firstly by utilizing numerical analysis ability of rough set theory, and then neurons number of the hidden layer were determined in terms of these rules to obtain the original topology of the rough neural network. Meanwhile, the input to the model was mapped into the output subspace by using rules acquired from the rough set and the expectation output could be approximated. Thus, a neural network that provides with good under2standability and rapid convergence could be constructed. Experiments show that the proposed approach can deal with problems of neural network topology architecture, sample size and quality which directly influence the gen2eralization ability and accuracy of neural network. While greatly reducing training time, the prediction precision of the model can be achieved 9614 % which is 316 % higher than RBF (radial basis function) neural network model under the same conditions.

Keywords: *rough set; neural network; rough set data analysis; rough neuron*

粗糙集理论[1]是由波兰科学家 Pawlak 在 1982 年提出的处理模糊和不确定性问题的数学工具,近年来在人工智能、机器学习、数据挖掘、模式识别和智能信息处理等领域取得了较为成功的应用[2-4].

本文将粗糙集理论与神经网络结合起来,研究了一种粗糙神经网络(RNN)构造方法. 首先利用粗糙集理论对属性进行约简,抽取精简规则,然后通过分析输入数据确定输出子空间,并在这些子空间上通过神经网络进行定量逼近,最后构建整个系统的模型. 实验结果表明,利用粗糙集理论对神经网络的输入数据进行预处理,提取其中的关键成分作为神经网络的输入以及初始规则的组成元素,可以简化

收稿日期: 2004 - 02 - 11. 作者简介: 何 明(1975～),男,博士生;冯博琴(联系人),男,教授,博士生导师. 基金项目: 国家高技术研究发展计划资助项目(2003AA1Z2610).

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

神经网络的拓扑结构,提高训练样本的质量,缩短训练时间.

1　粗糙集智能数据分析

1.1　基本概念

设一个信息系统 S 可以表示为 $S = \langle U, A, V, f \rangle$. 其中: $U = \{x_1, x_2, \cdots, x_n\}$ 是非空有限对象的集合(也称为论域); A 是属性集合; $V = \bigcup_{a \in A} V_a$, V_a 表示属性 a 的值域; $f: U \times A \to V$ 是一个信息函数,它指定 U 中每一个对象 x 的属性值,即对 $x \in U$, $a \in A$,有 $f_a(x) \in V_a$. 如果属性集 A 可以分为条件属性集 C 和决策属性集 D,即 $C \cup D = A$, $C \cap D = \S$,则该信息系统被称为决策系统或决策表,简记为 $S = (U, C, D)$,其中 D 一般只含有一个属性,记为 d.

定义 1　设 P, Q A A,如果等价关系 θ_Q 定义的每个等价类都属于等价关系 θ_P 定义的等价类,则称 P 依赖于 Q,记作 $Q \to P$. 依赖关系 $Q \to P$ 表达了如下规则:假设 $Q = \{q_1, q_2, \cdots, q_n\}$, $P = \{p_1, p_2, \cdots, p_k\}$,则对每一个 $t = \{t_1, t_2, \cdots, t_n\}$, $t_i \in Vq_i$,惟一决定了属性值集合 $s = \{s_1, s_2, \cdots, s_k\}$, $s_i \in Vp_i$,即 $(\mathrm{P}\ x \in U)[(f(x, q_1) = t_1, \cdots, f(x, q_n) = t_n)]\ (f(x, p_1) = s_1, \cdots, f(x, p_k) = s_k)]$.

定义 2　元素 $u \in U$ 在关系 R 下对集合 X 的粗糙隶属函数(RMF)[5]为

$$\mu_X^R(u): U \to [0,1] = \frac{|[u]_R \cap X|}{|[u]_R|} \tag{1}$$

式中: $|\cdot|$ 表示集合中元素的个数; $[u]_R$ 为包含元素 u 的等价类. 显然, $0 \le [u]_R \le 1$.

1.2　规则的匹配度和适用度

假设有 m 条规则,分别用 $R^1, R^2, \cdots, R^m$ 表示,若从原始数据集中抽取 m 条 $Q \to P$ 规则,则第 i 条规则 R^i 可表示为

$$\text{if } (f(x, q_1) = t_1^i, \cdots, f(x, q_n) = t_n^i) \text{ then}$$
$$(f(x, p_1) = s_1^i, \cdots, f(x, p_r) = s_r^i) \tag{2}$$

式中: $t_j^i \in V_{q_j}$; $s_k^j \in V_{p_k}$; $i = 1, 2 \cdots, m$; $j = 1, 2, \cdots, n$; $k = 1, 2, \cdots, r$.

对于一组输入 $I\{I_1, I_2, \cdots, I_n\}$,规则最大匹配函数为

$$M_i = 1 - \min \frac{\|c_{\exp} - c_i\|}{\|c_i\|}, \quad i = 1, 2, \cdots, m \tag{3}$$

式中: $c_{\exp}$ 是根据输入 $I\{I_1, I_2, \cdots, I_n\}$ 构造的条件向量; c_i 是根据粗糙集理论抽取的第 $i\ (i \le m)$ 条规则的条件向量.

从原始数据中抽取的规则可靠程度是不同的,可用粗糙隶属函数来表示规则的可靠程度,并与匹配度相结合得出规则的适用度. 根据定义 2,第 i 条规则的第 j 个属性值相对于结论等价类 X 的粗糙隶属函数为

$$\mu_X^{q_j}(t_j^i) = \frac{|X \cap [t_j^i]_{q_j}|}{|[t_j^i]_{q_j}|}, \quad i = 1, 2, \cdots m$$
$$j = 1, 2, \cdots, n, \quad X = [s_1^i, s_2^i, \cdots, s_r^i]_P \tag{4}$$

$\mu_X^{q_j}$ 越大说明由属性 t_j^i 推出结论的可能性越大,特别对于 $\mu_X^{q_j} = 1$,说明当 $f(x, q_j) = t_j^i$ 时,结论肯定成立. 输入 $I\{\cdot\}$ 对于第 i 条规则的适用度为

$$\mu_i = \max(\mu_X^{q_j}(t_j^i)\ M_i) \tag{5}$$

2　粗糙神经网络的结构和学习算法

2.1　粗糙神经网络的基本结构

本文构造的粗糙神经网络共有下述 4 层.

第 1 层:输入层,它的值为实际的精确值,表示输入向量 $x = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$.

第 2 层:隶属度函数层,分别将 n 个输入分量 $(x_1, x_2, \cdots, x_n)$ 依照某种不可分辨关系进行划分,确定其与每个相应分类之间的联系,再将每一个输入分量离散化为 r_i 个不同的值,这些值在 $[0,1]$ 之间. 与第 t 个输入节点相连接的一组神经元的作用是对输入向量的第 t 个分量进行解释. 本文定义该层神经元的作用函数为粗糙隶属函数,而神经元 N_{jt} 的输出

$$N_{jt}^{\text{out}} = \mu_{A_{jt}}(x_t) \tag{6}$$

式中: N_{jt} 是与第 t 个输入节点相连接的第 j 个神经元,其含义是与在第 t 个分量相关的分类中的第 j 个等价类; A_{jt} 为 x_t 与 N_{jt} 之间的粗糙隶属函数值连接权,含义为第 j 个类所代表的等价类.

第 3 层:推理层,该层的每个节点代表一条规则. 假设有 $m\ (m \le n)$ 条规则,该层节点的作用函数

$$\pi_i = \mu_{1i}\mu_{2i}\cdots\mu_{ni} = \prod_{j=1}^{n}\mu_{ji}, \quad 1 \le i \le m \tag{7}$$

第 4 层:清晰化层,这一层的节点代表输出变量. 在多输入单输出系统中,该层的节点数为 1,权值 ω_i 的初始值预先设为各规则粗糙隶属度值. 该层节点的输出

$$y = \sum_{i=1}^{m}\omega_i\pi_i \tag{8}$$

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd.　All rights reserved.

2.2 粗糙神经网络的分支预测算法

设 i 与 j 的上下神经元是完全连接方式，$f(x)$ 为传递函数，那么粗糙神经元 r 的输入、输出可通过下式计算得到

$$\left.\begin{aligned} I_j &= \sum \omega_{ij} o_i \\ o_{\bar{r}} &= \max(f(I_{\bar{r}} - \theta_{\bar{r}_j}), f(I_{\underline{r}} - \theta_{\underline{r}_j})) \\ o_{\underline{r}} &= \min(f(I_{\bar{r}} - \theta_{\bar{r}_j}), f(I_{\underline{r}} - \theta_{\underline{r}_j})) \end{aligned}\right\} \quad (9)$$

为简化问题，以下只讨论输出层只有一个粗糙神经元的情形. 对应于任一输入模式 k 和输出神经元 p 的实际输出 $o_p = (o_p, o_{\underline{p}})$，其中 o_p、$o_{\underline{p}}$ 分别为粗糙神经元 p 的上、下近似神经元的实际输出，而 $\hat{o}_p$、$\hat{o}_{\underline{p}}$ 为粗糙神经元 p 的上、下近似神经元的期望输出. 为了使学习以尽可能快的方式进行且误差最小，对误差的计算采用了广义的 δ 规则，而粗糙神经网络误差函数

$$E^k = \frac{1}{2}[(\hat{o}_p - o_p)^2 + (\hat{o}_{\underline{p}} - o_p)^2] \quad (10)$$

如果样本数为 m，则输入全部样本后的总误差函数

$$E = \frac{1}{2}\sum_{k=1}^{m} E^k = \frac{1}{2}\sum_{k=1}^{m}[(\hat{o}_p - o_p)^2 + (\hat{o}_{\underline{p}} - o_{\underline{p}})^2] \quad (11)$$

连接权值可通过

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \eta o_i e_j f'(I_i) \quad (12)$$

进行调整，其中 $f'(I_i)$ 是传递函数的导数，$e_j = \hat{o}_j - o_j$，η 为学习速率.

3 实验过程及结果比较

为了验证本文方法的有效性，采用文献[6]中的数据进行分析，该决策信息系统有 12 个条件属性 $A = \{S_1, S_2, \cdots, S_{12}\}$，具体步骤如下.

步骤 1：选取训练和测试数据集. 本文从编号为 1～55 的 55 组数据中选取样本号为 1，3，5，7～12，14，15，17，23，26～30，32，33，35～40，42～54 共 39 组数据作为训练数据集，其余的 16 组数据作为测试数据集.

步骤 2：离散化处理. 采用 Rosetta[7]系统对决策信息系统中的各个连续属性进行离散化处理，各属性值离散化后的取值区间为 $S_1[0,1]$，$S_2[0]$，$S_3[0,1]$，$S_4[0]$，$S_5[0,1,2]$，$S_6[0,1,2]$，$S_7[0]$，$S_8[0,1,2]$，$S_9[0]$，$S_{10}[0,1]$，$S_{11}[0,1,2]$，$S_{12}[0,1,2]$，$d[0,1]$.

步骤 3：对步骤 1 离散化后的结果进行数据过滤，消除对所有规则都不必要的属性. 通过约简，得到属性核$\{S_6\}$，并且获得 3 个最小条件属性集$\{S_5, S_6\}$、$\{S_6, S_8\}$、$\{S_6, S_{11}\}$和一个决策属性 d. 本文抽取下述 5 条确定性规则.

规则 1：$(S_6 = 2) \rightarrow (d = 0)$；[9，9，40191 %，100100 %] [9，0] [{2，12，21，29，30，32，35，36，37}，{}].

规则 2：$(S_{12} = 0) \rightarrow (d = 0)$；[11，11，50100 %，100100 %] [11，0] [{1，4，5，6，7，10，13，14，18，19，31}，{}].

规则 3：$(S_5 = 2) \rightarrow (d = 0)$；[8，8，36136 %，100100 %] [8，0] [{2，12，21，22，29，30，35，36}，{}].

规则 4：$(S_6 = 0) \rightarrow (d = 0)$；[10，10，45145 %，100100 %] [10，0] [{4，5，6，7，10，13，14，18，31，34}，{}].

规则 5：$(S_5 = 1) \& (S_6 = 1) \rightarrow (d = 1)$；[17，17，100100 %，100100 %] [0，17] [{}，{3，8，9，11，15，16，17，20，23，24，25，26，27，28，33，38，39}].

再根据式(3)计算各条规则的粗糙隶属度.

步骤 4：将剩余的 16 组测试数据进行分类，并与一个径向基函数(RBF)神经网络(输入层 10 个节点、隐含层 5 个节点和输出层 1 个节点)的分类结果和性能进行比较，结果见表 1 和表 2.

表 1 测试数据集分类结果比较

测试数据集编号	粗糙神经网络		RBF 神经网络		测试数据集编号	粗糙神经网络		RBF 神经网络	
	绝对误差	相对误差/ %	绝对误差	相对误差/ %		绝对误差	相对误差/ %	绝对误差	相对误差/ %
2	01057 6	5100	0115	10100	21	01001 2	0110	0108	5133
4	01051 6	4148	0106	4100	22	01076 0	6160	0110	6168
6	01104 0	9100	0109	6100	24	01001 2	0110	0109	6100
13	01001 2	0110	0103	2100	25	01032 0	2178	0104	2167
16	01106 0	9120	0107	4167	31	01036 5	3117	0106	4100
18	01018 7	1162	0104	2167	34	01012 6	1109	0103	2100
19	01032 5	2182	0120	1133	41	01002 5	0122	0102	1133
20	01046 0	3199	0106	4100	45	01001 2	0110	0103	2100

(下转第 1246 页)

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

从图 1b 和图 2b 中可以看出,原始图像经轮廓提取后均得到清晰的轮廓图,可顺利进行直线提取. 由于图像中机场跑道分别呈水平和倾斜状态,直线模板的角度也因此作相应变化,分别取 0°和 45°. 最终识别完成后,在跑道中心用十字线表示,见图 1d 和图 2d. 对呈垂直角度(90°)的跑道识别亦可得到满意的结果(图略).

实验结果显示,本文提出的算法可以有效识别不同角度的机场跑道.

3 结束语

本文提出了一种识别雷达图像中机场跑道的方法. 首先提取机场轮廓,然后使用 Radon 变换及逆变换提取直线,再使用直线模板滤除不符合条件的短线,最后正确识别出机场跑道. 该方法的特点是识别能力强,即使图像中有较多杂乱的背景,也能迅速识别出正确目标. 实验结果表明,本文所提方法能够准确识别出不同角度的机场跑道.

参考文献:

[1] Han J W, Euo L, Bao Y S. A method of automatic finding airport runways in aerial images [A]. Proc of 6th International Conference on Signal Processing [C]. Beijing, China: People's Posts and Telecommunications Publishing House, 2002. 731 - 734.

[2] Huerlas A, Cole W, Nevatia R. Detecting runways in complex airport scenes [J]. Computer Vision Graphics and Image Processing, 1990, 51(2): 107 - 145.

[3] Venkateswar V, Chellappa R. Extraction of straight lines in aerial images [J]. IEEE Trans on Pattern Analysis and Machine Intellegence, 1992, 14(11): 1 111 - 1 114.

[4] 罗 军,杨卫平,沈振康. 红外图像中机场跑道的自动目标识别 [J]. 红外技术,2003,25(3): 13 - 17.

[5] 何国威,刘建国,周春晓,等. 基于知识的红外数字图像中铁路的识别 [J]. 红外与激光工程,2000,29(6): 55 - 59.

[6] Finch I, Antonacopoulos A. Identification of airfield runways in synthetic aperture radar images [A]. Proc of 14th Int Conf on Pattern Recognition [C]. Los Alamitos, USA: IEEE-CS Press, 1998. 1 633 - 1 636.

(编辑 苗 凌)

(上接第 1242 页)

表 2 测试数据集主要性能比较

	平均误差	最大相对误差/ %	训练次数	计算时间 / min
RBF 神经网络	01072	10100	> 2 000	50
粗糙神经网络	01036	9120	< 100	3

从表 1 和表 2 中可以看出,本文提出的基于粗糙集的神经网络模型,在训练时间和预测精度上较单一的 RBF 神经网络模型要优越.

4 结 论

本文提出了一种基于粗糙集的神经网络模型,该模型综合了粗糙集理论在知识获取方面的能力和神经网络在数值逼近上的优势. 通过粗糙集智能数据分析,可消除初始决策表中的冗余信息和噪声数据的干扰,减少了粗糙神经网络中输入层和代表规则层的神经元个数,简化了神经网络的拓扑结构,减少了训练所需的计算量和时间,提高了模型的正确率.

参考文献:

[1] Pawalk Z. Rough sets [J]. International Journal of Computer and Information Science, 1982, 11(5): 341 - 356.

[2] Pawalk Z. Rough sets [M]. Norwell, Netherlands: Kluwer Academic Publisher, 1991.

[3] Skowron A. Rough sets and Boolean reasoning [M]. New York: Physica Verlag, 2001. 95 - 124.

[4] Skowron A, Peters J. Rough sets: trends and challenges [A]. Proceedings of the 9th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing [C]. Berlin: Springer Verlag, 2003. 25 - 34.

[5] Wu Y W, Zhang C N. A rough neural network for material proportioning system [A]. IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions, Chengdu, China, 2002.

[6] Ryszard N. Evaluation of vibroacoustic diagnostic symptoms by means of the rough set theory [J]. Computers in Industry, 1992, 20(2): 141 - 152.

[7] Ohrn A. Rosetta software system [EB/OL]. http://rosetta.lcb.uu.se/, 2003 - 12 - 08.

(编辑 苗 凌)

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.