

基于概率神经网络的广东省土地资源评价

薛月菊¹, 胡月明², 杨敬锋¹, 陈其昌¹

(¹华南农业大学工程学院, 广东 广州 510642; ²华南农业大学信息学院, 广东 广州 510642)

摘要: 用概率神经网络的方法, 以广东省土地资源为对象进行了土地资源评价的研究, 并与BP网络的评价结果作了比较。结果表明, 用概率神经网络进行评价的结果比BP网络的评价结果更加准确, 更加切合实际情况, 一定程度上避免了人为因素的干扰, 提高了土地资源评价的准确性。

关键词: 概率神经网络; 土地资源评价; BP网络

中图分类号: TP181 文献标识码: A 文章编号: 1001-411X(2006)03-0108-03

Guangdong Province Land Resource Evaluation Based

on Probabilistic Neural Networks

XUE Yue-ju¹, HU Yue-ming², YANG Jing-feng¹, CHEN Qi-chang¹

(¹ College of Engineering, South China Agric. Univ., Guangzhou 510642, China;

² College of Information, South China Agric. Univ., Guangzhou 510642, China)

Abstract: A method of probabilistic neural network (PNN) was used to evaluate Guangdong Province land resource. The obtained results of PNN were compared with those of BP neural networks (BPNN), indicating that the PNN method is more accurate than the BPNN, with less influence of human factors and improved validity.

Key words: probabilistic neural networks; land resource evaluation; BP neural networks

传统的土壤资源评价主要采用指数和法、指数积开方法、模糊综合评判法、灰关联综合分析法等方法^[1-5]。神经网络有表示任意非线性关系和学习能力, 将它应用于土地资源评价, 可以克服传统土地资源评价方法所面临的实际困难。一些利用神经网络研究土地资源评价的方法被相继提出^[6-8], 概率神经网络是基于统计原理的神经网络模型, 在分类功能上与最优 Bayes 分类器等价^[9], 其实质是基于贝叶斯最小风险准则发展而来的一种并行算法, 同时它不像传统的多层前向网络那样需要用BP算法进行反向误差传播的计算, 而是完全前向的计算过程, 它训练时间短, 不易收敛到局部最优, 而且它的分类正确率高, 无论分类问题多么复杂, 只要有足够多的训练数据, 可以保证获得

贝叶斯准则下的最优解。因此, 对于分类问题, 概率神经网络表现出良好的性能。而土地资源评价实质上也是一种分类问题, 概率神经网络模型在土地资源评价中的应用鲜见报道, 本文将概率神经网络模型引入土地资源评价, 利用该方法对广东省的土地样本予以评价, 并将其结果与BP网络的评价结果作了比较。

1 材料与方法

以广东省第2次土壤普查成果资料为主要数据源, 结合广东省统计年鉴等社会经济资料, 建立土壤资源数据库^[10]。

根据广东省土壤资源的实际特点和前人的研究成果^[10-11], 采用高程、地面坡度、土壤质地、土壤酸碱

收稿日期: 2005-03-02

作者简介: 薛月菊(1969-), 女, 副教授, 博士; 通讯作者: 胡月明(1964-), 男, 教授, 博士, E-mail: ymhu@scau.edu.cn

基金项目: 广东省自然科学基金(04300504, 980150); 国家科技攻关项目(2002BA516A08); 广东省科技计划项目(2005B20701008)

度、土壤有机质含量、土地利用类型等6个指标作为评价因子。所有评价因子的属性数据均从已经建立的广东省土壤资源数据库中获取^[10],有关农用地的样本,共有16 652条记录。

所获取的各属性的数据中,部分为连续属性(如高程、地面坡度、土壤酸碱度和土壤有机质含量),部

分为标称属性(如土地利用类型和土壤质地)。依据《土壤资源信息系统的建立与应用》文献的资料^[10],将连续属性划分为不同的区段值,并将各区段用质量等级来表示,例如分级为1表示该土壤质量指标最优,分级为5表示最差;将标称属性也用相应的等级来表示(表1)。

表1 各评价因素的分级
Tab.1 The gradation of evaluation factors

高程 altitude/m	地面坡度 slope/(°)	土壤质地 soil texture	w(土壤有机质 soil organic material)/%	pH	土地利用类型 land utilization type	等级 grade
<200	<3	轻壤	>4	6.5~7.5	耕地	1
200~500	3~6	沙壤、中壤	3~4	5.5~6.5 或 7.5~8.5	园地	2
500~800	6~15	重壤、轻粘	1.0~3.0	8.5~9.0	林地	3
800~1 000	15~25	中粘	0.6~1.0	4.5~5.5	牧草地	4
>1 000	>25	沙土	<0.6	<4.5	滩涂、未利用地	5

采用概率神经网络进行模式分类。确定 X 多维向量所属的类别 C_n ,假设已知各类别的概率密度函数 $f_n(X)$,根据 Bayes 分类规则,假如下列不等式成立, X 将被分到 C_i 类:

$$P_i L_i f_i(X) > P_j L_j f_j(X), i \neq j \quad (1)$$

其中 P_i 为 X 属于 C_i 类的先验概率, L_i 为 C_i 类的错

误分类的代价。

概率神经网络是统计方法与前馈神经网络相结合的一种神经网络模型,它利用一个指数函数来代替S型函数,是一个有2个隐层的网络(图1)^[9]。该网络结构的第1层是输入层,它将输入样本完全不变地传给下一层;第2层是模式层,各模式单元有相

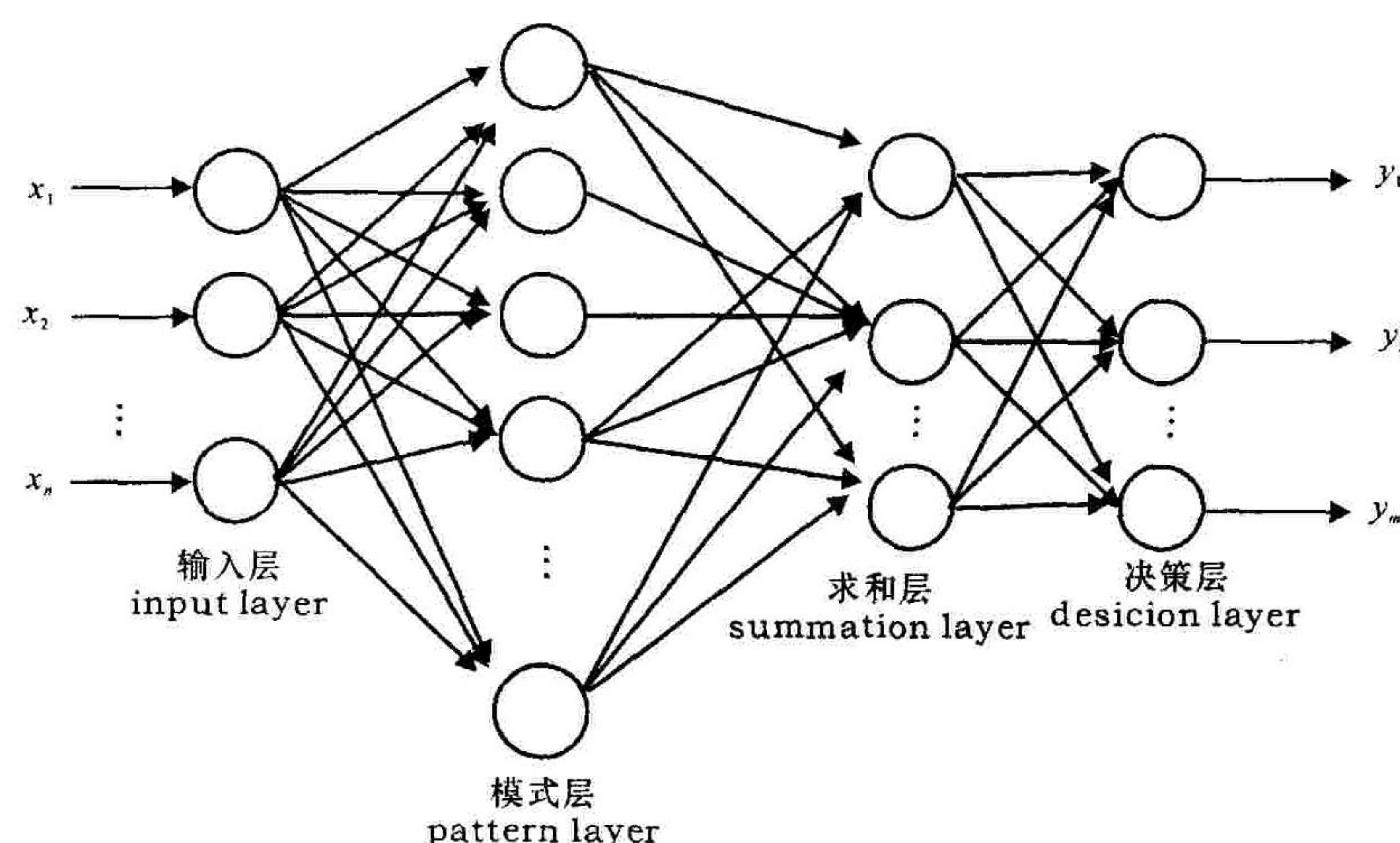


图1 概率神经网络

Fig.1 Probabilistic neural networks

同的输入,通常模式单元个数等于训练样本数,该层每个模式单元的 outputs 为

$$f(X, W_i) = \exp\left[-\frac{(X - W_i)^T(X - W_i)}{2\delta^2}\right], \quad (2)$$

其中, W_i 为输入层到模式层连接的权值。 δ 为平滑因子,它对分类的起着至关重要的作用;第3层是累加层,是将属于某类的概率累计,按下式计算

$$f_i(X) = \frac{1}{2\pi^{n/2}\delta^n k_i} \sum_{j=1}^{k_i} \exp\left[-\frac{(X - X_{ij})^T(X - X_{ij})}{2\delta^2}\right], \quad (3)$$

其中, k_i 为属于类 C_i 的模式样本个数, X_{ij} 为属于类别 C_i 的第 j 个样本,从而得到输入样本属于 C_i 类的最大可能性,通常求和层单元数等于类别数;第4层是决策层,作用是实现(1)式。概率神经网络具有

Bayes 后验概率输出的特点. 对于传统的概率神经网络, 它的权值是不需要训练的, 输入层与模式层的权值被设置为各类训练样本.

2 结果与分析

根据广东省土地资源的实际情况, 土壤质量被分成 4 个等级. 由于通过实际调查验证模糊变权方法的评价结果较贴切实际^[10,12], 因此本文以模糊变权评价方法评价结果作为神经网络训练的导师信号, 由于数据库中的记录较多, 分别从各等级的样本中随机地抽取 10% 的样本作为训练样本, 其他的样本作测试样本. 神经网络的输入为高程、地面坡度、土壤质地、土壤酸碱度、土壤有机含量、土地利用类型等 6 个评价因子, 则土地资源评价的神经网络为 6 个输入单元、1 665 个模式层单元、4 个求和层单元、4 个输出单元的神经网络. 其中的平滑因子 $\delta =$

0.65, 并将各类的输入样本作为输入层与模式层的权值, 从而得到土地资源评价的概率神经网络.

首先利用概率神经网络对测试集进行评价, 正确评价率为 92.762%; 然后用其对所有 16 652 个样本进行评价, 结果见表 2. 从表 2 可以看出, 概率神经网络对土地资源评价的结果比 BP 网络的评价结果^[10]更加符合土地的实际情况. 这是由于概率神经网络具有与 Bayes 分类器等价的优良分类性能, 而 BP 网络会由于训练算法和结构选择等原因引起训练结果的局部最优, 所以在利用它们进行分类时, 概率神经网络的效果优于 BP 网络的结果, 这是有理论依据的. 另外, 由于概率神经网络的权值可直接由训练样本设定, 权值无需训练, 更无需反向误差传播计算训练过程. 而 BP 网络中权值的训练往往速度较慢, 因此概率神经网络的权值确定也较 BP 网络更节省时间.

表 2 广东土壤质量评价结果

Tab. 2 The results of land resource evaluation for Guangdong province

分类 type	实际情况 practical situation ^[10,12]		BP 网络 BPNN ^[10]		概率神经网络 PNN	
	面积 area/km ²	占比 rate/%	面积 area/km ²	占比 rate/%	面积 area/km ²	占比 rate/%
I	25 630.987	14.760	30 585.537	17.613	21 776.383	12.540
II	42 375.034	24.402	39 173.510	22.558	45 098.506	25.970
III	76 495.022	44.050	68 644.663	39.530	79 819.652	45.965
IV	29 152.974	16.788	35 250.307	20.299	26 959.473	15.525
合计 total	173 654.014	100.000	173 654.014	100.000	173 654.014	100.000

3 结语

利用概率神经网络, 对广东省的土地样本评价的结果表明, 该方法比 BP 网络更加有效. 因此, 利用概率神经网络评价土地不失为一个可行的方法. 下一步, 我们将进一步研究土地样本的属性部分缺失的情况下, 如何利用概率神经网络进行土地资源评价.

参考文献:

- [1] 倪绍祥. 近 10 年来中国土地评价研究的进展[J]. 自然资源学报, 2003, 18(6): 672-683.
- [2] 张萍, 刘高焕, 邢立新, 等. 农业土地资源动态评价模型研究[J]. 国土资源遥感, 2000(1): 51-56.
- [3] 王令超, 王国强, 王国灵, 等. 农用地定级的总分值计算模型研究[J]. 地域研究与开发, 2001, 20(3): 10-12.
- [4] 陈良. 模糊综合评判在农用地评价中的应用[J]. 甘肃教育学院学报: 自然科学版, 1999, 13(3): 22-25.
- [5] HU Yue-ming, DAI Jun, WANG Ren-chao. GIS-based red

soil resources classification and evaluation [J]. Pedosphere, 1999, 9(2): 131-138.

- [6] 付强, 王志良, 梁川. 自组织竞争人工神经网络在土壤分类中的应用[J]. 水土保持通报, 2002, 22(1): 39-43.
- [7] 焦利民, 刘耀林. 土地适宜性评价的模糊神经网络模型[J]. 武汉大学学报: 信息科学版, 2004, 29(6): 513-516.
- [8] 唐南奇, 谭明军. 基于人工神经网络农用地分等研究: I 分等模型与精度检测[J]. 福建农林大学学报: 自然科学版, 2004, 33(2): 32-36.
- [9] ROMERO R, TOURTZKY D. Probabilistic neural networks classifiers for Optical chinese recognition [J]. Pattern Recognition, 1997, 8(30): 1 279-1 292.
- [10] 胡月明. 土壤资源信息系统的建立与应用[M]. 北京: 中国科学技术出版社, 2002: 10.
- [11] 广东省土壤普查办公室. 广东土壤[M]. 北京: 科学出版社, 1993: 4.
- [12] 胡月明, 万洪富, 吴志峰, 等. 基于 GIS 的土壤质量模糊变权评价[J]. 土壤学报, 2001, 28(3): 266-274.

【责任编辑 周志红】