

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Information Systems

School of Information Systems

10-2008

Determining the number of BP neural network hidden layer units

Huayu SHEN

Zhaoxia WANG

Singapore Management University, zxwang@smu.edu.sg

Chengyao GAO

Juan QIN

Fubin YAO

See next page for additional authors

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research



Part of the [Numerical Analysis and Scientific Computing Commons](#), and the [Operations Research, Systems Engineering and Industrial Engineering Commons](#)

Citation

1

This Journal Article is brought to you for free and open access by the School of Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

Author

Huayu SHEN, Zhaoxia WANG, Chengyao GAO, Juan QIN, Fubin YAO, and Wei XU

文章编号:1673-095X(2008)05-0013-03

BP神经网络隐含层单元数的确定

沈花玉¹, 王兆霞¹, 高成耀², 秦娟¹, 姚福彬¹, 徐巍¹

(1. 天津理工大学 电子信息与通信工程学院, 天津 300191; 2. 天津大学 电子信息工程学院, 天津 300072)

摘要: 本文针对BP神经网络隐含层单元数难以确定的问题,提出了一种改进的方法,并通过实验证明该方法有效的减少了验证次数,提高了确定隐含层最佳单元数的效率,具有较高的应用价值。

关键词: BP神经网络; 隐含层; 单元; 误差

中图分类号: TP183 **文献标识码:** A

Determining the number of BP neural network hidden layer units

SHEN Hua-yu¹, WANG Zhao-xia¹, GAO Cheng-yao²,
QIN Juan¹, YAO Fu-bin¹, XU Wei¹

(1. School of Electronics Information and Communications Engineering, Tianjin University of Technology, Tianjin 300191, China; 2. School of Electronic and Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract: This paper puts forward an improved method to contravene the problem which is difficult to determine the number of BP neural network hidden layer units. It is proved that the method is effective in reducing the frequency of test through experiments, and improves the efficiency of determining the best number of hidden layer units, which is more valuable in the application.

Key words: BP Neural Network; hidden layer; unit; error

BP(Back Propagation)神经网络^[1-3]是1986年由Rumelhart和McClelland为首的科学家小组提出,是一种单向传播的多层前馈网络,是目前应用最广泛的神经网络模型^[4]之一。但BP神经网络至今仍存在一个难以解决的问题:难以确定隐含层单元的个数。本文提出一种确定隐含层单元数的方法,并加以验证,证明了其可行性,具有一定的应用价值。

1 BP神经网络

BP神经网络结构如图1所示。由图1可见,BP神经网络是一种具有3层或3层以上的神经网络,包括输入层(input layer)、隐含层(hidden layer)和输出层(output layer),上下层之间实现全连接,而每层神经元之间无连接。当一对学习样本提供给网络后,神经元的激活值从输入层经各隐含层向输出层传

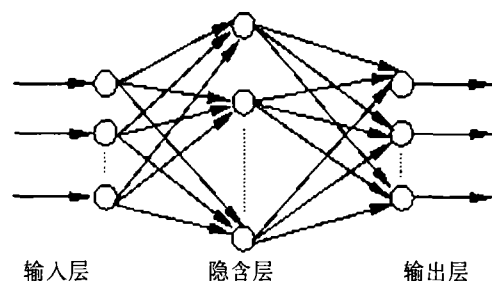


图1 BP神经网络结构示意图

Fig.1 The structure of BP neural network

播,在输出层的各神经元获得网络的输入响应。接下来,按照减少期望输出与实际输出之间误差的方向,从输出层经过各中间隐含层逐层修正各连接权值,最后回到输入层,随着这种误差逆的传播修正不断进行,网络对输入模式响应的正确率也不断上升^[5]。

收稿日期:2007-11-26.

基金项目:天津市高等学校科技发展基金(20041325);中国博士后科学基金(2005037529).

第一作者:沈花玉(1981—),女,硕士研究生.

2 隐含层单元数的确定

隐含层的单元数目选择是个十分复杂的问题, 往往需要根据设计者的经验和多次实验来确定, 因而不存在一个理想的解析式来表示. 隐含层单元的数目与问题的要求、输入/输出单元的数目都有着直接关系. 若数目太少, 则网络所能获取的用以解决问题的信息太少; 若数目太多, 不仅增加训练时间, 更重要的是隐含层单元数过多会导致学习时间过长, 且误差不一定最佳, 也会导致容错性差、不能识别以前没有训练过的样本, 还可能出现所谓“过渡吻合”(Overfitting)问题, 即测试误差增大导致泛化能力下降. 因此合理选择隐含层单元数非常重要, 一定存在一个最佳的隐单元数.

2.1 隐含层单元数的确定方法

以下3种途径可用于选择最佳隐含层单元数时的参考公式^[6].

1) $\sum_{i=0}^n C_{n_i}^i > k$, 其中, k 为样本数, n_1 为隐含层单元数, n 为输入单元数, i 为 $[0, n]$ 之间的常数, 如果.

2) $n_1 = \sqrt{n+m} \div a$, 其中 n_1 为隐含层单元数, n 为输入单元数, m 为输出单元数, a 为 $[1, 10]$ 之

的常数.

3) $n_1 = \log_2 n$, 其中, n_1 为隐含层单元数, n 为输入单元数.

在实际问题中应该选择哪种方法来确定最佳隐含层单元数? 通常的选择就是每种方法分别验证其最佳隐含层单元数, 然后将各方法得出的最佳隐含层单元数综合比较, 然后确定网络最终的最佳隐含层单元数, 这样就需要验证几十次甚至几百次. 作者提出一种方法: 由方法 1), 2), 3) 综合确定最佳隐含层单元数的边界数, n_{1_min} , n_{1_max} , 从最小单元数 n_{1_min} 开始训练网络, 逐个增加单元数, 验证到最大单元数 n_{1_max} , 一共验证 $n_{1_max} - n_{1_min} + 1$ 次, 在每种隐含层单元数情况下, 待网络收敛后, 比较在训练集相同的情况下网络的收敛速度; 最后, 根据训练结果得到的训练误差和测试误差来选择隐含层最佳单元数. 这样可以有效的减少验证次数, 从而以最快的速度找到隐含层最佳单元数.

2.2 实例验证

下面给出一个具体问题, 根据这个具体问题来验证这种确定隐含层最佳单元数的方法.

1) 问题描述. 设定样本数据^[6]如表 1 所示, 其中 1~12 为训练样本, 13~15 为测试样本:

表 1 样本数据
Tab.1 Sample data

样本序号	样本输入特征数据									类别
1	-1.781 7	-0.278 6	-0.295 4	-0.239 4	-0.184 2	-0.157 2	-0.158 4	-0.199 8		1
2	-1.871 0	-0.295 7	-0.349 4	-0.290 4	-0.146 0	-0.138 7	-0.149 2	-0.222 8		1
3	-1.834 7	-0.281 7	-0.356 6	-0.347 6	-0.182 0	-0.143 5	-0.177 8	-0.184 9		1
4	-1.880 7	-0.246 7	-0.231 6	-0.241 9	-0.193 8	-0.210 3	-0.201 0	-0.253 3		1
5	-1.415 1	-0.228 2	-0.212 4	-0.214 7	-0.127 1	-0.068 0	-0.087 2	-0.168 4		2
6	-1.287 9	-0.225 2	-0.201 2	-0.129 8	-0.024 5	-0.039 0	-0.076 2	-0.167 2		2
7	-1.523 9	-0.197 9	-0.109 4	-0.140 2	-0.009 4	-0.139 4	-0.167 3	-0.281 0		2
8	-1.678 1	-0.204 7	-0.118 0	-0.153 2	-0.173 2	-0.171 6	-0.185 1	-0.200 6		2
9	0.160 5	-0.092 0	-0.016 0	0.124 6	0.180 2	0.208 7	0.223 4	0.100 3		3
10	0.204 5	0.107 8	0.224 6	0.203 1	0.242 8	0.205 0	0.070 4	0.040 3		3
11	-1.024 2	-0.146 1	-0.101 8	-0.077 8	-0.036 3	-0.047 6	-0.016 0	-0.025 3		3
12	-0.791 5	-0.101 8	-0.073 7	-0.094 5	-0.095 5	0.004 4	0.046 7	0.071 9		3
13	-1.473 6	-0.284 5	-3.072 4	-0.210 8	-0.190 4	-0.146 7	-0.169 6	-0.200 1		1
14	-1.600 2	-0.201 1	-0.102 1	-0.139 4	-0.100 1	-0.157 2	-0.158 4	-0.279 0		2
15	-1.031 4	-0.152 1	-0.110 1	-0.080 1	-0.034 7	-0.048 2	-0.015 8	-0.030 1		3

对任何在闭区间内的连续函数,都可以用一个隐含层的BP神经网络来逼近,因而一个3层的BP神经网络可以完成任意的 n 维到 m 维的映照.因此本文从含有一个隐含层的网络开始进行训练.

下面需要确定样本数、输入单元数和输出单元数,由上表可知:一共12个样本,即样本数,每个样本中有8个输入特征数据,即输入单元数,该12组样本数据一共有1,2,3种类别,为了简化网络,用(0,1)表示第1类状态,用(1,0)表示第2类状态,用(1,1)表示第3类状态,因此在网络输出层中只设计两个输出单元就可以表示这3种状态类别,即输出单元数.

2) 隐含层单元数边界的确定.由以上分析可知:样本数,输入单元数3,输出单元数.根据上述3个参考公式可知:最佳隐含层单元数应处在3~13之间,依次从最小单元数 $n_{1_min}=3$ 开始到最大单元数 $n_{1_max}=13$ 训练网络.

3) 训练结果.本文采用MATLAB^[8]软件,训练结果如表2,图2所示.其中,表2为隐含层单元数与训练误差、测试误差的关系,图2表明,当隐含层单元数 $n_1=9$ 时,网络训练至第9次达到了目标误差,网络训练结束.

表2 隐含层单元数与训练误差、测试误差的关系1

Tab.2 The relationship 1 between the number of hidden layer units and the training error, test error

隐含层单元个数	训练误差	测试误差
3	9.583 1e-005	1.060 4e-005
4	8.410 5e-005	1.542 0e-005
5	7.978 2e-005	1.320 4e-005
6	2.630 0e-005	4.726 1e-006
7	2.847 2e-005	1.474 7e-005
8	1.903 7e-005	6.169 0e-006
9	1.269 9e-005	2.443 9e-007
10	3.933 0e-005	1.824 1e-006
11	3.210 7e-005	1.869 6e-007
12	3.104 3e-005	4.108 1e-006
13	2.313 2e-005	4.116 9e-007

由图2可以看出,增加隐含层单元数可以减少训练误差,超过10以后测试误差产生微小波动,综合比较各隐含层单元数的训练误差和测试误差,决定隐含层单元数选用9.并非隐含层单元的个数越多,网络的性能就越好,在测试本组样本数据时,训练误差随着隐含层单元的个数增加而逐渐减小,而

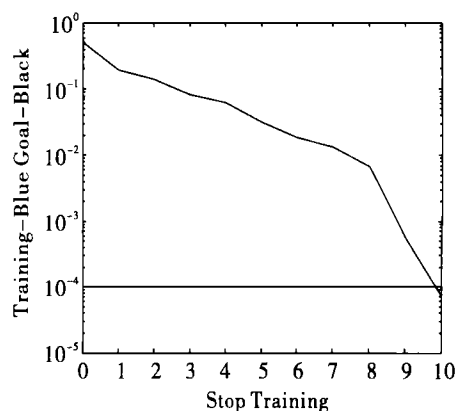


图2 隐含层单元数 $n_1=9$ 时的训练图

Fig.2 The training figure of the number of hidden layer units, $n_1=9$

在个数为9~13时,训练误差随着隐含层单元的个数增加而逐渐增大,虽然增大的幅度不是很大,但足以影响网络的性能.

对于普通的确定隐含层单元数的方法,即从隐含层单元数 $n_1=1$ 开始训练,至少一直训练到隐含层单元数 $n_1=2n+1=17$,对于该具体问题来说,需要验证17次,才能决定隐含层单元数为9.而通过本文提出的方法只验证了11次隐含层单元数便可得出最佳隐含层单元数为9,结果相同.

3 结束语

通过实验证明,用此方法减少了验证次数,可以有效且直观的来确定隐含层单元数,提高了确定隐含层最佳单元数的效率,从而也提高了确定网络结构的速度,具有一定的可行性.

参 考 文 献:

- [1] 袁曾任. 人工神经网络及其应用[M]. 北京:清华大学出版社,1999.
- [2] Dayhoff J E, Deleo J M. Artificial neural networks[J]. Cancer, 2001, 91(8):1615-1634.
- [3] 蒋宗礼. 人工神经网络导论[M]. 北京:高等教育出版社,2001.
- [4] 焦李成. 神经网络系统理论[M]. 西安:西安电子科技大学出版社,1996.
- [5] 阎平凡,张长水. 人工神经网络与模拟进化计算[M]. 北京:清华大学出版社,2000.
- [6] 飞思科技产品研发中心. 神经网络理论与MATLAB7实现[M]. 北京:电子工业出版社,2005.
- [7] 郑阿奇. MATLAB实用教程[M]. 北京:电子工业出版社,2005.