

文章编号: 1007-6492(1999)01-0036-03

故障诊断神经网络隐层结点数的优化方法

薛斌党¹, 孙惠国¹, 薛文芳², 时 斌¹

(1. 郑州工业大学机械与电子工程学院, 河南 郑州 450002; 2. 河南农业大学机电工程学院, 河南 郑州 450002)

摘要: 针对故障诊断神经网络隐层结点数过少或过多而造成的网络学习过程可能不收敛、网络泛化能力弱和网络容错性下降等问题, 根据隐层结点对输出层输出影响的显著性差异, 提出了一种在权值学习过程中动态的删除隐层冗余结点的优化学习方法. 该方法在故障诊断神经网络实际应用中提高了收敛速度和诊断精度.

关键词: 神经网络; 隐层结点数; 信息熵; 故障诊断; 优化方法

中图分类号: TH 806.3 **文献标识码:** A

0 引言

在多层前馈神经网络应用中, 结构优化和联接权值学习是两个重要的研究课题. 故障诊断神经网络的结构优化主要是隐层结点数的优化. 隐层结点数过少, 学习过程可能不收敛, 网络对训练模式的特征提取、识别能力降低, 训练样本外的数据其精度急剧下降, 网络的泛化能力弱, 故障诊断精度不高; 隐层结点数过多, 网络长时间不收敛, 接近收敛时易出现振荡, 而且由于过拟合, 造成网络的容错性能下降. 然而, 隐层结点数的确定目前尚无严格的理论依据. 对此, 人们已提出了许多效果较好的算法, 其中以 Bhat 等人提出的网络剪枝法^[1]是比较有代表性的一类算法. 本文针对故障诊断网络的联接权值学习和结构优化问题, 提出了在网络权值学习过程中根据隐层结点对网络输出值影响的显著性不同, 来调整网络隐层结点数的优化方法, 其中联接权值的学习采用信息熵为代价函数的神经网络学习算法^[2]. 优化前后网络的对比结果表明: 优化后的网络不仅收敛速度快, 而且诊断精度高.

1 动态删除冗余结点法的优化思想^{3~9]}

假设神经网络输入层、隐层、输出层结点数分别为 N, L, M , 学习样本数为 P , 采用因素平方总和作为评判隐层结点对输出层输出影响大小的优

化指标函数.

对 P 个学习样本, 第 j 个隐层结点对输出层所有结点输出影响的因素平方总和为

$$SS_j = \sum_{p=1}^P SS_j(p), \quad (1)$$

$$SS_j^{(p)} = (C_j^{(p)} - \bar{t})^2. \quad (2)$$

$SS_j^{(p)}$ 表示第 j 个隐层结点对第 p 个学习模式的因素平方和. 式(2)中, \bar{t} 表示所有期望输出的均值; $C_j^{(p)}$ 表示第 j 个隐层结点与所有输出层结点联接后, 输出层所有结点的实际输出值的均值.

$$C_j^{(p)} = \frac{1}{M} \sum_{s=1}^M f(W_{js} b_j^{(p)} - \theta_s)$$

($f(\cdot)$ 为非线性激励函数). (3)

式(3)中, W_{js} 表示第 j 个隐层结点与第 s 个输出层结点的联接权值; $b_j^{(p)}$ 为第 p 个学习模式第 j 个隐层结点的输出; θ_s 为第 s 个输出层结点的阈值. 对第 j 个隐层结点, 以其对输出层所有结点输出影响的因素平方总和作为优化指标函数. 预先设定精度 ϵ 的值, 如果

$$SS_j < \epsilon, \quad (4)$$

则删除第 j 个隐层结点, ϵ 的值可根据具体情况, 由经验事先确定^[3].

隐层结点数的初始值通常根据经验选定. 对于 $N-L-M$ 的三层神经网络, P 个学习模式, 隐层结点数的初始值经验公式^[3]为

收稿日期: 1998-09-21; 修订日期: 1998-11-06

基金项目: 河南省科技攻关项目(98112037)

作者简介: 薛斌党(1971-), 男, 河南省新野县人, 郑州工业大学硕士研究生.

$$L > \frac{M(P-1)}{(N+M-1-P)} = A \quad (5)$$

一般而言, 隐层结点数的取值范围为 $A < L < 3A$

下面给出故障诊断神经网络联接权值学习和隐层结点数优化的基本过程:

(1) 根据诊断网络输入层、输出层的结点和训练样本数, 首先由经验公式(5)确定 A 的近似值, 然后在 $A < L < 3A$ 范围内选定一个较大的隐层结点数的初始值. 网络各连接权值赋予 $(0, 1)$ 间的随机数, 变尺度矩阵赋予单位矩阵.

(2) 随机选取一学习模式对, 提供给网络, 采用经信息熵为代价函数的神经网络学习算法^[3]对网络联接权值进行学习.

(3) 在网络联接权值学习过程中, 以因素平方总和作为评判指标函数来评判每一个隐层结点对输出层输出值的影响程度, 如果某隐层结点的因素平方总和小于给定的精度值, 则删除该结点.

(4) 在网络连接权值学习和结构优化过程中, 当隐层所有结点对输出值影响的因素平方和均大于给定的限定值时, 在以后的学习过程中, 可跳过第(3)步, 网络结构保持不变, 只对连接权值进行学习.

(5) 随机选取下一个模式对送给网络, 直至全部 P 个学习模式对训练完毕.

(6) 重新从 P 个学习模式对中随机选取一个模式对, 返回步骤(2), 直至, 网络全局信息熵函数值小于预先设定的根限值.

2 算例

旋转机械常见故障的标准模式见表 1. 首先根据经验公式确定隐层结点数的取值范围为 $9 \sim 27$. 取初始值为 25, 故障诊断网络初始拓扑结构为 $8-25-9$ 型, 非线性激励函数为 Sigmoid 函数. 采用信息熵为代价函数的神经网络学习算法和本文的隐层结点数优化方法对表 1 数据进行学习. 用优化学习后的故障诊断网络再对表 1 数据进行学习, 从表 2 结果可以看出, 优化后的网络学习速度明显高于根据经验公式确定的网络速度.

利用学习、优化前的诊断网络和学习、优化后的诊断网络对表 3 的 3 种故障样本进行诊断, 从表 4、表 5 两个诊断结果可以看出, 原始拓扑结构的网络诊断精度比学习优化后的网络诊断精度低, 模糊度相对大些.

表 1 标准训练样本

故障模式	$(0.00 \sim 0.39)f$	$(0.40 \sim 0.49)f$	$(0.50 \sim 0.99)f$	$1f$	$2f$	$3f$	$4f$	$\geq 5f$	输出向量
油膜振荡	0.1	0.5	0.1	0.3	0.0	0.0	0.0	0.0	100000000
旋转失速	0.2	0.1	0.2	0.3	0.2	0.1	0.1	0.0	010000000
喘振	0.7	0.0	0.0	0.3	0.0	0.0	0.0	0.0	001000000
碰摩	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	000100000
支承松动	0.1	0.2	0.1	0.4	0.1	0.1	0.0	0.0	000010000
轴裂纹	0.0	0.0	0.0	0.4	0.2	0.0	0.0	0.2	000001000
不平衡	0.0	0.1	0.0	0.5	0.1	0.0	0.0	0.0	000000100
油膜涡动	0.1	0.4	0.1	0.3	0.1	0.0	0.0	0.0	000000010
不对中	0.0	0.0	0.0	0.4	0.5	0.1	0.0	0.0	000000001

说明: f 为工频, Hz

表 2 优化前后网络学习结果比较

网络结构	原始拓扑结构(8-25-9)	优化后的网络拓扑结构(8-20-9)
精度	10^{-6}	10^{-6}
迭代次数	618	44
训练时间	1'06"54	16" 53

表 3 3 种故障模式

故障模式	$(0.0 \sim 0.39)f$	$(0.40 \sim 0.49)f$	$(0.50 \sim 0.99)f$	$1f$	$2f$	$3f$	$4f$	$\geq 5f$
1	0.6	0.05	0.05	0.3	0.04	0.04	0.0	0.0
2	0.0	0.0	0.0	0.3	0.6	0.1	0.0	0.0
3	0.15	0.2	0.18	0.65	0.2	0.2	0.0	0.0

说明: f 为工频, Hz

表 4 原始拓扑结构网络诊断结果

故障模式	油膜振荡	旋转失速	喘 振	碰 摩	支承松动	轴裂纹	不平衡	油膜涡动	不对中
1	0.00000	0.209931	0.89616	0.00172	0.07036	0.00001	0.14609	0.01598	0.00007
2	0.00001	0.070496	0.00005	0.00062	0.00016	0.00335	0.00003	0.00001	0.87033
3	0.00384	0.002904	0.04663	0.00078	0.72625	0.00934	0.00241	0.498336	0.07942

表 5 优化后网络诊断结果

故障模式	油膜振荡	旋转失速	喘 振	碰 摩	支承松动	轴裂纹	不平衡	油膜涡动	不对中
1	0.00000	0.021469	0.97767	0.00046	0.03344	0.00002	0.00001	0.01439	0.00006
2	0.00001	0.09543	0.00001	0.00026	0.00028	0.00088	0.00005	0.00002	0.98343
3	0.00027	0.01138	0.01268	0.00008	0.99062	0.00748	0.47692	0.00361	0.00952

3 结束语

动态删除隐层冗余结点优化方法是一种删除法,从上面故障诊断实例可以看出该方法在故障诊断领域是可行的,在其他领域的应用还有待验证.网络的结构学习与参数学习相比难度更大,如何动态的增加、删除隐层结点,较快地找到最优网络结构,还有待进一步的研究.

参考文献

[1] BHAT N V, MC AVOY T J .Determining model structure for neural models by network stripping[J] .Computers and Chem Engng, 1992, 16(4) : 271—281.
 [2] 孙惠国.信息熵为代价函数的神经网络算法及诊断

应用[J] .振动工程学报, 1998, 11(3) : 135—138.

[3] GAO Da -qi, WU Shou -yi .An optimization method for the topological structure of feed -forward multi -layer neural networks[J] .Pattern Recognition, 1998, 31(9) : 1337—1342.
 [4] DOERING Axel, GALICKI Mirosław, WITTE Herbert . Structure optimization of neural networks with the A * -algorithm[J] .IEEE Transaction on Neural Networks, 1997, 8(6) : 1434—1445.
 [5] 付 英,曾 敏,李兴源. 隐含层人工神经网络电压安全评估的影响[J] .电力系统自动化, 1996, 20(11) : 13—16.
 [6] 王志鹏,孙惠国,时 斌.故障诊断的多层神经网络研究[J] .郑州工业大学学报, 1998, 19(2) : 20—23.

Optimization Algorithm for Hidden Layer Nodes of Fault Diagnosis Neural Network

XUE Bin -dang¹, SUN Hui -guo¹, XUE Wen -fang², SHI Bin¹

(1.College of Mechanical & Electronic Engineering, Zhengzhou University of Technology, Zhengzhou 450002, China; 2.College of Mechanical & Electronic Engineering, Henan Agriculture University, Zhengzhou 450002, China)

Abstract :The problem that the hidden layer nodes of neural network to fault diagnosis are insufficient or redundant can result in many problems such as neural network study unconvergence, weak ability of generalization and tolerance of error. In order to solve these problems on the basis of the differences of the hidden nodes' great effect on the outputs of the output layer, an optimization algorithm for dynamically deleting redundant hidden nodes during the weight learning process is presented. The results of its application in neural network to fault diagnosis show that the optimization algorithm can increase the training speed and promote accuracy of diagnosis.

Key words :neural network; numbers of hidden nodes; entropy; fault diagnosis; optimization algorithm