

并联混合神经网络模型及应用研究

曹云忠

CAO Yun-zhong

四川农业大学 信息与工程技术学院, 四川 雅安 625014

Information & Engineering Technology Institute, Sichuan Agricultural University, Yaan, Sichuan 625014, China

E-mail: caoyz1108@163.com

CAO Yun-zhong. Model and application research on parallel hybrid neural network. *Computer Engineering and Applications*, 2009, 45(21): 218-221.

Abstract: Single neural network is difficult in performing accurate predictions for complex model. A hybrid model, which involves a radial basis function network, a multi-layer perception network with back-propagation and a control module, is proposed and used for forecasting complex system. The control module serves as a linear mapping network which combines the outputs of two neural networks to gain the final output value. The prediction methods of the hybrid model are mainly discussed: Firstly, the improved algorithm is taken to train two networks respectively and the output values are obtained; Secondly, the linear mapping network is optimized by self-adaptive genetic algorithm to gain higher prediction accuracy; Finally, this paper has carried out two experiments to compare the prediction performance of a single network and the proposed model. The experimental results show that the proposed hybrid neural network provides a superior performance in prediction accuracy than other methods and offers a common tool for complex prediction.

Key words: radial basis function; Back Propagation (BP) neural network; hybrid network model; data prediction; linear mapping

摘要: 单一神经网络难以对复杂模型做出准确的预测, 提出了一种并联型混合神经网络模型用于对复杂的系统进行预测, 该模型由径向基函数网络、BP 网络和控制模块组成。控制模块用于线性映射层, 将两种单一神经网络的输出结合并得到最终的输出结果。详细地给出了混合模型的预测方法: 首先, 利用改进算法分别训练径向基函数网络和 BP 网络; 其次, 采用自适应遗传算法优化线性映射层以获得更好的预测精度; 最后, 利用两个实例比较单一神经网络和提出的混合网络的预测性能。实验表明, 混合神经网络在预测精度上比单一网络具有更优的性能, 同时, 该混合模型为复杂系统提供了一种通用的预测工具。

关键词: 径向基函数; BP 神经网络; 混合网络模型; 数据预测; 线性映射

DOI: 10.3778/j.issn.1002-8331.2009.21.063 文章编号: 1002-8331(2009)21-0218-04 文献标识码: A 中图分类号: TP311

1 引言

人工神经网络已广泛应用于模式分类、函数逼近、时间序列预测等领域^[1], 成为了人工智能领域的研究和应用热点。对于复杂问题的建模, 往往需要大样本集对神经网络进行训练, 但大样本训练既容易造成拟合过度, 导致网络泛化能力降低, 也需要大量的训练时间。在当前日益活跃的快速预测方法中, 比如商业销售形势预测、股票价格预测等, 传统的单一神经网络往往难以胜任。鉴于此, 一些文献提出了利用复合或混合神经网络处理一些复杂的建模。陈德军等提出了用 RBF 网络作为单元并联, 构成复合网络, 将问题分割给各个子网分别处理, 最后合并得到整体的结果^[2]。孙俊等建立了串并行的两级网络, 第一级为 BP 并行网络, 第二级是一个单一 BP 网络, 处理第一级的拒识模式^[3]。田雨波等提出了基于 PSO 的多神经网络集成方法^[4]。但在时间序列预测方面, 特别是网络在线预测应用中, 上述方法在实时性方面表现不尽如人意。

对复杂系统进行预测, 并联型的混合网络模型是一种理想

的选择。利用两种异构的神经网络构造混合结构模型, 其中包括一个能进行快速建模和仿真的径向基函数网络 (RBF) 和一个 BP 神经网络, 前者能对变化的数据样本进行快速匹配处理, 后者能对样本数据进行准确的学习, 获取系统的特征, 最后, 利用线性映射网络对前述单一网络的预测结果进行修正和优化, 以实现复杂系统的精确预测。

2 并联混合神经网络模型

2.1 RBF 和 BP 神经网络

径向基神经网络是典型的前向神经网络, 它具有训练速度快、能收敛到全局最优、可最佳逼近等优点, 在函数拟合和分类上得到了广泛的应用。从原理上说, 神经网络能逼近任意复杂的非线性函数。RBF 具有很强的非线性并行处理能力和容错能力, 网络的性能主要取决于隐含层基函数的中心和宽度。RBF 网络学习需要确定 3 个参数: 网络中心、核函数宽度和隐节点到输出层的权值。

作者简介: 曹云忠 (1970-), 男, 讲师, 主要研究方向: 信息管理、智能计算。

收稿日期: 2009-02-10 **修回日期:** 2009-04-17

BP 网络是一种反向传递并修正误差的多层前馈式映射神经网络, 其研究和应用非常广泛。包括输入层、隐含层和输出层。前后层之间实现全连接。只要有足够的隐神经元节点, 网络就能实现输入到输出的映射, 具有很强的非线性映射能力, 以及自学习、自组织和自适应能力。对神经网络进行训练, 学习样本提交网络后, 神经元的激活值从输入层经中间层向输出层传播, 在输出层获得网络的输出。

2.2 混合神经网络模型结构

RBF 网络因其结构简单, 具有快速学习的能力, 对研究模型的快速变化能进行实时的跟踪, 特别是对快速跳跃函数的逼近具有十分优良的性能, 因此特别适合于快速分类和函数逼近。而 BP 网络通用性很好, 能够准确地学习样本模式, 能进行准确的仿真, 但其结构复杂, 运行速度较慢, 换句话说, BP 网络并不擅长适应模式的快速改变, 但能对函数做精确的逼近^[5]。因此将这两种神经网络的优点结合起来, 设计了图 1 所示的混合神经网络(HNN)模型。

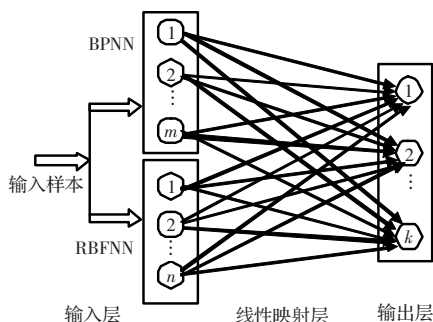


图1 并联混合神经网络模型

HNN 模型由 RBF 神经网络、BP 神经网络和一个控制模块构成, 当输入向量 X 同时送入 RBF 网络和 BP 网络, 经过控制模块 CM 进行映射后得到混合网络 HNN 的输出向量 Y 。因此, CM 可以设计成线性映射层, 参数用矩阵 C 表示, 该映射层和 HNN 描述如下:

$$C=[c_{i,j}], i=1, 2, \dots, m+n, j=1, 2, \dots, k \quad (1)$$

$$y_j = \sum_{i=1}^{m+n} c_{i,j} o_i, j=1, 2, \dots, k \quad (2)$$

其中, m 是 BP 网络输出节点的个数, n 是 RBF 网络输出节点的个数, k 是混合神经网络 HNN 输出神经元的个数, $O=[o_1, o_2, \dots, o_{m+n}]$ 是一个向量, 其元素包含 BP 网络和 RBF 网络的全部输出节点, y_j 是 HNN 第 j 神经元的输出值。

3 并联混合神经网络的算法设计

3.1 参数化 BP 神经网络算法设计

BP 神经网络中激励函数是影响网络性能的重要因素, 采用参数化的激励函数构造 BP 神经网络, 通过网络训练确定最佳网络参数。激励函数定义如下:

$$f(a, b, x) = \frac{a(1 - e^{-bx})}{1 + e^{-bx}} \quad (3)$$

其中, a 和 b 是可变参数。BP 网络参数包括神经元连接权值、阈值和激励函数的参数等, 通过梯度下降法可以推导出神经网络的参数学习公式^[6], 如下所示。

$$w_{i,j,k}^{(r)} = w_{i,j,k}^{(r-1)} - \eta \delta_{i,k} o_{j,k-1} \quad (4)$$

$$\theta_{i,k}^{(r)} = \theta_{i,k}^{(r-1)} + \eta \delta_{i,k} \quad (5)$$

$$a_{i,k}^{(r)} = a_{i,k}^{(r-1)} - \eta \xi_{i,k} \frac{o_{i,k}}{a_{i,k}} \quad (6)$$

$$b_{i,k}^{(r)} = b_{i,k}^{(r-1)} - \eta \delta_{i,k} \frac{net_{i,k}}{b_{i,k}} \quad (7)$$

其中 $\theta_{i,k}$ 是 k 层神经元 i 的阈值, $net_{i,k}$ 表示 k 层神经元 i 的输入, $w_{i,j,k}$ 是 $k-1$ 层神经元 j 与 k 层神经元 i 的连接权值, $a_{i,k}$ 和 $b_{i,k}$ 分别是 k 层神经元 i 激励函数的可变参数。 $\delta_{i,k}$ 和 $\xi_{i,k}$ 定义如下:

$$\delta_{i,k} = \frac{1}{2} \xi_{i,k} a_{i,k} b_{i,k} (1 - \frac{o_{i,k}}{a_{i,k}}) (1 + \frac{o_{i,k}}{a_{i,k}}) \quad (8)$$

$$\xi_{i,k} = \begin{cases} \sum_j \delta_{j,k+1} w_{j,i,k+1} & 1 \leq k < n \\ o_{i,n} - d_i & k = n \end{cases} \quad (9)$$

网络训练时, 在给定网络误差的条件下, 按照式(4)~(7)修改网络参数, 重复这一过程直到参数不再变化为止, 最终完成 BP 网络的训练。

3.2 RBF 网络学习算法

为了提高 RBF 网络的学习效率, 这里采用两阶段训练算法, 第一阶段完成输入层-隐含层神经元训练, 第二阶段实现隐含层-输出神经元训练^[7]。首先, 保持输出层权值不变, 采用梯度下降法训练隐含层神经元的参数, 然后保持隐含层神经元参数不变, 训练输出层神经元权值。交替进行两阶段的训练, 直到 RBF 网络满足预定的精度要求。网络的能量函数可以定义如下:

$$E = g(c_i, b_i, w_{k,i}) = \frac{1}{2} \sum_{k=1}^m (f(x) - y_k(x))^2 \quad (10)$$

其中, c_i 是隐含层第 i 神经元的中心, b_i 是第 i 神经元中心的宽度, $w_{k,i}$ 表示输出层神经元 i 和隐含层神经元 j 的连接权值。 $f(x)$ 是 RBF 神经网络的期望输出, $y_k(x)$ 是网络实际输出。两阶段神经网络训练算法表示如下:

步骤 1 初始化 c_i 和 b_i , 计算能量函数 E , 根据式(11)确定 $w_{k,i}$

$$w_{k,i}^{(t)} = w_{k,i}^{(t-1)} - \eta \frac{\partial g(c_i^{(t-1)}, b_i^{(t-1)}, w_{k,i}^{(t-1)})}{\partial w_{k,i}^{(t-1)}} \quad (11)$$

步骤 2 保持 $w_{k,i}$ 不变, 为了降低网络误差利用梯度下降法确定 b_i, c_i, b_i 和 c_i 定义如下:

$$c_i^{(t)} = c_i^{(t-1)} - \alpha \frac{\partial g(c_i^{(t-1)}, b_i^{(t-1)}, w_{k,i}^{(t-1)})}{\partial c_i^{(t-1)}} \quad (12)$$

$$b_i^{(t)} = b_i^{(t-1)} - \beta \frac{\partial g(c_i^{(t-1)}, b_i^{(t-1)}, w_{k,i}^{(t-1)})}{\partial b_i^{(t-1)}} \quad (13)$$

其中, $\eta^{(t-1)}, \alpha^{(t-1)}$ 和 $\beta^{(t-1)}$ 表示 $t-1$ 步的学习速率。

3.3 线性映射层算法设计

遗传算法具有很强的全局寻优性能, 但是传统的遗传算法也有收敛速度慢和早熟等缺陷。在最优参数搜索中, 自适应遗传算法(Self-Adaptive Genetic Algorithm, AGA)能根据个体的适应能力自适应地改变交叉和变异概率以保持个体的多样性, 因此, AGA 能利用搜索空间知识控制优化过程, 并获得求解问题的全局最优解^[8]。该文将利用 AGA 来学习混合神经网络的线性映射层参数。

3.3.1 交叉概率和变异概率的确定

传统的遗传算法中, 对于选中的个体, 无论其适应度高低,

均采用相同的交叉概率和变异概率, 迭代中当种群达到最优时, p_c 和 p_m 将会相应增加, 使得其他个体将加速趋于最优个体, 这样会破坏邻近的优良个体, 进而破坏种群的多样性, 导致早熟。因此, 需要对优良个体进行必要的保护。对于适应度高的个体, 可以降低其交叉概率和变异概率, 以保证算法收敛, 而对于适应度低的个体, 可以增加其交叉概率和变异概率, 以避免早熟发生。因此, p_c 应该由待交叉的两个个体、最大适应度 f_{\max} 和平均适应度 f_{avg} 共同确定, 相应地, 变异概率 p_m 也由个体自身适应度和最大适应度 f_{\max} 和平均适应度 f_{avg} 共同确定。 p_c 和 p_m 定义如下:

$$p_c = \begin{cases} \frac{k_1(f_{\max} - f')}{f_{\max} - f_{\text{avg}}} & f' > f_{\text{avg}} \\ k_3 & f < f_{\text{avg}} \end{cases} \quad (14)$$

$$p_m = \begin{cases} \frac{k_2(f_{\max} - f)}{f_{\max} - f_{\text{avg}}} & f > f_{\text{avg}} \\ k_4 & f < f_{\text{avg}} \end{cases} \quad (15)$$

其中, f' 是两个交叉个体中适应度的较大者, f 是变异个体的适应度。

3.3.2 染色体的编码

染色体编码方式主要包括二进制编码和十进制编码。二进制编码具有操作形象、易于理解等优点, 缺点是算法精度取决于码串的长度及待优化参数的取值范围, 而且在操作中需要不断地对码串进行数制转换, 计算量大。当处理多维数据且精度要求高的问题时, 二进制编码的遗传算法的性能非常差。该文采用十进制数字串取代二进制数字串来直接表征染色体, 既避免了编码译码过程, 又提高了算法的速率, 而且算法的计算精度大大提高。混合神经网络中性线性映射层中的权值可以定义如下:

$$C = [c_{ij}] \quad 0 \leq c_{ij} \leq 1 \quad (16)$$

其中 $i=1, 2, \dots, q, j=1, 2, \dots, k, q=m+n$ 。因此, 每个染色体中有 $q \times k$ 个基因, 如式(17)所示。

$$a_1 a_2 \dots a_m a_{m+1} \dots a_{m+n} \dots a_{q \times k} \quad (17)$$

这里, k 是混合神经网络输出节点的个数。

3.3.3 确定适应度函数

为确定遗传算法的适应度函数, 定义 E 为混合网络的误差函数:

$$E = \frac{1}{2p} \sum_p \sum_k (t_k^p - o_k^p)^2, \quad o_k = \sum_{i=1}^q y_i \cdot c_{i,k} \quad (18)$$

因此, 适应度函数可以定义为:

$$f = \begin{cases} G_{\max} - E & E < G_{\max} \\ 0 & E > G_{\max} \end{cases} \quad (19)$$

其中 G_{\max} 取进化迭代中误差 E 的最大值, 即当样本使 E 取值越大时, 其适应度就越小。

3.3.4 基于 AGA 的线性映射算法

如前所述, 可以用轮盘赌算法选中个体 s_1 和 s_2 , 按式(14)计算交叉概率, 再按式(20)得到新的个体。

$$s'_1 = r s_1 + (1-r) s_2, \quad s'_2 = r s_2 + (1-r) s_1 \quad (20)$$

这里 s'_1 和 s'_2 即交叉所得的新个体。然后对每个个体按前述算法计算变异概率, 并按式(21)执行变异操作。

$$s_i = v + (u-v) \times r \quad (21)$$

综上, 提出的线性映射层学习算法描述如下:

步骤 1 初始化线性映射层权值, 即随机建立映射层矩阵 C , 并生成遗传算法初始群体。

步骤 2 采用轮盘赌算法从群体中选出适量个体, 按式(19)计算每个个体的适应度值。

步骤 3 利用自适应方法, 根据式(14)(15)计算个体的 p_c 和 p_m , 并因此进行相应的交叉和变异操作。

步骤 4 若算法满足预定的结束条件, 再收敛退出, 否则转步骤 2。

4 算例

为验证提出的混合神经模型 HNN 在预测方面的有效性, 这里分别预测一种非连续的一元多峰函数隐含点和某股票的价格变化。

4.1 函数隐含点预测

测试样本函数定义如下:

$$y = 0.4 + \text{sinc}(4x) + 1.1 \text{sinc}(4x+2) + 0.8 \text{sinc}(6x-2) + 0.7 \text{sinc}(6x-4)$$

$$\text{sinc}(x) = \begin{cases} 1 & x=0 \\ \frac{\sin(\pi x)}{\pi x} & x \neq 0 \end{cases} \quad (22)$$

在取值范围 $[-2, 2]$ 中, 提取了 80 个数据点, 这些点均匀地分布在 $[-2, -1], [-1, 0], [0, 1]$ 和 $[1, 2]$ 中, 每个子区间有 20 个点, 隐去其中 5 个点。为了比较各种方法的预测性能, 这里分别用 RBF、BP 和该文提出的 HNN 预测隐去的 20 个数据点, 预测等式定义如下:

$$x(k+1) = f(x(k), x(k-1), \dots, x(k-w+1)) \quad (23)$$

这里, w 称为预测滑动窗口的尺度。当得到一个新的预测点时, 最久的点从滑动窗口中移除, 新的预测点则加入窗口, 并成为当前值, 以预测下一个点。窗口大小取为 10, 使用方差作为评价上述各种方法预测性能的标准, 定义如式(24)所示, \bar{x}_i 是 x_i 的预测值, n 是测试样本的个数。实验结果见表 1 和图 2 所示。

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad (24)$$

表 1 三种网络的测试结果

网络类别	评估标准(方差)	
	一步预测误差	二步预测误差
RBF	0.082 4	0.132 6
BP	0.143 7	0.237 1
HNN	0.035 7	0.085 7

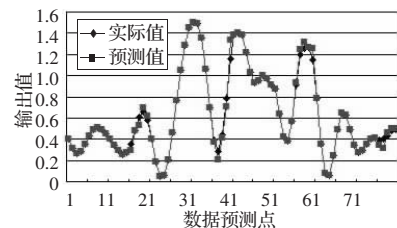


图 2 混合网络 HNN 一步预测结果

4.2 股票价格预测

股票市场是一个复杂的非线性系统, 股票价格受多种因素影响, 如国家政策、价格指数、利率和投资者信心等。已有一些

利用单一神经网络进行股票价格预测的方法, 该文采用基于技术指标模型的股票预测方法构造预测样本^[9], 分别利用 RBF、BP 和 HNN 进行预测。

该实验从某一股票 2007 年 4 月 2 日至 6 月 22 日的价格数据序列中选用 60 个数据作为训练样本, 测试时预测连续未来 5 日的股票价格, 为了比较三种方法的预测性能, 这里分别对 RBF、BP 和该文提出的 HNN 做了一步预测和第二步预测, 实验结果如表 2 和 3 所示。实验结果也表明了该文提出模型的预测精度在一步预测和三步预测方面均优于上述单一神经网络。

表 2 不同方法的测试结果(一步预测)

实际价格	三种网络及评价准则					
	HNN		BP		RBF	
	预测价格	相对误差	预测价格	相对误差	预测价格	相对误差
5.57	5.52	0.009 1	5.45	0.022 0	5.51	0.010 9
5.68	5.63	0.008 9	5.74	-0.010 5	5.63	0.008 9
5.76	5.71	0.008 6	5.59	0.030 4	5.61	0.026 7
5.77	5.88	-0.187 0	5.64	0.023 0	5.67	0.017 6
5.71	5.71	-0.007 0	5.76	-0.015 6	5.74	-0.012 2

表 3 不同方法的测试结果(二步预测)

实际价格	三种网络及评价准则					
	HNN		BP		RBF	
	预测价格	相对误差	预测价格	相对误差	预测价格	相对误差
5.57	5.44	0.023 9	5.39	0.033 4	5.41	0.029 6
5.68	5.62	0.010 7	5.51	0.030 9	5.53	0.027 1
5.76	5.61	0.026 7	5.55	0.037 8	5.58	0.032 3
5.77	5.61	0.028 5	5.54	0.041 5	5.56	0.037 8
5.71	5.79	-0.020 7	5.41	0.048 1	5.92	-0.042 2

(上接 205 页)

- [12] Lund C, Reingold N, Westbrook J, et al. On-line distributed data management[C]//Lecture Notes in Computer Science 855: Proceedings of the 2nd European Symposium on Algorithms (ESA), 1994: 202-214.
- [13] Baev I, Rajaraman R. Approximation algorithms for data placement in arbitrary networks[C]//Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms(SODA), 2001: 661-670.

(上接 211 页)

参考文献:

- [1] 张杰, 邹继刚, 李文秀. 多输入多输出系统的神经网络 PID 解耦控制器[J]. 哈尔滨工程大学学报, 2000(10): 6-9.
- [2] 陶永华, 尹怡欣, 葛芦生. 新型 PID 控制及其应用[M]. 北京: 机械工业出版社, 1998.
- [3] 刘金琨. 先进 PID 控制 MATLAB 仿真[M]. 2 版. 北京: 电子工业出版社, 2004.
- [4] 杨青, 党选举. 基于神经网络 2 自由度 PID 的解耦控制实现[J]. 计算

5 结语

单一神经网络在函数逼近和高速高精度预测方面受到了极大挑战, 提出了一种并联型混合神经网络用于实现通用的预测, 为了比较和评价不同网络模型在预测方面的性能, 分别做了复杂函数逼近预测和数据序列预测两个实验并进行验证。实验结果验证了混合神经网络具有更高的预测性能, 因此提出的混合模型是一种通用和有效的预测工具。

参考文献:

- [1] Looney C G. Pattern recognition using neural networks[M]. New York: Oxford University Press, 1997.
- [2] 陈德军, 胡华成. 基于径向基函数的混合神经网络模型研究[J]. 武汉理工大学学报, 2007, 29(2): 122-125.
- [3] 孙俊, 李正明. 基于串并行混合神经网络的汽车胎号识别方法[J]. 计算机工程与应用, 2006, 42(23): 198-200.
- [4] 田雨波, 李正强. 基于混沌 PSO 算法的选择性神经网络集成方法[J]. 计算机应用, 2008, 28(11): 2844-2846.
- [5] Barron A R. Universal approximation bounds for superpositions of a sigmoidal function[J]. IEEE Trans Information Theory, 1993, 39(3): 930-945.
- [6] 贵忠华, 刘振凯. 前馈神经网络混合学习算法研究[J]. 西北工业大学学报, 1999, 17(2): 210-215.
- [7] 杨胡萍, 白慧. 基于改进 RBF 神经网络的电力系统短期负荷预测[J]. 东华大学学报: 自然科学版, 2008, 34(2): 204-207.
- [8] Davis L. Adaptive operator probabilities in genetic algorithm[C]// Proceedings of 3rd International Conference on Genetic Algorithm. San Francisco: Morgan Kaufmann, 1989: 61-69.
- [9] 缪凯, 陈芳, 赵志刚. 基于菌群 RBF 神经网络的股票价格预测[J]. 青岛大学学报: 自然科学版, 2007, 20(2): 50-54.
- [10] Shmoys B, Tardos E, Aardal K. Approximation algorithms for facility location problems[C]//Proceedings of the 29th Annual ACM Symposium on Theory of Computing, 1997: 265-274.
- [11] 潘锐, 朱大铭, 马绍汉. 一般设施定位问题计算复杂度和近似算法研究[J]. 计算机研究与发展, 2007, 44(5): 790-797.

机工程与应用, 2004, 40(26): 197-199.

- [5] 吴星刚, 姜长洪, 姜楠. 基于双隐层 DRNN 的自整定 PID 解耦控制[J]. 计算机仿真, 2005(6): 112-114.
- [6] 陈渭红, 李守智, 王昆. 基于准对角递归神经网络(DRNN)的 PID 解耦控制研究[J]. 冶金自动化, 2007(S1): 852-855.
- [7] 舒怀林. PID 神经元网络对强耦合带时延多变量的解耦控制[J]. 控制理论与应用, 1998, 12(6): 920-924.
- [8] 杨青, 党选举. 基于 DRNN 的多变量解耦控制系统[J]. 控制理论与应用, 2004, 23(3): 14-17.
- [9] 徐丽娜. 神经网络控制[M]. 北京: 电子工业出版社, 2003.