

文章编号: 1001-0920(2007)10-1177-04

一种改进的 RAN 网络结构优化算法

韩 敏, 穆云峰

(大连理工大学 电子信息与工程学院, 辽宁 大连 116023)

摘 要: 针对常用的资源分配网络(RAN)存在的问题,提出一种改进的 RAN 网络学习算法(PIRAN).该算法可避免常用的 RAN 网络由于没有初始隐层节点而造成的不足,在一定程度上改善新性条件对网络结构和参数的影响,并采用 FPE 剪枝策略使网络复杂度与拟合度得到较好的结合.最后通过对无机建筑材料成分分析的仿真,表明该算法可有效简化网络结构,实现样本的正确分类,获得较好的校验性能.

关键词: 资源分配网络; 相似度; FPE 剪枝策略; 材料成分

中图分类号: TP183

文献标识码: A

Improved learning algorithm for optimizing RAN network structure

HAN Min, MU Yun-feng

(School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116023, China.

Correspondent: HAN Min, E-mail: minhan@dlut.edu.cn)

Abstract: An improved resource-allocating network (RAN) learning algorithm is proposed for avoiding the existing problems of RAN network. The algorithm avoids the drawback that the RAN has no initial hidden node, improves the influence of the RAN novelty on the architecture and parameters, and also achieves a better combination between the complexity and precision applying the FPE pruning strategy. Finally, simulations by analysis of the composition of building materials show that the algorithm can simplify the network architecture, realize correct classification of patterns, and improve better testing performance.

Key words: RAN; Similarity; FPE pruning strategy; Composition of materials

1 引 言

近年来,径向基函数神经网络(RBFNN)以其拓扑结构简单和易于学习等优点,广泛应用于模式识别、函数逼近等领域^[1].构造 RBFNN 的关键是网络隐层单元数的确定,过多或过少都会对网络造成很大的影响^[2].但目前仍没有一种有效的方法来确定适当的隐层节点数^[3].常用的学习方法是在学习过程中根据某种准则动态地添加或删除隐节点,以达到网络结构适当的要求.其中最著名的是 Platt 提出的资源分配网络(RAN)学习算法.

RAN 网络学习算法可避免固定网络结构造成的欠拟合或过拟合,同时利用样本输出信息提高学习精度,从而避免聚类方法产生过于庞大的网络结构^[4].但该学习算法仍存在不足之处,例如常用的 RAN 网络由新性条件决定网络结构,而新性条件受输入样本的影响较大^[4];又如常用的 RAN 网络没

有设定任何初始节点,不仅增加了网络学习时间和复杂度,而且降低了校验效果^[5].

本文针对 RAN 网络的上述缺点,提出一种改进的 RAN 网络学习算法.该算法在均值聚类的基础上,学习过程中选取产生最大误差的样本,当满足新性条件时分配新的隐层节点,当不满足时则使用相似度^[6]参数对隐层中心进行调整.为降低网络结构的复杂度,根据 FPE 剪枝策略^[7]对网络性能影响不大的情况对网络进行剪枝.最后将这种改进方法应用于建筑材料系统的材料成分分析,以验证改进学习算法的分类精度和校验性能.

2 改进的 RAN 学习算法(PIRAN)

2.1 原始 RAN 结构算法简介

原始 RAN 网络学习算法是 Platt 于 1991 年提出的,其主要学习行为是根据新性条件来决定分配新的隐层节点或调整现有隐层节点的参数.开始时

收稿日期: 2006-06-06; 修回日期: 2006-10-30.

基金项目: 国家自然科学基金项目(60674073, 60374064).

作者简介: 韩敏(1959—),女,吉林延吉人,教授,博士生导师,从事神经网络、专家系统等研究;穆云峰(1981—),男,辽宁鞍山人,硕士生,从事神经网络、模式识别的研究.

网络没有任何的隐层节点,依据检验顺序输入各样本输入输出对,当其满足新性条件时,分配新的隐层节点;否则,使用梯度法调整隐层参数.本文针对原始 RAN 网络的一些缺点,提出一种 PIRAN 学习算法.

2.2 PIRAN 学习算法简介

2.2.1 初始中心的选取

假设学习样本的类别为 M ,初始隐层节点数与学习样本的类别相同,则初始隐层中心 c_j 及其宽度 d_j ($j = 1, 2, \dots, M$) 分别由下式求得:

$$c_j = \frac{1}{N_j} \sum_{x_i \in A_j} x_i, \quad (1)$$

$$d_j = d_{j \max}, \quad j = 1, 2, \dots, M. \quad (2)$$

式(1)中: A_j 是各样本归属的样本集, x_i 是各类样本集中的样本, N_j 是各类中样本的数量;式(2)中: $d_{j \max}$ 是一常数, $d_{j \max}$ 是各类样本集中学习样本到中心的最远距离.从上面两式可以看出,这种方法可保证每类中至少有一个聚类中心,并且充分利用了学习样本的输入输出信息.

2.2.2 隐层中心的调整

隐层中心的调整根据新性条件包括两部分:分配网络新的隐层节点和调整网络已有的网络参数.在学习过程中,选取当前产生最大误差的学习样本代替输入网络的顺序样本,则新性条件可重新描述如下:

$$r_{\max} = \|x_{\maxerror} - c_{\text{nearest}}\| > e_{\max}, \quad (3)$$

$$|e_{\max}| = |y_{\maxerror} - f(x_{\maxerror})| > e_{\min}. \quad (4)$$

其中: x_{\maxerror} 是产生最大误差的学习样本, y_{\maxerror} 和 $f(x_{\maxerror})$ 分别代表 x_{\maxerror} 所对应的期望输出和实际输出, c_{nearest} 是距离 x_{\maxerror} 最近的隐层中心,和 e_{\min} 都是人为设定的阈值.式(3)和(4)分别对应新性条件中的距离准则和误差准则.

将 x_{\maxerror} 代入新性条件,如果同时满足式(3)和(4),即该样本与各隐层中心均不相似,则根据下式添加隐层节点:

$$\begin{cases} M = M + 1, N_M = 1, \\ c_M = x_{\maxerror}, d_M = \|x_{\maxerror} - c_{\text{nearest}}\|. \end{cases} \quad (5)$$

其中 d_M 为一常数.当 x_{\maxerror} 不满足式(3)或(4)时,采用下式对隐层中心进行调整:

$$N_j = N_j + 1, \quad j = 1, 2, \dots, M; \quad (6)$$

$$c_{ji} = 2 \sum_j \frac{(x_{\maxerror} - c_{ji})}{d_j} \phi_j(x_{\maxerror}) w_{sj} \times (f(x_{\maxerror})_s - y_{\maxerror}_s), \quad s = 1, 2, \dots, K; \quad (7)$$

$$c_{ji} = c_{ji} + c_{ji}, \quad i = 1, 2, \dots, n. \quad (8)$$

其中: n , K 和 M 分别为输入节点个数、输出节点个

数和隐层节点个数, N_j 为各类样本个数, d_j 是步长, ϕ_j 是表征 x_{\maxerror} 与 c_j 相似程度的参数, ϕ_j 定义如下:

$$\phi_j = \frac{\|x_{\maxerror} - c_j\| - \|x_{\maxerror} - c_{\text{nearest}}\|}{\|x_{\maxerror} - c_{\text{farthest}}\| - \|x_{\maxerror} - c_{\text{nearest}}\|}. \quad (9)$$

其中: c_{nearest} 是距离 x_{\maxerror} 最近的中心, c_{farthest} 是距离 x_{\maxerror} 最远的中心.

选取产生的最大误差样本,其优点是新性较为突出,并以此为基础改变新性条件,可有效改善网络的隐层结构.距离准则和误差准则缺一不可,如果只满足误差准则,则可能对异常样本或受干扰的样本分配新的隐层节点,而距离准则在一定程度上弥补了这一缺点.此外,为了得到更小的网络结构,删除由于样本异常或干扰所添加的隐节点.本文使用剪枝策略对网络结构进行简化.

2.2.3 剪枝策略

剪枝策略的基本思想是先取一个规模较大的网络,然后将那些不重要的权和神经元逐渐删掉,最终达到缩小网络规模的目的.

本文所用的剪枝策略是 Akaike 的最终误差标准,它由下式给出:

$$FPE = \frac{1 + (n_p/N)}{1 - (n_p/N)} R_e. \quad (10)$$

其中: R_e 为网络经验风险函数(误差函数), n_p 为网络的隐层节点数, N 为学习样本数, λ 为一权值因子.

由式(10)可以看出,FPE是一种经验风险函数的加权形式,对于结构简单而精度不高的网络或精度虽高但结构过于复杂的网络,它们的 FPE 值都很高.这种剪枝策略的目的是在精度与结构复杂度之间寻求最佳的结合点,因此该标准兼顾了模型拟合度与模型复杂度的均衡问题.

FPE 剪枝策略的具体作法是:在一个由某种学习算法得到的既定中心集上,依次删除每个中心,计算在缺少该中心情况下的 FPE 值,当 FPE 达到最小时,剪枝策略即可结束.

2.2.4 算法流程

综上所述,PIRAN 学习算法流程可归结如下:

Step1: 根据式(1)和(2)选取隐层初始中心和宽度,对网络进行初始化.

Step2: 计算网络隐层输出.

Step3: 根据最小二乘法求出网络输出权值 w 及网络的学习输出.

Step4: 根据均方差误差计算网络的学习误差.若误差小于期望值或学习步骤达到最大,则停止隐层节点的添加,转向 Step6;否则,转向 Step5.

Step5: 根据新性条件,判断是否满足添加节点条件.若满足,则根据式(5)添加节点;若不满足,则根据式(6)~(9)调节隐层节点,转向 Step2.

Step6: 根据式(10)对冗余节点进行剪枝,直到 FPE 值达到最小,结束学习.

3 仿真实例

本文方法可用于一类模式分类问题,现以建筑材料 CaO-Al₂O₃-SiO₂ 三元系统为例.加以仿真说明.

3.1 CaO-Al₂O₃-SiO₂ 系统及其特征提取

在 CaO-Al₂O₃-SiO₂ 三元系统中,CaO,Al₂O₃和 SiO₂ 以不同比例结合,在不同温度下可能生成 15 种化合物. CaO-Al₂O₃-SiO₂ 系统是一个复杂的非线性系统,特别适合于用神经网络对其进行处理.因此,本节使用 PIRAN 学习算法对 CaO-Al₂O₃-SiO₂ 三元系统相图进行数据的学习和校验.

使用神经网络建立这一系统模型的重点是输入变量和输出变量的确定.根据 CaO-Al₂O₃-SiO₂ 三元系统中各个化合物的分布,采用如下表示方式:输入矢量 $x = [x_1, x_2, x_3, x_4]$, x_1, x_2, x_3, x_4 分别代表 CaO, Al₂O₃, SiO₂ 的百分含量及系统温度;输出矢量 $y = [y_1, y_2, y_3]$, y_1, y_2, y_3 分别表示形成物质的 C, A, S 的组成比例.当有一种物质生成时,使用该物质的 C, A, S 组成比例表示;当有多个物质生成时,则用各个物质中 C, A, S 的和表示.这样,可通过 C, A, S 的组成比例唯一表示所形成的物质.

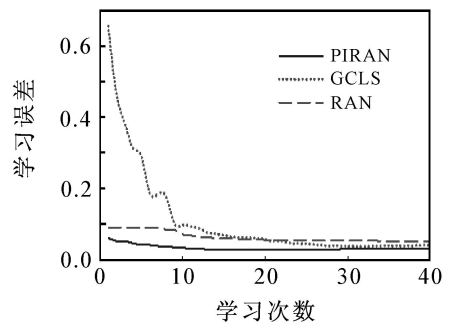
3.2 仿真结果

从 CaO-Al₂O₃-SiO₂ 系统相图中选取 157 组数据,其中 130 组用于学习,27 组用于校验.建模之前,先将所有样本输入都归一化到 [0, 1] 范围内,评价函数采用均方差函数来检验网络的输出误差.为说明改进算法的有效性,本文采用 3 种学习算法对该系统进行仿真:GCLS 算法^[8]、RAN 算法和本文方法(PIRAN).为保证算法的一致性,输出权值均采用最小二乘法求得.

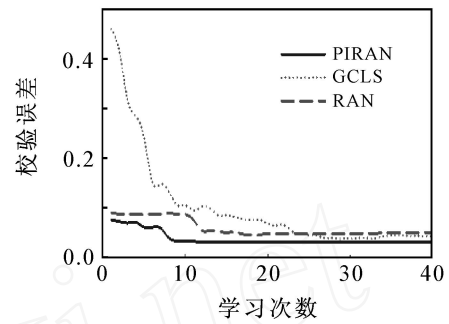
仿真结果如表 1 所示,3 种算法的仿真曲线如图 1 所示.图 1(a)~(d)分别表示 3 种算法所得网络的学习误差、校验误差、学习精度和校验精度.

表 1 3 种算法的仿真结果

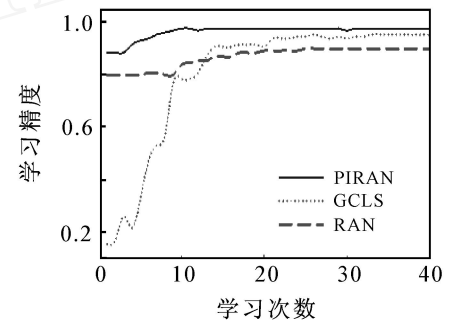
| | GCLS | RAN | PIRAN |
|---------|-----------|-----------|-----------|
| 中心数 | 35 | 28 | 18 |
| 学习误差 | 0.051 246 | 0.039 051 | 0.030 459 |
| 学习精度/ % | 90.0 | 95.4 | 97.7 |
| 校验误差 | 0.050 050 | 0.042 882 | 0.031 615 |
| 校验精度/ % | 88.9 | 88.9 | 96.3 |



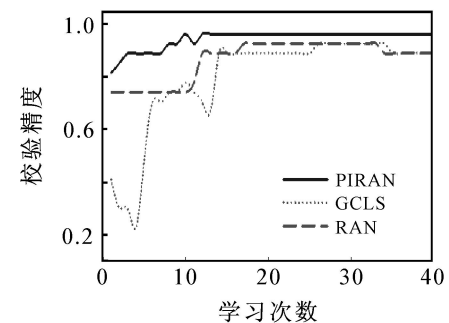
(a) 学习误差曲线



(b) 校验误差曲线



(c) 学习精度曲线



(d) 校验精度曲线

图 1 3 种算法仿真曲线比较

3.3 仿真结果分析

从以上仿真结果可以看出,在该系统模型仿真中,本文的改进算法在明显简化网络结构的同时,可得到比其他两种算法更高的精度.

GCLS 算法是一种根据距离量度进行节点添加的算法,它只用到了样本的输入信息;RAN 网络学习算法是根据新性条件,即根据距离量度和误差量

度进行隐层的调整,同时用到了样本的输入输出信息,可以有效地提高学习效率;PIRAN 网络学习算法在 RAN 网络学习算法的基础上,针对自身不足的几个特点进行改进,使其在网络复杂度与拟合度上得到较好的结合.

PIRAN 学习算法与原始 RAN 网络学习算法的不同之处(除剪枝策略外)表现在以下几方面:网络初始化,采用最大误差样本代替顺序样本输入网络,采用相似度参数调整隐层中心.为分析这些不同之处对网络的影响,分别对其进行仿真,并与原始 RAN 网络的学习结果进行比较,结果如表 2 所示.

表 2 各不同点对 RAN 网络的影响

| | RAN | 仅最大误差样本 | 仅初始化 | 仅相似度 |
|---------|-----------|-----------|-----------|-----------|
| 中心数 | 28 | 17 | 22 | 31 |
| 学习误差 | 0.039 051 | 0.032 856 | 0.042 443 | 0.035 468 |
| 学习精度/ % | 95.4 | 96.2 | 94.7 | 96.2 |
| 校验误差 | 0.042 882 | 0.052 079 | 0.030 662 | 0.035 688 |
| 校验精度/ % | 88.9 | 85.2 | 92.6 | 92.6 |

由表 2 看出,使用产生最大误差的样本代替顺序样本输入网络,对网络结构影响较大,这是因为新性条件对输入样本较为敏感;对网络进行初始化,可提高网络的学习效率,在一定程度上减小网络结构;相似度参数的使用,有利于隐层中心的改变,可提高网络的学习能力和校验能力.

4 结 语

本文提出一种改进的 RAN 网络学习算法(PIRAN).根据样本的输入输出信息对网络进行初始化,在学习过程中采用最大误差的样本代替顺序样本输入网络,在不满足新性条件时,采用相似度参数对网络的隐层中心进行调整,并根据 FPE 剪枝策略对网络结构进行简化.通过对 $\text{CaO-Al}_2\text{O}_3\text{-SiO}_2$ 系统进行仿真,结果表明:该学习算法可避免常用的 RAN 网络无初始隐层节点的不足,改善输入样本对网络的影响,有效地调整了隐层中心,简化了网络结构.该方法可以实现样本的正确分类,在简化网络结构的同时,得到更好的网络性能,使网络复杂度与拟合度得到较好的结合.

参考文献(References)

- [1] 刘铁男,段玉波,刘志德,等.带优选聚类算法的 RBF 网络辨识器及应用[J].控制与决策,2003,18(2):232-236.
(Liu Tie-nan, Duan Yu-bo, Liu Zhi-de, et al. RBF network identifier with optimal selection cluster algorithm and its application[J]. Control and Decision, 2003, 18(2): 232-236.)
- [2] 刘志远,吕剑虹,陈来九.基于 RBF 神经网络的单元机组负荷系统建模研究[J].控制与决策,2003,18(5):637-640.
(Liu Zhi-yuan, Lv Jian-hong, Chen Lai-jiu. Modelling of unit power plant load system based on RBF neural network[J]. Control and Decision, 2003, 18(5): 637-640.)
- [3] Parekh R, Yang J, Honavar V. Constructive neural-network learning algorithms for pattern classification [J]. IEEE Trans on Neural Networks, 2000, 11(2): 436-451.
- [4] 魏海坤,丁维明,宋文忠,等. RBF 网的动态设计方法[J].控制理论与应用,2002,19(5):673-680.
(Wei Hai-kun, Ding Wei-ming, Song Wen-zhong, et al. Dynamic method for designing RBF neural network[J]. Control Theory and Applications, 2002, 19(5): 673-680.)
- [5] Manolis W, Nicolas T, Stefanos K. Intelligent initialization of resource allocating RBF networks [J]. Neural Networks, 2005, 18(2): 117-122.
- [6] Kyoung M L, Street W N. An adaptive resource allocating network for automated detection, segmentation and classification of breast cancer nuclei topic area: Image processing and recognition [J]. IEEE Trans on Neural Networks, 2003, 14(3): 680-687.
- [7] Gomm J B, Yu D L. Selecting radial basis function network centers with recursive orthogonal least squares training [J]. IEEE Trans on Neural Networks, 2000, 11(2): 306-314.
- [8] Theodoridis S, Koutroumbas K. Pattern recognition [M]. Beijing: China Machine Press, 2003: 552-554.