

层次分析的神经网络集成方法

陈文宇¹, 刘井波^{1,2}, 孙世新¹

(1. 电子科技大学计算机科学与工程学院 成都 610054; 2. 重庆三峡学院数学与计算机科学学院 重庆 万州区 404000)

【摘要】提出了基于层次分析法的神经网络集成方法;介绍了四种不同类型的神经网络,并用这四种不同类型的神经网络构造了五个独立并行处理的模式识别子系统,对每个子系统进行独立训练并分别进行模式识别,统计每一输入模式被各子系统识别成的模式及次数,按识别次数进行两两成对比较;计算各输入模式的识别结果与典型故障模式识别结果间的期望与方差的大小,进行两两成对比较并进行一致性检验。选取层次分析法中准则层对目标层的权重,综合各神经网络子系统的识别情况与输入模式和目标模式的差异大小计算输入模式对各典型故障模式的整体权值并以此作出决策。使用matlab软件进行仿真计算,得到的故障正确识别率。层次分析法神经网络集成方法取得了比相对多数集成方法好的识别效果,还可集成其他模式的识别方法。

关键词 层次分析法; 集成; 泛化; 神经网络
中图分类号 TP 183 文献标识码 A

Method for Neural Network Ensemble Based on Analytical Hierarchy Process

CHEN Wen-yu¹, LIU Jing-bo^{1,2}, and SUN Shi-xin¹

(1. School of Computer Science and Engineering, University of Electronic Science and Technology of China Chengdu 610054;

2. School of Mathematics and Computer Science, Chongqing Three Gorges University Wanzhou Chongqing 404000)

Abstract This paper presents a method based on analytical hierarchy process (AHP) for neural network ensemble. In the method, five subsystems are constructed. Those neural networks compute parallel which have been trained solely. The decisions are made. Correct classification rates are obtained from the mat lab simulation. In experiments, the proposed methods have been successfully evaluated using thirteen different datasets, it is more effective than the relative majority voting scheme, it can be applicable to other method of pattern recognition.

Key words analytical hierarchy process; ensemble; generalization; neural network

神经网络集成(neural network ensemble)方法可以合成多个神经网络的训练和仿真结果,极大地提高神经网络系统的泛化能力。神经网络集成主要进行构造神经网络子系统和集成子系统的结果两个方面的工作。构造子系统最重要的技术可参见文献[1]。对子系统结果的集成技术^[2]有少数投票法、简单平均值法、加权平均值法、基于遗传算法的进化集成技术等。神经网络的泛化性能由集成中个体神经网络的精度和差异性决定,常用相同类型的有监督学习神经网络作为个体神经网络。个体神经网络子系统的生成方法^[3]为:对同种神经网络采用不同的初始权重、不同的层次结构、不同的训练数据、不同的训练算法。本文采用不同种类、不同学习方式(有教师、无教师)的神经网络为子系统,结合输入模式与目标模式运用层次分析法进行模式集成。

1 基本方法与原理

1.1 SOM神经网络

文献[4]提出了自组织特征映射网络(SOM)。SOM模拟大脑中具有相同功能的脑细胞聚集在一起的特性。SOM学习完毕后,输出处理单元相邻者会具有相似的功能。当神经网络接受外界输入模式时,神经元会分成不同的反应区域,各区域对输入模式具有不同的响应特性。反应区域由输入层、竞争层组成,是无教师自组织网络^[4]。第1层维数同输入向量;等2层为 $m \times n$ 的神经元阵列。两层间是全连接的。SOM网络的学习过程如下:

(1) 初始化。对输入层到竞争层的连接权值 $w_{ij}(0)$ 赋予较小的权值,对邻域函数 $s_j(0)$ 赋值。

(2) 提供新的输入向量 x 。

收稿日期: 2007-07-06; 修回日期: 2008-01-10

基金项目: 国家自然科学基金(60471055)

作者简介: 陈文宇(1968-), 男, 博士生, 副教授, 主要从事模式识别、编译技术、自动机理论等方面的研究。

(3) 计算输入向量与每个输出神经元 J 之间的距离 d_j , 求出最小者 $d_k = \min(d_j)$, 其中 $d_j = \|X -$

$$W_j\| = \sqrt{\sum_{i=1}^N (x_i(t) - w_{ij}(t))^2}。$$

(4) 给出随时间变化而变小的邻域函数 $S_k(t)$ 。

(5) 调整权值 $w_{ij}(t+1) = w_{ij}(t) + \eta(t)[x_i(t) - w_{ij}(t)]$,

其中 $\eta(t) = 1/t$ 或 $\eta(t) = 0.2(1 - t/10\ 000)$ 。

(6) 计算输出 $O_k = f(\min\|X - W_j\|)$, 其中 f 函数为0-1函数或其他非线性函数。

(7) 如达到训练总次数, 输出 O_k , 否则返回过程(2)。

2.2 概率神经网络

文献[5]提出的概率神经网络(PNN)属于有监督的前馈(feedforward)神经网络, 由输入层、模式层、求和层、输出层组成。输入层与模式层为全连接; 模式层每一个神经元表示一个训练样本; 求和层的每个神经元表示一个目标模式。若训练样本 i 属于目标模式 j , 则模式层神经元 i 到求和层神经元 j 间存在连接。输出层对求和层使用WTA (winner-take-all)方法得到输出。PNN计算过程如下:

(1) 构造神经网络。输入层神经个数与输入向量维数相同; 模式层神经个数与训练样本数相同; 求和层神经个数与目标模式个数相同。根据训练样本所属目标模式建立模式层与求和层间的连接。

(2) 计算输入模式向量对 q 种目标模式 A^q 的概率密度函数 $f_{A^q}(x) = P(x|A^q)$, 其中 x 为待分类的输入数据; $f_{A^q} = \frac{1}{(2\pi)^{0.5 \times p}} \frac{1}{n_i} \sum_{i=1}^{n_i} \exp((x'X_{A^q} - 1)/(2\sigma^2))$; p 为输入数据维度; n_i 为训练样本个数; X_{A^q} 为类别 A^q 中的第 i 个训练数据; σ 为平滑系数。

(3) 输出 $O = \max(f_{A^q}(x))$ 。

1.3 径向基函数神经网络

径向基函数(RBF)神经网络^[6]是一种性能良好的前向神经网络结构, 它具有全监督式和全局逼近的性质^[7], 是由输入层、隐含层、输出层构成的3层前馈网络, 隐含层采用径向基函数 $f(n) = e^{-n^2}$ 作为激励函数。RBF网络的训练学习步骤为: (1) 无教师学习以确定输入层与隐含层间的权值 w_1 。(2) 有教师学习, 确定隐含层与输出层间的权值 w_2 。

1.3.1 第一步学习步骤

(1) 初始化, 随机给出权值 $w_{1i}(0)$ 的值。

(2) 相似度计算, 求与训练样本最接近的权值

$$d_r(t) = \min(\|x(t) - w_{1i}(t-1)\|)。$$

(3) 权值调整为:

$$w_{1i}(t) = w_{1i}(t-1) \quad 1 \leq i \leq h; i \neq r$$

$$w_{1r}(t) = w_{1r}(t-1) + \beta(x(t) - w_{1r}(t-1)) \quad i = r$$

式中 h 为隐含层神经元数; β 为学习率, $0 < \beta < 1$ 。

(4) 令 $t = t + 1$, 返回(2)进行计算, 直到 $\|w_{1i}(t) - w_{1i}(t-1)\| < \varepsilon$ 时为止。

1.3.2 第二步学习步骤

r^q 为隐含层对第 q 个输入模式计算后的输出。

(1) 初始化, 对 w_{2ij} 赋予 $[-1, 1]$ 内的随机值。

(2) 计算网络的输出层输出 $y^q = \sum_{i=1}^h r_i^q \times w_{2i}$ 。

(3) 计算输出与目标的差 $d_j^q = t_j^q - y_j^q$, 其中 $j = 1, 2, \dots, s$, s 为输出神经元数。

(4) 权值调整 $w_{2ij}(t+1) = w_{2ij}(t) + \alpha y_i^q d_i^q$, 其中 $i = 1, 2, \dots, h$ 。

(5) 取下一个训练模式重复(2)~(4), 直到 d_j^q 变得足够小为止。

1.4 向量量化神经网络^[8]

向量量化(LVQ)神经网络是有监督的训练竞争层的方法, 由输入层、竞争层、输出层组成。输入层与竞争层间为完全连接, 竞争层自组织学习并对输入向量进行分类, 结果只与输入向量间的欧氏距离有关; 竞争层与输出层间为部分连接, 每个输出神经元与竞争层神经元的不同组连接权值固定为1。竞争层将输入向量分成多个子类别; 输出层将子类别合并而形成目标分类模式。

令 x^i 代表第 i 个训练模式矢量; C_{x^i} 代表输入向量 x^i 所属的模式类别; C_j 为第 j 个输出神经元代表的类别。LVQ 网络学习算法如下。

(1) 初始化, 随机地进行初始化权值向量 $w_j(0)$, 选取学习率 η 的值, $k = 0$ 。

(2) 检查训练结束条件, 不满足继续, 否则退出。

(3) 对每个训练样本 x^i 进行(4)~(5)两个步骤。

(4) 确定获胜权值向量 w_q , 使满足:

$$d(x^i, w_q) = \min_{\forall j} d(x^i, w_j) = \min_{\forall j} \|x^i - w_j\|_2^2 \quad j = 1, 2, \dots, n$$

(5) 更新权值 w_q :

$$w_q(k+1) = w_q(k) + \eta(k)(x^i - w_q(k)) \quad C_{w_q} = C_{x^i}$$

$$w_q(k+1) = w_q(k) - \eta(k)(x^i - w_q(k)) \quad C_{w_q} \neq C_{x^i}$$

式中 $0 < \eta(k) < 1$ 。

(6) 令 $k = k + 1$, 减小学习率 $\eta(t)$, 返回(2)。

1.5 层次分析法

文献[9]创立的层次分析法(AHP)保证了定性分析的科学性和定量分析的精确性。在决策类问题中必须考虑多方面的因素或判断准则，最终通过判断准则作出决策。层次分析法把问题条理化、层次化，构造出一个如图2所示的三层结构模型。(1) 目标层：决策问题的唯一目标或理想结果。(2) 准则层：包含为实现目标所需考虑的所有准则、子准则。(3) 方案层：包括可选的各种措施、决策方案等。

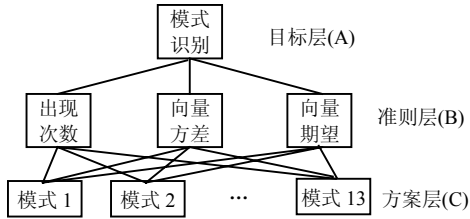


图2 层次分析法集成模型

两两成对比较确定各因素对同一目标的影响。比较 $\{y_1, y_2, \dots, y_n\}$ 各因素对同一目标的影响，每次取两个因素 y_i 和 y_j ， a_{ij} 表示 y_i 与 y_j 对目标的影响程度之比； $a_{ji} = 1/a_{ij}$ ， a_{ij} 取值如表1所示。

表1 a_{ij} 取值表

取值	含意	取值	含意
1	相同重要	7	强烈重要
3	稍微重要	9	极端重要
5	明显重要	2, 4, 6, 8	取中间值

定义 1 若矩阵 $A=(a_{ij})_{n \times n}$ 满足 $a_{ij} > 0, a_{ji} = 1/a_{ij}$ ，则称 A 为正互反矩阵，其中 $i, j=1, 2, \dots, n$

定义 2 满足关系式 $a_{ij}a_{jk} = a_{ik}$ 的正互反矩阵称为一致矩阵，其中 $\forall i, j, k=1, 2, \dots, n$ 。

一致性检验步骤如下。(1) 计算一致性指标 $CI = (\lambda_{max} - n) / (n - 1)$ 。(2) 查找一致性指标 RI，如表1所示。(3) 计算一致性比例 $CR = CI / RI$ 。当 $CR < 0.1$ 时，矩阵的一致性是可以接受的，否则应对判断矩阵作修正。(4) 层次总排序、一致性检查。上一层(A层)包含 A_1, A_2, \dots, A_m 共 m 个因素，层次总排序权重分别为 a_1, a_2, \dots, a_m 。下一层(B层)包含 n 个因素 B_1, B_2, \dots, B_n ，它们关于 A_j 的层次单排序权重为 $b_{1j}, b_{2j}, \dots, b_{nj}$ 。B 层中各因素关于总目标的权重

b_1, b_2, \dots, b_n 计算公式为 $b_i = \sum_{j=1}^m b_{ij} a_j$ ，其中， $i=1, 2, \dots, n$ ；B 层总排序随机一致性指标为

$$CR = \left(\sum_{j=1}^m CI(j) a_j \right) / \left(\sum_{j=1}^m RI(j) a_j \right)$$

当 $CR < 0.1$ 时，矩阵的一致性是可以接受的；否则应对判断矩阵作出修正。以总体权重最大值对应的方案作为决策。

表2 一致性指标RI

n	1	2	3	4	5	6	7	8	9	10	11	12	13
RI	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54	1.56

2 集成系统模型

集成系统中各子系统的差异性越大其集成的泛化能力越好。为了使各子系统相互独立且模式识别互不相关，本文使用有教师学习、无教师学习的不同类型的神经网络作为识别子系统，各子系统的识别模式经模式映射标准化后用层次分析法进行集成。神经网络是稳定的^[10]。

有教师学习的神经网络识别结果是唯一的，无教师学习的神经网络训练完成后识别结果是不确定的。神经网络训练完后用典型的故障对无教师学习的子系统进行仿真得到映射关系 $y, y: I_i \rightarrow O_j^{(i)}$ ；

I_i 为典型故障输入； $O_j^{(i)} (i=1, 2, \dots, m; j=1, 2, \dots, k)$ 为 m 个子系统的输出，它是一一对应的，其逆映射记为 y^{-1} ；各子系统经模式映射后的的识别输出为 $y^{-1}(x) = I_i$ ，其中 $I_i \rightarrow O_j^{(i)}$ 且满足：

$$\|x - O_j^{(i)}\| = \min(\|x - O_z^{(i)}\|) \quad z=1, 2, \dots, k$$

2.1 层次分析集成模型

神经网络子系统的识别模式出现的次数能一定程度反映输入向量的模式，输入向量与目标向量之差的期望与方差也能反应输入向量的模式。

构造层次分析法集成模型如图3所示。方案层为需识别出的模式，准则层为子系统识别的模式出现次数、输入模式与目标模式之差向量的方差和期望。

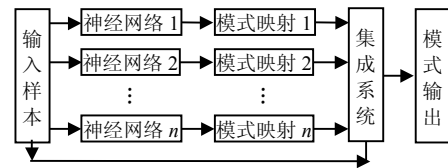


图3 集成系统框架图

2.2 集成计算过程

(1) 构造各神经网络子系统，独立地对每个子系统进行训练，计算子系统的模式映射。分别对输入模式向量进行模式识别，统计识别模式出现的次数。

(2) 计算输入模式与目标模式之差向量的方差和期望。(3) 分别构造模式出现次数、差向量的方差与期望的两两成对比较矩阵。(4) 计算成对比较矩阵的最大特征值对应的特征向量并归一化，对每个准则进行一致性检验。(5) 计算方案层模式对目标层的总体权值向量，并进行总体一致性检验；根据总体权值向量排序得到最终识别模式。

3 实验

使用本系统对某设备的典型故障样本进行训练和仿真。典型故障样本如表3所示, 其中 x 为征兆集; y 为故障集。

表3 典型故障样本

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
y_1	0.25	0.60	0.25	0.50	0.50	0.25	0.50	0.25
y_2	0.25	0.60	0.5	0.50	0.50	0.75	0.50	0.50
y_3	0.75	0.40	0.75	0.50	0.50	0.50	0.50	0.50
y_4	0.25	0.60	0.50	0.25	0.25	0.75	0.50	0.50
y_5	0.25	0.60	0.50	0.00	0.00	1.00	0.50	0.50
y_6	0.75	0.40	0.75	0.25	0.25	0.75	1.00	0.25
y_7	0.25	0.60	0.50	0.50	0.50	0.75	0.75	0.50
y_8	0.75	0.40	0.50	0.75	0.75	0.25	0.00	1.00
y_9	0.25	0.60	0.50	0.75	0.25	0.25	0.50	0.50
y_{10}	0.25	0.60	0.50	0.25	0.25	0.75	1.00	0.50
y_{11}	0.50	0.50	0.50	0.75	0.75	0.25	1.00	0.50
y_{12}	0.00	0.75	0.75	1.00	0.75	0.25	1.00	0.50
y_{13}	0.50	0.50	0.50	0.75	0.75	0.25	0.50	0.50

在典型故障数据中加入方差为0.05、0.08、0.12、0.20、0.30、0.40、0.50的随机噪声信号, 得到7组测试数据。每种故障100个, 测试数据为 8×1300 的矩阵。用4种不同的神经网络构造5个识别子系统: 第1个子系统为SOM, 输出层为 5×5 的阵列, 训练样本为典型故障数据, 共训练100次; 第2个子系统为PNN, 训练样本为13种典型故障数据, 共训练100次, 径向基函数的分布密度为0.1; 第3、4个子系统为参数不同的两个LVQ神经网络, SPREAD为1.0、20, 隐层神经元数为30、20, 训练样本为13种典型故障数据, 训练300次; 第5个子系统为RBF, 训练误差平方和目标为0.01, SPREAD为40, 训练样本为每种故障选10个共130个。每组测试数据计算10次, 得到模式正确识别率的平均值。层次分析法中准则层对目标层的权重为0.3、0.3、0.4。模式出现次数之差为0、1、2、3、4、5时对应的成对比较 $a_{ij} = 1, 2, 3, 4, 5, 6$; 设 $7e$ 为输入向量与13个目标向量之差的期望、方差的最大值与最小值之差, 差的可能取值为0、 e 、 $2e$ 、 $3e$ 、 $4e$ 、 $5e$ 、 $6e$ 、 $7e$, 对应的成对比较 $b_{ij}, c_{ij} = 1, 2, 3, 4, 5, 6, 7, 8$ 。层次分析集成方法模式识别正确率如图4所示。

实验结果表明:(1) 集成系统的识别率高于各子系统, 且识别率增加量在低噪声情况下高于高噪声情况。(2) RBF子网络和SOM子网络与其他子网络相比在低噪声时识别率较高, 而其他情况下识别率低得多。(3) 初始化参数不同的LVQ网络与PNN网络的识别率相当。(4) 层次分析集成方法的识别率比相对多数集成法高, 在中低噪声条件下提高效果突出, 在高噪声情况下识别率相当。

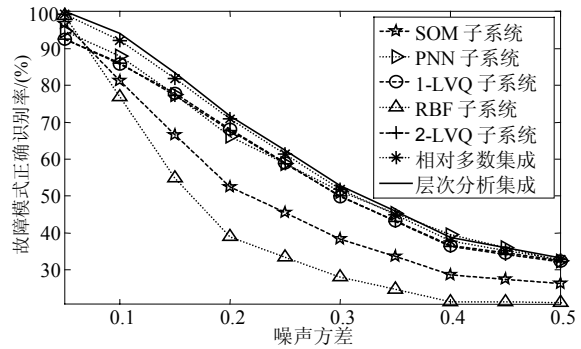


图4 分类正确率对比图

5 结束语

本文建立了神经网络的层次分析集成方法模型使用某型设备的典型故障数据进行测试试验。子系统采用了SOM、PNN、1-LVQ、2-LVQ、RBF四种类型的五个神经网络, 对每个子网络进行独立训练并行处理, 使用层次分析法对各神经网络子系统的识别模式进行集成提高了模式识别率, 集成方法计算量很小, 计算结果符合实际情况; 表明层次分析集成方法是一种很有效的集成方法, 除神经网络本方法还可集成其他模式识别方法。

本文研究工作得到电子科技大学青年基金(JX7019)的资助, 在此表示感谢。

参考文献

- [1] 周志华, 陈世福. 神经网络集成[J]. 计算机学报, 2002, 25(1):
- [2] SYLVESTER J, NITESH V C. Evolutionary ensemble creation and thinning[J]. Neural Networks, 2006, 7 16-21: 5148-5155.
- [3] SHARKEY A J C. On combining artificial neural nets[C]// Connection Science, 1996, 8(3-4): 299-314.
- [4] KOHONEN T. The self-organizing map[C]//In: Proceedings of IEEE, 1990,78(9): 1464-1481.
- [5] SPECHT D F. Probabilistic neural networks[J]. Neural Networks, 1990, 1(3): 109-118.
- [6] 李国友, 姚 磊, 李惠光, 等. 基于优化的RBF神经网络模式识别新方法[J].系统仿真学报, 2006, 18(1): 181-184.
- [7] 曹安照, 田 丽. 基于RBF神经网络的短期电力负荷预测[J]. 电子科技大学学报, 2006, 35(4): 507-509.
- [8] 楼顺天. 基于Matlab的系统分析与设计—神经网络[M]. 西安: 西安电子科技大学出版社, 1999.
- [9] SAATY T L. The analytic hierarchy process: Planning, priority setting, resource allocation[M]. [S.l.]: McGraw Hill, 1980.
- [10] 陈中柘. 神经网络定性分析[J]. 电子科技大学学报, 2002, 31(3): 250-254.