

混合粒子群算法优化神经网络的研究

田雨波, 潘朋朋

(江苏科技大学 电子信息学院, 江苏 镇江 212003)

摘 要: 针对 BP 神经网络初始权阈值的确定所具有的随机性和各个隐含层神经元数的不确定性, 通过利用混合粒子群优化算法来同时优化神经网络的初始权阈值和结构. 首先通过混合粒子群优化算法来确定一个较好的搜索空间, 然后在这个解空间里再通过 BP 算法对网络进行训练和学习, 搜索出最优的网络结构和权阈值. 通过 Iris 模式分类、Wine 模式分类问题和广义异或问题来对该模型进行训练和测试, 相比遗传算法等其他算法, 该模型可以获得更高的正确识别率, 结果表明此方法是可行的.

关键词: 混合粒子群; 神经网络; 分类; 广义异或

中图分类号: TP183

文献标识码: A

文章编号: 1000-7180(2011)06-0156-04

The Research of Neural Network Based on Hybrid Particle Swarm Optimization

TIAN Yu-bo, PAN Peng-peng

(School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: Aiming at the random of the determination of initial weight and the number of neurons of the hidden layers for BP neural network, this paper utilizes hybrid particle swarm optimization to optimize the initial weight and structure for the neural network. It is used to get a better search space by hybrid particle swarm optimization firstly, and then the network is trained and learned in the solution space by BP. Thereby, it searches out the optimal network structure and weight. This model is used to train and test by the use of Iris pattern classification, Wine pattern classification and generalized XOR problem, compared with other algorithm such as genetic algorithm, it can get higher recognition rate. The result shows that the way of this paper is feasible.

Key words: hybrid particle swarm optimization; neural network; classification; generalized XOR problem

1 引言

BP 神经网络是最重要的神经网络模型之一, 标准的 BP 模型具有输入层、隐含层(可以是一层或多层)、输出层, 相邻层各个神经元之间两两连接. 而同一层的神经元之间没有连接, 可以实现对任何复杂函数的映射和逼近. BP 算法的功能强大, 易于理解, 它主要用来进行分类和预测.

近年来, 一些智能算法和神经网络的融合也取得了很大的进展, 文献[1-2]主要通过遗传算法来优化神经网络的权值和结构. 文献[3]通过粒子群算法

来优化神经网络的权值. 文献[4]通过粒子群算法来优化神经网络的权值和结构.

粒子群优化算法是一种基于群体智能的进化计算技术. 算法最早是由美国的 Kennedy 和 Eberhart 教授受鸟群觅食行为的启发而提出的. 粒子群优化算法是以模拟鸟的群集智能为特征, 以求解连续变量优化问题为背景的一种优化算法. 它是通过利用个体间的协作和竞争来实现对问题最优解的搜索.

BP 算法虽然具有简单可塑的优点, 但它是基于梯度下降的方法, 对初始权阈值异常敏感, 不同的初始权阈值可能导致完全不同的结果. 而粒子群优化

收稿日期: 2010-09-07; 修回日期: 2010-12-25

基金项目: 船舶工业国防科技预研基金项目(10J3. 5. 2); 江苏省高校自然科学基金基础研究项目(07KJB510032)

算法采用实数求解,并且需要调整的参数较少,易于实现,是一种通用的全局搜索算法。

基于粒子群算法和神经网络^[5]的特点,本文选取具有两个隐含层的BP神经网络,通过混合粒子群算法来确定神经网络的初始权阈值和结构,主要表现在通过十进制粒子群优化算法优化神经网络的初始权阈值,同时通过二进制粒子群优化算法来优化各个隐含层的神经元数.将该模型应用到模式分类问题数值实验和广义异或问题上,取得了较好的效果。

2 基本粒子群优化算法

基本粒子群优化算法^[6]是对鸟群的捕食行为的一种模拟.如果一群鸟在搜索食物,在这个区域里只有一块食物,所有的鸟都不知道食物在那里,但是它们知道当前的位置离食物还有多远,那么找到食物的最优策略是寻找目前距离食物最近的鸟的周围区域。

基本粒子群优化算法的数学描述为:假设在一个 n 维的搜索空间中,由 m 个粒子组成的种群 $x = (x_1, x_1, \dots, x_m)^T$,其中,第 i 个粒子位置为 $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,n})^T$,其速度为 $v_i = (v_{i,1}, v_{i,2}, \dots, v_{i,n})^T$.它的个体极值 $p_i = (p_{i,1}, p_{i,2}, \dots, p_{i,n})^T$,种群的全局极值 $p_g = (p_{g,1}, p_{g,2}, \dots, p_{g,n})^T$.粒子根据下面两个公式来更新自己的速度与位置。

$$v_{i,d}^{k+1} = v_{i,d}^k + c_1 \text{rand}() (p_{i,d}^k - x_{i,d}^k) + c_2 \text{rand}() (p_{g,d}^k - x_{i,d}^k) \quad (1)$$

$$x_{i,d}^{k+1} = x_{i,d}^k + v_{i,d}^{k+1} \quad (2)$$

式中, c_1 和 c_2 为学习因子。 $\text{rand}()$ 为介于(0,1)的随机数; $v_{i,d}^k$ 和 $x_{i,d}^k$ 分别为粒子 i 在第 k 次迭代中第 d 维的速度和位置; $p_{i,d}^k$ 和 $p_{g,d}^k$ 分别为粒子 i 在第 d 维的个体极值和全局极值的位置。

本文针对基本粒子群在算法后期易陷入局部极值的问题,在PSO进化过程中加入了混沌的思想.混沌具有遍历性,随机性,规律性等优点,与一般的随机搜索算法相比,混沌搜索具有较强的局部搜索能力,细致搜索的有效性更强.本文取一个典型的混沌系统——Logistic映射作为混沌信号发生器,迭代公式如下:

$$x^{k+1} = \mu x^k (1 - x^k) \quad (3)$$

式中, μ 为控制参量,当 $\mu = 4$ 时, $0 \leq x^0 \leq 1$ 时,Logistic完全处于混沌状态。

混沌粒子群更新公式为

$$v_{i,d}^{k+1} = v_{i,d}^k + c_1 * \text{chaos} * (p_{i,d}^k - x_{i,d}^k) +$$

$$c_2 * \text{chaos} * (p_{g,d}^k - x_{i,d}^k) \quad (4)$$

$$x_{i,d}^{k+1} = x_{i,d}^k + v_{i,d}^{k+1} \quad (5)$$

式中, $\text{chaos} = 4 * \text{rand}() * (1 - \text{rand}())$ 。

基本粒子群算法只能应用于连续问题的求解,为解决离散问题的优化问题,Kennedy等人提出了二进制PSO算法,其中,粒子的位置的每一维只有0和1两种状态,速度更新的方法与连续基本粒子群算法相同,位置的更新则取决于由粒子速度决定的状态转移概率,速度大于一定的数值,粒子取1的可能性大,反之越小.其迭代公式如下:

$$v_{i,d}^{k+1} = v_{i,d}^k + c_1 \text{rand}() (p_{i,d}^k - x_{i,d}^k) + c_2 \text{rand}() (p_{g,d}^k - x_{i,d}^k) \quad (6)$$

$$x_{i,d}^{k+1} = \begin{cases} 1 & \rho_{i,d}^{k+1} \leq \text{sig}(v_{i,d}^{k+1}) \\ 0 & \rho_{i,d}^{k+1} > \text{sig}(v_{i,d}^{k+1}) \end{cases} \quad (7)$$

式中, $\rho_{i,d}^{k+1} \in [0,1]$ 是一随机数, $\text{sig}(\cdot)$ 为Sigmoid函数,是速度转换函数,本文选取 $\log \text{sig}(\cdot)$ 函数。

3 混合粒子群算法优化BP神经网络

3.1 编码思想

本文在通过十进制粒子群算法来优化神经网络权阈值的同时,利用二进制粒子群算法来优化网络的结构.优化网络结构主要是通过二进制粒子群算法来确定网络各个隐含层的神经元是否存在.根据式(7),若该位置变量等于1,则其对应的隐含层神经元存在,其对应的网络的权阈值存在.相反,若该位置变量等于0,则其对应的隐含层神经元不存在,其对应的网络的权阈值不存在.然后根据当前网络的结构和权阈值来计算适应度值。

$$\text{适应度评价函数为 } E = \frac{\sum_{i=1}^N (T - T')^2}{S_2 \times N} \quad (8)$$

其中, T 为网络期望输出, T' 为网络实际输出。 N 为训练样本的个数。 S_2 为输出层神经元个数。

3.2 混合粒子群算法优化BP神经网络的流程

(1) 初始化设置十进制粒子群算法和二进制粒子群算法中的基本参数.种群的大小 M ,惯性权重 w ,学习因子 c_1 和 c_2 ,在可行域内分别产生初始粒子的初始位置和初始速度,粒子群算法中的最大迭代次数 $T_{\max}$,BP算法的最大训练次数为 $t = 1$.同时根据经验初始化各个隐含层神经元的个数 $H_{\max}$ 。

(2) 依据初始化的网络结构和权阈值,计算每个粒子的初始适应度值,将 pbest_i 和 pbest_d 分别作为当前每个粒子的权阈值和结构的位置, gbest 和 gbest_d 分别作为初始化粒子中最好粒子的权阈值和

结构的位置。

(3) 通过式(4)和(5)更新所有粒子权阈值变量的速度和位置,产生一组新的粒子。

(4) 通过式(6)和(7)更新所有粒子结构的速度和位置,产生一组新的粒子。

(5) 依据当前网络的结构和权阈值来计算每个粒子的适应度值,如果第 i 个粒子的适应度值比更新前的适应度值优,则更新 $pbest_i$ 和 $pbestd_i$ 的值及其对应的位置;如果所有粒子中最好的适应值优于 $gbest$ 和 $gbestd$,则更新 $gbest$ 和 $gbestd$ 对应的位置, $t = t + 1$, 如果 $t > T_{max}$, 转到步骤(6), 否则转到步骤(3);

(6) 将通过混合粒子群优化后确定的权阈值和结构作为 BP 算法的初始权阈值和网络结构,然后再通过 BP 算法进行训练,若达到设定的最大的训练次数 T_{BP} , 则训练结束。

3.3 算法的基本参数设置

粒子种群大小为 200 个,十进制粒子群算法中,初始位置在 $[-10, +10]$ 间随机产生,初始速度在 $[-5, +5]$ 间随机产生,学习因子 $c_1 = 2.8, c_2 = 1.3$. 惯性权重从 0.9 到 0.4 线性递减. 二进制粒子群算法中,初始位置在 $[-1, +1]$ 间随机产生,初始速度在 $[-0.5, +0.5]$ 间随机产生,学习因子 $c_1 = c_2 = 2$, 惯性权重 $\omega = 1$.

4 分类问题数值实验

4.1 Iris 分类问题

Iris 分类问题是用于测试分类器性能的一个著名的基准函数数据集,它来源于 UCI 数据库. 三种不同的花各有 50 组数据,共有 150 组数据. 本文取每组数据中的前 20 组数据用于作为训练集,后 30 组数据用于作为选用测试集。

根据经验,初始两个隐含层的神经元数均取为 15 个,粒子群优化算法中的最大迭代次数 T_{max} 为 200 次,之后通过 BP 算法再对神经网络训练 5 000 次. 其中涉及 BP 算法的,均采用 MATLAB 语言神经网络工具箱中的默认值,为消除算法的随机性,通过 5 次仿真,结果如表 1 所示。

从表 1 可以看出,当网络结构为 4-8-11-3 时,总的正确分类率可以达到 99.33%,与文献[7]的结果 98% 相比,结果较好。

4.2 Wine 分类问题

Wine 数据集是测试神经网络性能经常采用的一种数据集,它来源于 UCI 数据库. Wine 数据集由

3 类共有 178 个样本数目组成,其中每个样本有 13 个属性. 本文选取其中的 89 个用于训练样本,另外 89 个用于测试样本。

表 1 Iris 数据集训练及测试结果

	网络结构	总的正确 分类率/%	训练集上正 确分类率/%	测试集上正 确分类率/%
第 1 次	4-6-6-3	98.667	100	97.778
第 2 次	4-8-11-3	99.33	100	98.889
第 3 次	4-6-9-3	98	100	96.667
第 4 次	4-8-6-3	98	100	96.667
第 5 次	4-8-7-3	98.667	100	97.778

根据经验,初始两个隐含层的神经元数均取为 25 个,粒子群优化算法中的最大迭代次数为 200 次,之后通过 BP 算法再对神经网络训练 5 000 次. 其中涉及 BP 算法的,均采用 MATLAB 语言神经网络工具箱中的默认值,为消除算法的随机性,通过 5 次仿真,结果如表 2 所示。

表 2 Wine 数据集训练及测试结果

	网络结构	总的正确 分类率/%	训练集上正 确分类率/%	测试集上正 确分类率/%
第 1 次	13-14-10-3	100	100	100
第 2 次	13-11-14-3	99.438	100	98.876
第 3 次	13-12-16-3	99.438	100	98.876
第 4 次	13-15-7-3	98.876	100	97.753
第 5 次	13-12-10-3	100	100	100

从表 2 可以看出,当网络结构为 13-14-10-3 或 13-12-10-3 时,总的正确分类率可以达到 100%,与文献[7]的结果 98.876% 相比,结果较好。

4.3 广义异或问题

异或问题是早期神经网络研究的经典问题,广义异或问题是标准异或问题的推广。

在标准异或问题中,输入 x_1 和 x_2 取离散值 -1 和 +1,而在广义异或问题中,输入 (x_1, x_2) 可以在区间 $[-1, +1] \times [-1, +1]$ 内任意取值,而输出为 $y = \text{sgn}(x_1 x_2)$, 式中, $\text{sgn}(\cdot)$ 为符号函数. 随机产生在 $[-1, +1]$ 内的 500 样本个作为训练样本,测试样本选取在 $[-1, +1]$ 内均匀分布产生的数据(间隔为 0.05)构成的不同组合作为测试样本集。

根据经验,初始两个隐含层的神经元数均取为 15 个,粒子群优化中算法的最大迭代次数 T_{max} 为 200 次,之后通过 BP 算法再对神经网络训练 5000 次. 其中涉及 BP 算法的,均采用 MATLAB 语言神经网络工具箱中的默认值,为消除算法的随机性,通

过5次仿真,结果如表3所示。

表3 广义异或问题测试集上错误分类率

	网络结构	测试集上的错误分类率/%
第1次	2-9-9-1	100
第2次	2-11-10-1	99.438
第3次	2-10-8-1	99.438
第4次	2-7-7-1	98.876
第5次	2-10-5-1	100

从表3中可以看出,当网络结构为2-9-9-1或2-11-10-1时,测试集上的错误分类率最小为1.205%,与文献[8]的结果1.324%相比,结果较好。图1为在测试集上对广义异或问题分类的结果的示意图。

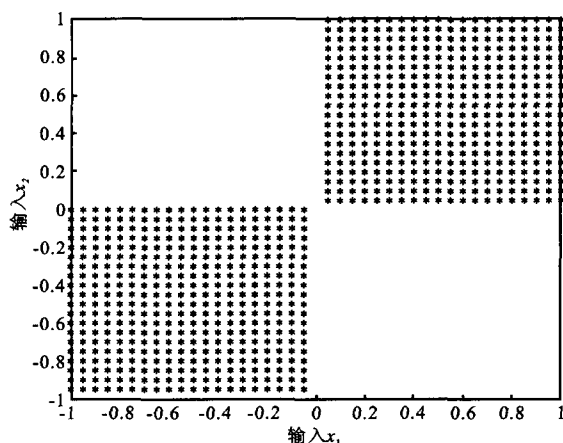


图1 测试集上的广义或问题分类示意图

5 结束语

本文通过混合粒子群优化算法同时来优化神经网络的权阈值和结构,并通过 Iris 分类、Wine 分类问题和广义异或问题进行验证,获得了较高的正确识别率,取得了较好的效果。这一方法在优化神经网络权阈值的的同时,为各个隐含层神经元数的确定给

出了一种可行的优化方法,避免了重复性的来试凑隐含层神经元数,从时间和精度上来说,时间较短,精度较高。

参考文献:

- [1] 魏志成,杨联祥,周激流. 遗传算法优化神经网络拓扑结构和权值[J]. 广西师范大学学报, 2003, 21(1): 62-65.
- [2] 田旭光,宋彤,刘宇新. 结合遗传算法优化 BP 神经网络的结构和参数[J]. 计算机应用与软件, 2004, 21(6): 69-71.
- [3] 宁东方,章卫国,田娜. 基于改进粒子群优化算法的神经网络设计[J]. 计算机应用研究, 2008, 25(11): 3343-3345.
- [4] 唐贤伦,庄陵,李银国,等. 混合粒子群优化算法优化前向神经网络结构和参数[J]. 计算机应用研究, 2007, 24(12): 91-93.
- [5] 袁冬根,刘晓东,蔡磊,等. 基于混沌蚁群算法的 BP 神经网络训练研究[J]. 微电子学与计算机, 2009, 26(4): 11-14.
- [6] 吴连贵,齐蕊,李肯立. 一种基于粒子群优化的传递函数设定方法[J]. 微电子学与计算机, 2009, 26(9): 187-195.
- [7] 周辉仁,郑丕谔,牛犇,等. 基于递阶算法和 BP 网络的模式分类方法[J]. 系统仿真学报, 2009, 21(8): 2243-2247.
- [8] 朱人杰,田雨波,贾则. 混沌搜索神经网络集成求解异或分类问题[J]. 微电子学与计算机, 2009, 26(5): 99-102.

作者简介:

田雨波 男,(1971-),博士,副教授. 研究方向为计算智能应用于电子学与电磁学,神经网络。

潘朋朋 女,(1985-),硕士研究生. 研究方向为计算智能与船舶避碰。

(上接第155页)

参考文献:

- [1] 章毓晋. 图像工程(中册),图像分析[M]. 2版. 北京:清华大学出版社,2005.
- [2] 贾晓琳,吴惠思,李爱国,等. 适用于机场跑道识别的改进 Hough 变换[J]. 小型微型计算机系统, 2005, 26(12): 2200-2202.
- [3] 朱娟,刘艳滢,王延杰. 一种基于 Hough 变换的新直线段检测算法[J]. 微电子学与计算机, 2008, 25(12): 22-25.
- [4] Matus F, Flusser J. Image representation via a finite radon

transform[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(10): 996-1006.

- [5] 曹玉珍,刘刚,杨海峰. 导盲系统中道路斑马线识别方法[J], 计算机工程与应用, 2008, 44(15): 176-178.

- [6] Farag Aly A, Delp Edward J. Edge linking by sequential search[J]. Pattern Recognition, 1995, 28(5): 611-633.

作者简介:

李竹林 女,(1972-),博士,副教授. 研究方向为数字图像处理与立体图像匹配。