

一种基于随机 GA 的提高 BP 网络泛化能力的方法

郭海如,李志敏,万兴,熊斌

(湖北工程学院 计算机与信息科学学院,湖北 孝感 432000)

摘要:LM-BP 网络对其初始权值和阈值敏感,泛化能力不强,针对该缺点,采用遗传算法(GA)对其初始权阈值进行优化,在一定程度上能提高 LM-BP 网络的泛化能力。为进一步扩展 GA 初始种群的覆盖范围,进一步提高 LM-BP 网络的泛化能力,采用多次随机产生初始种群多次优化的方法。以伦河孝感段氟化物含量为实例,建立随机 GA 的 LM-BP 网络模型,对原始数据进行拟合及测试,结果表明该方法基本能 100% 拟合,测试误差不超过 2.3%。经过对比实验,证明了该方法的有效性。

关键词:随机遗传算法;神经网络;测试误差;泛化能力

中图分类号:TP181

文献标识码:A

文章编号:1673-629X(2014)01-0105-04

doi:10.3969/j.issn.1673-629X.2014.01.027

A Method of Improving Generalization for BP Network Based on Random GA

GUO Hai-ru, LI Zhi-min, WAN Xing, XIONG Bin

(School of Computer and Information Science, Hubei Engineering University, Xiaogan 432000, China)

Abstract: The LM-BP neural network was sensitive to its initial weight values and threshold, and it had bad generalization ability. In view of its shortcomings, the initial weights and threshold of LM-BP neural network were optimized with GA. The generalization of LM-BP neural network was improved to a certain extent. To expand the coverage of initial population, the initial populations were randomly generated iteratively and the network was optimized multi times. Thus, the generalization of LM-BP network was further improved. Take the content of fluorine in Lun River from Xiaogan as an example, the LM-BP neural network model based on random GA was established, and the raw data were fitted and tested. The results showed that the accordance of fitting data is approximately 100%, and the testing errors were less than 2.3%. Through contrast experiments, the validity of this method was proved.

Key words: random genetic algorithm; neural network; test error; generalization capability

0 引言

LM-BP 网络具有拟合能力强、全局收敛、速度快等特点,但其对初始权值和阈值敏感,泛化能力不强。文献[1]采用多次拟合测试,反复实验在一定程度上可以提高其测试精度^[1-2],但有时拟合测试上万次也不一定取得很好的效果。针对 LM-BP 网络的初始权阈值敏感的特点,采用遗传算法对 LM-BP 网络的初始权阈值进行优化,能够大大提高其泛化能力^[3]。但实验证明,经过遗传算法优化后的 LM-BP 网络也有可能得不到满意的结果。文中提出并采用随机遗传算法对 LM-BP 网络的初始权阈值进行优化,经过多

次随机优化后的神经网络泛化能力得到了进一步提高。

1 基于随机 GA 的提高 BP 网络泛化能力的方法

1.1 LM-BP 网络及遗传算法简介

标准 BP 网络的算法具有思路清晰、结构严谨、可操作性强等特点,而且一个 3 层的 BP 网络可以完成任意的 n 维到 m 维的映射,但其收敛速度慢,且存在所谓“局部最小值”问题^[4-6]。可以通过 Levenberg—Marquardt (简称 LM) 算法来克服。LM^[7] 算法利用了近似

收稿日期:2013-03-21

修回日期:2013-06-25

网络出版时间:2013-11-12

基金项目:湖北省教育科研计划重点项目(D20122606);湖北工程学院项目(Z2011009)

作者简介:郭海如(1978-),男,湖北武穴人,硕士,讲师,研究方向为智能预测、智能控制。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20131112.1639.036.html>

的二阶导数信息,它比梯度下降法快得多,实践证明,针对标准的 BP 改进算法中,LM 算法是速度最快的一种,且能克服“局部最小值”。

遗传算法是借鉴达尔文进化思想和遗传学演化出的一种随机搜索算法^[8-10]。其基本思想是:随机产生若干个所求解问题的数字编码,形成初始群体;通过适应度函数对每个个体进行评价,淘汰适应度低的个体,选择适应度高的个体参加遗传操作,生成下一代新的种群,再对这个新种群进行下一轮进化。

1.2 随机 GA 的提高 LM-BP 泛化能力模型

LM-BP 神经网络学习能力强,但对其初始权值和阈值敏感,泛化能力差。为了提高其泛化能力,采用遗传算法(GA)对 LM-BP 神经网络的初始权阈值进行优化,提出随机遗传算法的 LM-BP 模型,算法流程如图 1 所示。

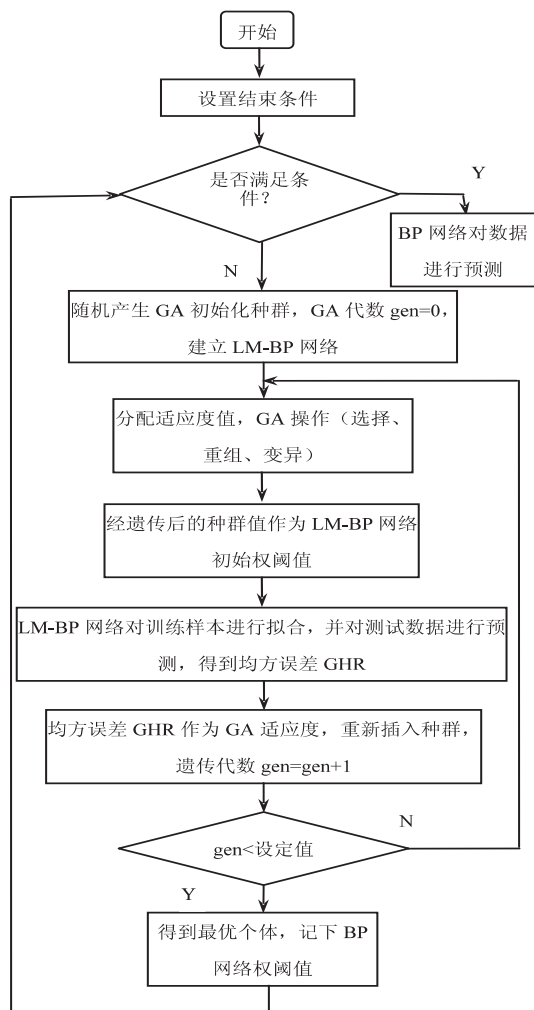


图 1 随机 GA 的 LM-BP 模型流程

其主要思想如下:

1) 首先设定结束条件,便于进行多次遗传优化,直到达到目标为止。编程时根据实际需要,可以直接设定循环次数。

2) 建立 BP 网络,设置好网络训练参数,随机初始

化 GA 种群。

3) 将 GA 种群的每组个体作为 BP 网络的初始权值和阈值,采用 LM 算法对神经网络进行训练。

4) 将训练好的神经网络对测试值进行预测,将测试值和预测结果进行比较,计算其均方误差,将均方误差作为 GA 的评价函数,计算出每个个体的适应度。

5) 根据适应度对个体进行遗传操作(选择、交叉、变异)。

6) 满足结束条件就停止,否则,转第三步,进入下一代遗传。

在编程^[11-12]的过程中,要求出所有代所有个体中适应度最小的值,并且记下此时神经网络的权值和阈值,这样就能保证网络经过多次训练多代遗传后,最终得到最优的网络参数。其中,记下最优适应度值和当时的权阈值非常重要,需要不断地比较适应度,若所记录的适应度值 ms 比当前适应度大,则记下此时的权阈值,从而保证经过 GA 优化后,能得到最优结果,其 Matlab 核心代码如下:

```

if (ms>ObjV(i,1))%记录最优适应度值
ms=ObjV(i,1);
iw=net. iw{1,1};
b1=net. b{1};
lw=net. lw{2,1};
b2=net. b{2};

```

经过随机 GA 对 LM-BP 网络进行多次多代优化后,编程过程中记下了最优适应度时的权阈值,最后需要将得到的权阈值重新赋值给 BP,此时得到的权阈值是最优结果。将权阈值赋值给 BP 网络的 Matlab 代码如下:

```

net. iw{1,1}=iw;%取最优权阈值
net. b{1}=b1;
net. lw{2,1}=lw;
net. b{2}=b2;

```

经过随机 GA 优化后,神经网络的泛化能力能得到大大增强,比文献[1]中采用多次盲目拟合测试效果好。在第四步中,GA 的适应度跟种群的个体之间是一种间接的关系,中间经过了神经网络的训练和预测,这是该模型实现的一个难点。

2 算例

采用随机 GA 的 LM-BP 模型,对近几年伦河孝感段所测的氟化物含量进行拟合、测试。取样数据每两个月一次,连续取 20 个数据,得到原始数据(经相关部门同意作过保密处理)分别为:0.25,0.32,0.37,0.3,0.35,0.3,0.27,0.32,0.33,0.92,0.56,0.57,0.55,0.34,0.55,0.57,0.36,0.35,0.41,0.46。其变化规律如图 2 所示,从图中可看出其变化规律性不强,采用 BP

网络的非线性映射能力能够很好地描述其变化规律。

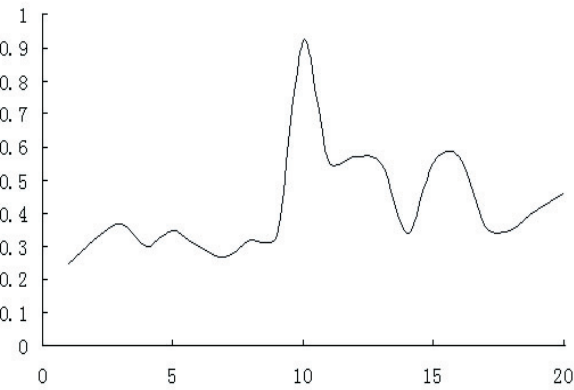


图2 原始氟化物含量变化规律

将原始氟化物含量的前 17 个数据作为训练数据,后 3 个数据作为测试数据,建立 BP 网络的训练样本如表 1 所示,测试样本如表 2 所示。训练样本和测试样本的前 8 个数据作为 BP 网络的输入,第 9 个数据作为 BP 网络的输出,测试样本不参加 BP 网络的学习训练。进一步建立 BP 网络,其输入层节点数为 8,隐含层节点数为 6(经过实验测试 6 个节点比较适合),输出层节点数为 1。经过 LM-BP 神经网络学习训练后,分别将三组测试样本的前 8 个数作为 BP 网络的输入,BP 网络将分别得到三个输出,最后分别将三组测试样本的第 9 个数据跟 BP 网络得到三个输出进行比较,得到测试误差。测试误差的大小是衡量该模型是否优越的标准。

表1 LM-BP 网络训练样本

样本1	样本2	样本3	样本4	样本5	样本6	样本7	样本8	样本9
0.25	0.32	0.37	0.3	0.35	0.3	0.27	0.32	0.33
0.32	0.37	0.3	0.35	0.3	0.27	0.32	0.33	0.92
0.37	0.3	0.35	0.3	0.27	0.32	0.33	0.92	0.56
0.3	0.35	0.3	0.27	0.32	0.33	0.92	0.56	0.57
0.35	0.3	0.27	0.32	0.33	0.92	0.56	0.57	0.55
0.3	0.27	0.32	0.33	0.92	0.56	0.57	0.55	0.34
0.27	0.32	0.33	0.92	0.56	0.57	0.55	0.34	0.55
0.32	0.33	0.92	0.56	0.57	0.55	0.34	0.55	0.57
0.33	0.92	0.56	0.57	0.55	0.34	0.55	0.57	0.36

设置遗传算法个体(31 个数据)变量的二进制位数为 20。随机产生遗传算法的初始种群,将遗传算法的个体作为 LM-BP 网络的初始权阈值,经过 LM-BP 网络训练后,对测试样本进行对比,求出其均方误差,将均方误差作为遗传算法的评价函数,进一步求出每个个体的适应度。经过反复实验将遗传的代数设为 10 较合适,再将图 1 中“结束条件”设置成循环 20 次,也就是随机产生 20 次初始种群,每次采用遗传算法对 LM-BP 网络进行 10 次(遗传代数为 10)优化。最后,

在所有的个体中记下适应度最小的值,即得到最优的测试结果。其中第 20 次经过 10 代遗传后,得到种群最优解和均值变化如图 3 所示。

表2 LM-BP 网络测试样本

测试1	测试2	测试3
0.920	0.560	0.570
0.560	0.570	0.550
0.570	0.550	0.340
0.550	0.340	0.550
0.340	0.550	0.570
0.550	0.570	0.360
0.570	0.360	0.350
0.360	0.350	0.410
0.350	0.410	0.460

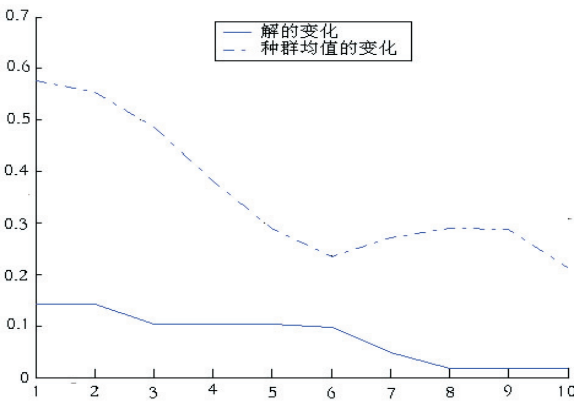


图3 第20次优化时各代种群的最优解和均值的变化

经过随机遗传算法优化后的 LM-BP 网络对表 1 中 9 个样本进行拟合,其拟合误差分别为: $-8.40\text{e}-10$, $-5.66\text{e}-11$, $-7.81\text{e}-11$, $2.79\text{e}-11$, $1.59\text{e}-11$, $1.08\text{e}-09$, $-8.27\text{e}-11$, $5.73\text{e}-09$, $2.57\text{e}-11$ 。可以看出,LM-BP 网络基本上 100% 拟合。分别将 3 个测试样本中的前 8 个数据作为 LM-BP 网络的输入,得到 3 个输出测试结果,将该测试值跟测试样本中应该得到的 3 个数据(即每个测试样本的最后一个数据)进行比较,得到测试误差见表 3 中“随机遗传算法的 LM-BP 模型”的测试误差,从测试结果可以看出,其误差不超过 2.3%,说明随机 GA 对 LM-BP 网络的初始权阈值进行优化后,其泛化能力很强。

为了进一步说明随机遗传算法的 LM-BP 网络的泛化能力,采用文献[1]中多次测试方法,未采用遗传算法对 LM-BP 网络的初始权阈值进行优化,测试 3 000 次,得到最优测试结果,如表 3 中“多次 LM-BP 模型”的测试误差。不采用随机遗传算法,采用文献[2]中方法即只用遗传算法对 LM-BP 网络的初始权阈值进行一次 10 代的遗传优化后,结果如表 3 中“遗传算

法优化的 LM-BP 模型”的测试误差。从表中可以看出,在该实例中多次 LM-BP 模型测试误差非常大,基本上属于失败的方法。经过遗传算法优化的 LM-BP 模型误差也比较大,但泛化能力能够得到大大提高。经过随机遗传算法优化的 LM-BP 模型比较容易找到合适的权阈值,测试效果比较理想,能够进一步提高其泛化能力。由于 LM-BP 网络的权阈值比较多,其初始权阈值都是随机产生,因此,采用不同方法测试时,每次结果都会不一样,这里只是取其中的一次测试数据。实验过程中也会出现不理想的情况,但是采用随机遗传算法的 LM-BP 网络更容易获得较理想的测试结果。

表 3 多种方法的测试误差对比 %

多次 LM-BP 模型	遗传算法优化的 LM-BP 模型	随机遗传算法优化的 LM-BP 模型
-10.238 0	3.198 9	2.236 9
25.383 0	4.066 9	-1.198 9
11.508 0	-10.191 0	0.531 1

3 结束语

LM-BP 网络具有全局收敛、速度快、拟合能力强的优点,采用 GA 优化 LM-BP 网络的初始权阈值,能克服 LM-BP 网络对初始权阈值敏感的缺点,能大大提高其泛化能力。初始权阈值的产生具有随机性,每次优化只局限在某次初始种群的范围进行,采用随机 GA 的方法能进一步提高 BP 网络的泛化能力。在实际应用中,为了取得更好的测试结果,可以适当提高优化的次数。

参考文献:

[1] 郭海如,冯 凯,邹 遵.基于 BP 网络的孝感学院未来数

(上接第 104 页)

参考文献:

[1] Pawlak Z. Rough sets[J]. International journal of information and computer science,1982,11:341-356.
[2] Pawlak Z. Rough set:Theoretical aspects of reasoning about data [M]. Boston, London: Kluwer Academic Publishers, 1991.
[3] Zhong N, Dong J Z, Ohsuga S. Using rough sets with heuristics for feature selection[J]. Journal of intelligent information systems,2001,16(2):199-214.
[4] 王国胤,于 洪,杨大春.基于条件信息熵的决策表约简[J].计算机学报,2002,25(7):759-766.
[5] 苗夺谦,胡桂荣.知识约简的一种启发式算法[J].计算机研究与发展,1999,36(6):681-684.
[6] Shannon C E. The mathematical theory of communication [M]. Illinois:University of Illinois Press,1963.

年招生规模预测[J].孝感学院学报,2010,30(3):60-63.
[2] Pai T Y,Tsai Y P,Loh H M,et al. Gray and neural network prediction of suspended solids and chemical oxygen demand in hospital wastewater treatment plant effluent [J]. Computers and chemical engineering,2007,31:1272-1281.
[3] Guo Hairu,Li Zhimin. A method of improving generalization ability for neural network based on genetic algorithm [C]//Proc of 2010 IEEE international conference on intelligent computing and intelligent systems(ICIS 2010). Beijing: Institute of Electrical and Electronics Engineers,Inc. ,2010:742-745.
[4] 杨建刚. 人工神经网络实用教程[M]. 杭州:浙江大学出版社,2001.
[5] Hecht-Nielsen R. Kolmogorov's mapping neural network existence theorem [C]//Proc of conf on neural networks. San Diego:IEEE,1987:11-14.
[6] 胡 康,万金泉. 基于遗传算法的控制系统在废水处理中的应用[J]. 计算机技术与发展,2011,21(2):18-21.
[7] 田 奕,乔俊飞. 基于遗传算法的 BOD 神经网络软测量[J]. 计算机技术与发展,2009,19(3):127-129.
[8] 郭海如,崔雪梅,董春玲. 一种基于神经网络模型的中国能耗预测[J]. 孝感学院学报,2007,27(3):76-79.
[9] Rumenhart D E,Hinton G E,Willians R J. Learning representation by backpropagation errors[J]. Nature,1986,323(9):533-536.
[10] 原思聪,江祥奎. 基于 GA-BP 神经网络的双目摄像机标定[J]. 西安建筑科技大学学报(自然科学版),2011,43(4):604-608.
[11] 雷英杰. Matlab 遗传算法工具箱及应用[M]. 西安:西安电子科技大学出版社,2005:107-118.
[12] FECIT. 神经网络理论与 MATLAB 7 实现[M]. 北京:电子工业出版社,2005:99-107.

[7] Skowron A. The discernibility matrix and function in information system [M]//Handbook of application and advances of the rough sets theory. Dordrecht: Kluwer Academic Publishers,1992:331-362.
[8] Guan J,Bell D. Rough computational methods for information systems[J]. Artificial intelligence,1998,105:77-103.
[9] 蒋 瑜,张 娟,林 和,等. 基于区分矩阵的决策表相容性的判断[J]. 甘肃科学学报,2006,18(2):59-61.
[10] 王国胤. Rough 理论与知识获取[M]. 西安:西安交通大学出版社,2001.
[11] 张文修,吴伟志. 粗糙集理论与方法[M]. 北京:科学出版社,2001.
[12] 杜 跃. 基于粗糙集理论的属性约简算法研究[D]. 兰州:西北师范大学,2008.

一种基于随机GA的提高BP网络泛化能力的方法

作者：[郭海如](#)，[李志敏](#)，[万兴](#)，[熊斌](#)，[GUO Hai-ru](#)，[LI Zhi-min](#)，[WAN Xing](#)，[XIONG Bin](#)

作者单位：[湖北工程学院 计算机与信息科学学院, 湖北 孝感, 432000](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

ISTIC

年，卷(期)：2014(1)

本文链接：http://d.wanfangdata.com.cn/Periodical_wjfz201401027.aspx