

优化 BP 神经网络的可靠性预测模型

王蓓, 刘桥

(贵州大学 计算机科学学院, 贵州 贵阳 550025)

摘要: 为了提供一种更加准确高效的算法, 对传统的 BP 神经网络模型进行优化。首先对初始权值的选取规则, 传统的 N-W 规则算法进行改进, 再对其它参数进行优化。在此基础上, 将神经网络理论应用于系统可靠预测评价之中, 提出了基于此理论的系统可靠预测评价模型、实现方法和优点; 评价实例证明此方法的可行性。

关键词: BP 神经网络; 网络优化; 可靠预测评价

中图分类号: TP183

文献标识码: A

文章编号: 1673-629X(2007)09-0102-04

Credible Prediction Model Based on BP Neural Network

WANG Bei, LIU Qiao

(College of Info. and Computer Sci., Guizhou Univ., Guiyang 550025, China)

Abstract: The aim is to provide a kind of rule which is more accurate and more effective. This optimization builds based on the traditional back propagation neural network model. Firstly, optimize the traditional N-W rule for choosing the rudimentary power-number. Secondly, optimize other parameters. Advance the optimization tactics with using this model to forecast the system's dependable. Giving the forecast model depend on ANN' BP theory, realise method and advantage. This way is feasible by the appraising example.

Key words: back propagation neural network; network optimization; credible prediction evaluation

0 引言

可靠性预测是在设计阶段进行的定量估计未来产品可靠性的一种方法。在复杂系统的预测方法的研究中, 现在研究较多的是传统统计方法和人工神经网络方法, 传统的统计学方法存在不适合动态系统、建模复杂等局限性, 相比之下, 作为新生事物的人工神经网络, 被普遍认为具有大规模非线性并行处理能力, 依据数据自身的内在建模、良好的适应性及自学习能力, 较强的抗干扰能力等优良性能。因此, 与传统统计学方法比较, 人工神经网络用于复杂系统的预测就似乎具有了明显优势。

1 人工神经网络研究的主要方法

人工神经网络^[1]受到学术界的高度重视和广泛应用, 已经成功地应用于众多领域, 如模式识别、图像处理、语音识别、智能控制、虚拟现实、优化计算、人工智能等领域。按照网络的拓扑结构和运行方式, 神经网络模型分为前馈多层式网络模型、反馈递归式网络模

型、随机型网络模型等。目前在模式识别中应用成熟较多的模型是前馈多层式网络中的 BP 反向传播模型, 基于 BP 算法的多层前馈型网络^[1]的结构如图 1 所示。这种网络不仅有输入层节点、输出层节点, 而且有一层或多层隐含节点。对于输入信息, 要向前传播到隐含层的节点上, 经过各单元的特性为 Sigmoid 型的激活函数运算后, 把隐含节点的输出信息输出到输出节点。最后输出结果。

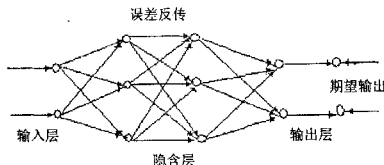


图1 基于 BP 算法的神经网络结构

具体的 BP 网络的算法描述如下^[2]:

- 1) 权值和阈值初始化;
- 2) 给定输入 X 和目标输出 Y ;
- 3) 计算实际输出 $y_j = f(\sum_i w_{ij}x_i)$, $j = 1, 2, \dots, n_j$;
- 4) 修正权值: 从输出层开始, 将误差信号沿连接通路返回传输方向传播, 通过修正各权值使误差向减

收稿日期: 2006-12-05

作者简介: 王蓓(1977-), 女, 贵州贵阳人, 硕士研究生, 研究方向为电路与系统; 刘桥, 博士生导师, 研究方向为 EDA。

小的方向变化;

5) 达到精度要求或循环次数要求,则输出结果,否则返回 2)。

BP 算法是一种梯度反向传播算法

$$\begin{aligned} W(l+i) &= W(l) + \eta \left(-\frac{\partial E}{\partial W} \right) |_{w=w_{l,i}} \\ \Delta W_{ij}(n+1) &= \eta \times \Phi_i \times O_j + a \times \Delta W_{ij}(n) \end{aligned} \quad (1)$$

其中, η 为学习因子; Φ_i 为输出节点 i 的计算误差; O_j 为输出节点 j 的计算输出; a 为动量因子。

2 BP 网络模型的缺陷分析及优化策略

1) 最主要的优点有^[3]:

(1) BP 网络采用的传递函数通常是 Sigmoid 型可微函数,所以可以实现输入和输出间的任意非线性映射。

(2) BP 网络的隐层节点足够多时,可以任意精度逼近任何一个具有有限间断点的非线性函数。但 BP 网络不可避免的存在缺点,主要体现在^[3]:

* 从数学上看,它可归结为一非线性尺度优化问题,因此不可避免地存在局部最小问题。

* 学习算法的收敛速度很慢,通常需要很长时间的训练。

* 网络结构为前向结构,没有反馈连接,因此它是一非线性映射系统。

* 网络隐层节点个数的选取,尚无理论的指导。

* 对新加入的样本要影响已经学完的样本,刻画每个样本特征的输入样本的特征数目要求必须相同。

2) 优化策略:

BP 网络的参数有很多,但对网络性能和识别结果有明显影响的主要有:初始权值、隐层节点数、网络训练步数、网络学习率和动量参数。下面重点讨论初始权值的优化问题。

2.1 算法优化

1) 初始权值对训练的影响。

普通的 BP 网络都是采用 S 型函数, S 型函数是一个奇函数,在原点附近近似为线性函数;当自变量大于 1 时,函数进入饱和区后变化非常缓慢,此时权值的变化不会对神经元的输出产生足够影响,导致网络不能收敛。

一般地,初始权值取较小的随机数,以保证每个神经元一开始都工作在线性区。如果权初值没有与网络结构联系起来,仅是任意选取小的随机数,仍然不能保证整个网络处于最佳工作状态,因此研究网络初始权值的选取,以及初始权值与网络各层节点的对应关系,

对于提高网络的学习速度有重要的意义。

2) 初始权值的 Nguyen - Widrow 规则^[4]。

关于网络初始权值选取的研究很少,文献报道的有遗传算法来优化神经网络的初始权值,采用厄米特样插值函数的方法,以及采用记忆初始权值进行优化等。但这些方法都过于繁琐,计算量和数学推导都比较复杂。而威得罗等人提出了 N-W 规则,使输入数据在运算中避开 S 型函数的饱和区。

在输入层到隐含层的映射中,如果隐含层的节点数少于输入的节点数,那么降维过程可能会丢失信息;反之,当隐含层的节点数大于输入的节点数时,升维过程会带来运算量的显著增加。因此,在 N-W 规则中,选取的隐含层节点数与输入层的节点数相等。S 函数的表达式为

$$\text{sigmoid}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

其中 $x \in [-1, 1]$ 的区域可以近似看作线性区,对应的直线用 $y = kx$ 来表示,如图 2 所示,图中长虚线表示 S 型函数与直线的偏差。用最小均方误差准则可以确定与 S 函数最接近的直线的斜率 K ,即满足:

$$\int_{-1}^1 \left[\frac{e^x - e^{-x}}{e^x + e^{-x}} - kx \right]^2 dx \rightarrow \min \quad (3)$$

计算得到 $k \approx 0.7$ 。

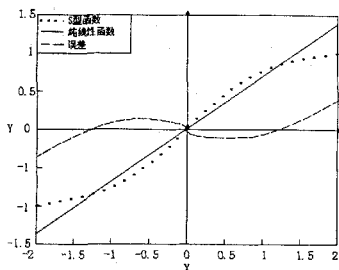


图 2 S 型传递函数

为了下面叙述方便,首先引入一些符号,假使 BP 网络的输入为 x ,网络有 s 个隐层节点,输入层到隐含层的权系数为 w_i, w_b 为偏置,则隐含层的节点为 $h_i = w_i x + w_b$ 。为了保证隐含层节点的输出落在线性区,

要求 $-1 < h_i < 1$,那么输入的变化范围为 $\frac{-1-w_b}{w_i} < x < \frac{1-w_b}{w_i}$,假设各隐层节点变化的范围相同,则

每个隐含层的节点的平均间隔为 $\frac{2}{s}$,所以, x 的取值在宽 $\frac{2}{w_i}$ 的区域内,又有 S 函数近似于直线 $y = 0.7x$,可得:

$$\frac{2}{s} = \frac{2}{w_i} \times 0.7 \Rightarrow w_i = 0.7s \quad (4)$$

对于多输入的情况,假设 r 为输入样本的行数,隐层节点被分为 m 个分段,没分段有 n 个间隔,则隐层节点数 $s = m \times n$,令

$$m = n^{r-1} \Rightarrow n = s^{\frac{1}{r}}$$

应用式(3)的结论可得: $W = n = 0.7s^{\frac{1}{r}}$

由 N-W 规则,初始权重矩阵大小应为 $s \times r$,系数为 $0.7s^{\frac{1}{r}}$,在 $[-1, 1]$ 之间均匀分布的随机数矩阵。其中 s 表示隐层节点数, r 表示输入样本的行数。

3) Nguyen-Widrow 规则改进算法^[4]

仔细研究可发现 N-W 规则存在两个问题:

* N-W 规则假设 S 函数在 $[-1, 1]$ 的区间内是线性的,但实际上两者之间存在误差。

* N-W 规则中 S 函数的线性区间 $[-1, 1]$ 是事先给定的,与网络的结构、与输入信号都没有建立联系,因此不能保证网络是最优收敛的。

在 Nguyen-Widrow 规则的基础上,提出一种新方案,即在网络训练的初期,直接应用线性函数训练网络,一方面采用线性函数计算,降低了运算量,并且可以加大每一步的更新量;另一方面根据网络误差的变化速率,自适应地确定线性区间。当网络误差变化量减小时,说明权重系数已偏离了线性区,此时将网络的权函数恢复为 S 函数。根据网络的训练情况确定线性区的宽度,增强网络训练的通用性,保证了对应具体的问题,网络都是按照最优方案进行收敛。

假设训练过程中第 l 步训练误差减小量为 e_l ,两步之间的误差减小量为 $\Delta e_l = e_{l+1} - e_l$, Δe_l 为前两步之间的误差减小量,线性传递函数的斜率为 k , α 为事先设定的一个传递函数的判决门限。具体算法如图3所示。

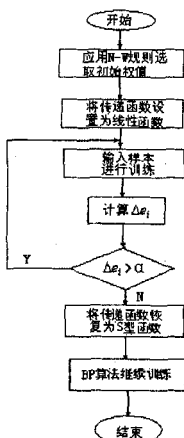


图3 改进算法流程图示意图

文中将取 α 为 $10\% \Delta e_l$, K 取为 3,改进后的网络权函数如图4所示。 k 值取较小值时,线性函数比 S 函数变化缓慢,每步更新量小,其对应的网络收敛也较慢,但 K 的取值过高,又可能导致网络陷入局部极小甚至震荡,反而使网络收敛状况变差。

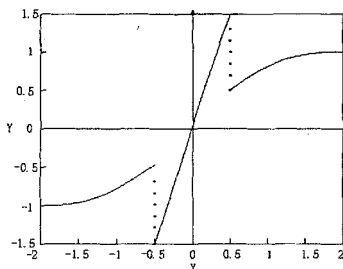


图4 改进的传递函数

2.2 学习因子 η 的优化

网络学习率 η 过大或过小,都可能会使网络无法收敛,这是因为 η 过小,会使网络的训练力度减小,使之发散;反之 η 过大,会使网络的调节作用过大,网络过分敏感,最终使网络训练处于震荡状态。采用逐步长法根据输出误差大小自动调整学习因子,来减少迭代次数和加快收敛速度。经过多次仿真实验,网络学习率 η 一般取典型值 0.9,不过也可根据具体情况进行调节。

2.3 隐层节点数的优化

隐层节点数的多少对网络性能的影响较大,当隐层节点数太多时,会导致网络学习时间过长,甚至不能收敛;而当隐层节点数过小时,网络的容错能力差。利用逐步回归分析法并进行参数的显著性检验来动态删除一些线性相关的隐节点。节点删除标准:当由该节点出发指向下一层节点的所有权值和阈值均落于死区(通常取 $\pm 0.1, \pm 0.05$ 等区间)之中,则该节点可删除。最佳隐节点数 L 可参考下面公式计算^[5]:

$$L = (m + n)^{1/2} + c \quad (5)$$

其中, m 为输入节点数; n 为输出节点数; c 为介于 1~10 的常数。

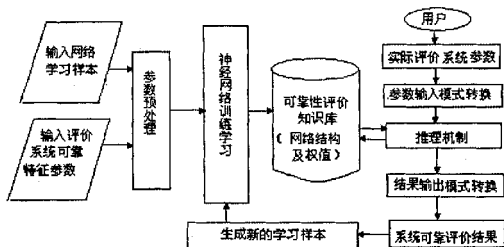
2.4 输入和输出神经元的确定

利用多元回归分析法对神经网络的输入参数进行处理,删除相关性强的输入参数,来减少输入节点数。优化 BP 神经网络可靠性预测模型如图5所示^[6]。

3 可靠性评价实例

1) 网络学习样本的选择。

根据上面提出的 BP 神经网络可靠性预测模型的建立步骤,选择了 8 台设备进行可靠性分析。将前 4 台设备的反映设备可靠状态和可靠条件的 6 个可靠评价参数作为学习样本,见表 1,根据可靠度预计的数学模型: $y_i = f(y_{i-1}, y_{i-2}, \dots, y_{i-m})$ ^[7] 对参数数据进行处理。处理如下:将 6 个可靠参数作为输入变量,以可靠



表中结果表明基于优化 BP 神经网络的系统可靠评价模型的可行性。

4 结束语

BP 神经网络理论应用于系统可靠评价中,可利用神经网络并行结构和并行处理的特征,通过适当选择评价项目,能克服可靠评价的片面性,可以全面评价系统的可靠状况和多因素共同作用下的可靠状态。运用神经网络知识存储和自适应特征,通过适应补充学习样本,实现历史经验与新知识完美结合,在发展过程中动态地评价系统的可靠状态。且利用神经网络理论的容错特征,通过选取适当的作用函数和数据结构,可以处理各种非数值性指标,实现对系统可靠状态的模糊评价。文中将优化后的 BP 神经网络应用于系统可靠评价中,人工神经网络具有自组织、自学习和非线性逼近的能力,将神经网络运用于函数逼近,应用十分广泛,为系统的可靠性预计提供了一种新的方法。

图5 优化 BP 神经网络可靠性预测模型

度 R 作为输出变量。根据上述的可靠参数处理以及对 BP 神经网络的要求,将表 1 的数据转换为对 BP 神经网络可靠性预计模型的训练样本,训练达到一定精度后停止训练。当学习因子 $\eta = 0.9$, 动量因子 $\alpha = 0.6$, 预设误差为 0.00001, 单隐层, 其隐节点数为 $L = 6$, 所得到的网络评价的结果见表 1。

表 1 设备可靠评价学习样本

网络学习样本	设备 1	设备 2	设备 3	设备 4
评价参数				
1. 资金投入	-1.5	0.5	-0.5	0.5
2. 危险源状况	-0.5	-1.5	0.5	1.5
3. 工作时间	-0.5	0.5	0.5	-0.5
4. 故障概率	-1.5	-0.5	-0.5	0.5
5. 防灾能力	0.5	0.5	1.5	-1.5
6. 安全记录	-0.5	-0.5	0.5	1.5
评价结果 R	-0.5	0.5	0.5	1.5
	可接受	好	好	非常好

2) 评价结果。

为检验网络训练的效果,将表 2 中的另 4 个设备的同样 6 个可靠参数数据输入,得到表 2。

表 2 设备神经网络可靠性评价结果

网络可靠评价结果	设备 1	设备 2	设备 3	设备 4
设备				
1	-0.5	-1.5	0.5	1.5
2	-1.5	0.5	-0.5	0.5
3	-0.5	0.5	0.5	-0.5
4	-1.5	-0.5	-0.5	0.5
5	0.5	0.5	1.5	-1.5
6	-0.5	-0.5	0.5	1.5
结论	-0.5	0.5	0.5	1.5
	可接受	好	好	非常好

(上接第 101 页)

在知识获取中有一定的使用价值。由于受计算机自动识别的能力以及规则匹配算法的限制,本系统还难以达到非常理想的效果,所以对这方面的改进将是笔者下一步的工作。

参考文献:

[1] 王文杰,叶世伟. 人工智能原理与应用[M]. 北京:人民邮电出版社,2004.

参考文献:

[1] 蒋宗礼. 人工神经网络导论[M]. 北京:高等教育出版社,2001.
 [2] 田景文,高美娟. 人工神经网络算法研究及应用[M]. 北京:北京理工大学出版社,2006.
 [3] 杜黎,陈陶. BP 网络预测模型仿真与分析[J]. 昆明理工大学学报,2003,28(5):78-80.
 [4] Widrow B, Nguyen D. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights[C]//In Proceedings of the IEEE International Conference on Networks. Amsterdam, The Netherlands: Elsevier Science Publishers B. V., 2000:21-26.
 [5] 丛爽,赵何. 反向传播网络的不足与改进[J]. 自动化博览,1999(1):25-26.
 [6] 楼顺天,施阳. 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,1998.
 [7] Reflex Software Co. & Intellect. 可靠性实用指南[M]. 北京:北京航空航天大学出版社,2005.

[2] 蔡自兴,徐光祐. 人工智能及其应用[M]. 第3版. 北京:清华大学出版社,2004.
 [3] 徐波,孙茂松,靳光瑾. 中文信息处理若干重要问题[M]. 北京:科学出版社,2003.
 [4] 曹倩,丁艳,王超等. 汉语自动分词研究及其在信息检索中的应用[J]. 计算机应用研究,2003(5):71-74.
 [5] 沈达阳. 汉语分词系统中的信息集成和最佳路径搜索方法[J]. 中文信息学报,1998,11(2):34-47.