

文章编号: 1004-4353(2021)01-0070-05

基于 RBF-BP 神经网络融合的 医学数据分类研究

金丹丹¹, 闻辉^{2*}

(1. 莆田学院 护理学院, 福建 莆田 351100; 2. 莆田学院 信息工程学院, 福建 莆田 351100)

摘要: 为改善传统的反向传播(BP)神经网络和径向基函数(RBF)神经网络的学习能力和分类性能的不足,提出一种融合 RBF 网络与 BP 网络的混合神经网络算法(HRBF-BP),并将其应用到医学数据分类问题中. 在网络结构的实现上,将 RBF 隐藏层与 BP 隐藏层进行级联融合,即在连接 BP 网络输入层与隐藏层之间加入 RBF 核映射层;在学习算法的实现上,先采用 k -均值聚类算法来实现 RBF 核参数的估计,然后再使用基于随机梯度下降的 BP 算法实现级联 BP 网络的权值优化. 将该算法与 SGBP、KMRB、PFRBF 等算法在不同的医学数据集上进行分类实验对比表明,该方法的网络训练精度以及测试精度均优于 SGBP、KMRB、PFRBF 算法;因此,该方法对提高 BP 网络和 RBF 网络的学习能力和分类性能具有良好的参考价值.

关键词: 径向基函数; 反向传播; 网络融合; 神经网络

中图分类号: TP183

文献标识码: A

Research on medical data classification based on RBF-BP neural network fusion

JIN Dandan¹, WEN Hui^{2*}

(1. School of Nursing, Putian University, Putian 351100, China;

2. School of Information Engineering, Putian University, Putian 351100, China)

Abstract: To improve the learning ability and classification performance of traditional back-propagation (BP) neural network and radial basis function (RBF) neural network, a neural network algorithm (HRBF-BP) which combines RBF network and BP network is proposed and applied to medical data classification. In the realization of the network structure, the RBF hidden layer and BP hidden layer are cascaded and fused, that is, a new RBF kernel mapping layer is added between the original BP network input layer and hidden layer; in the realization of the learning algorithm, the k -means clustering algorithm is used to realize the estimation of RBF kernel parameters, and then the BP algorithm based on random gradient descent is used to realize the weight estimation of the subsequent cascaded BP network optimization. Compared with SGBP, KMRBF and PFRBF algorithms in different medical data classification experiments, the results show that the network training accuracy and test accuracy of this method are better than SGBP, KMRBF, PFRBF. Thus, this method has a good reference value to improve the learning ability and classification performance of BP network and RBF network.

Keywords: radial basis function; back propagation; network fusion; neural network

收稿日期: 2020-12-07

* 通信作者: 闻辉(1981—),男,博士,副教授,研究方向为机器学习及神经网络.

基金项目: 福建省自然科学基金(2019J01815);莆田市科技局项目(2018RP4004);福建省教育科学“十三五”规划项目(FJJKCG20-101)

0 引言

目前,神经网络已经被广泛应用于医学诊断过程中,其中反向传播(BP)神经网络和径向基函数(RBF)神经网络是医学诊断过程中最为常用的两种网络模型. BP 神经网络的隐节点核函数一般采用统一的 sigmoid 核对输入样本进行映射,该方法虽然具有良好的泛化性能,但存在收敛速度慢和容易陷入局部最小的缺点. 目前优化 BP 神经网络的方法主要有基于全局优化的初始化权值^[1]、自适应调整学习速率^[2]、增加动量项^[3]、修正误差代价函数^[4]以及动态调整网络结构^[5]等,这些方法虽然在一定程度上改善了 BP 神经网络的不足,但对于较复杂的非线性问题其效果仍十分有限. RBF 网络隐节点核函数一般采用高斯核对输入样本进行映射,具有良好的局部响应特性. RBF 网络的优化过程主要包括隐层核参数的优化以及线性输出权值的优化,其实现过程可以分为两个阶段:一是先将原始样本通过隐层高斯核的映射,以此改善样本的可分性;二是通过优化线性超平面来完成模式分类. 目前优化 RBF 神经网络典型的方法有 k -均值聚类^[6]、模糊 c 均值聚类^[7-8]、敏感度分析^[9]、势函数优化^[10]等,但这些方法在优化复杂的非线性问题时,往往会加大 RBF 网络线性输出权值优化的负担,进而影响网络的性能. 为此,本文针对医学诊断数据集的特点(样本量少、特征复杂),提出了一种将 RBF 神经网络与 BP 神经网络相融合的 HRBF-BP 算法,并利用 UCI 机器学习数据库^[11]中的 3 个基准医学诊断数据集(Heart Disease、Diabetes、Breast Cancer)验证了本文算法的有效性.

1 HRBF-BP 网络模型的构建

HRBF-BP 模型的构建原理为:首先将原始样本输入到 RBF 隐层中不同参数下高斯核函数的映射中,以此提取原始样本在不同空间区域中的局部特征,从而形成新的特征向量;然后利用 RBF 隐层所级联的 BP 网络来完成特征空间中样本的有效分类. 相对于 BP 网络,HRBF-BP 改善了输入样本的可分性,因此可以加快网络的权值收敛速度,减少陷入局部最小的风险;相对于 RBF 网络,HRBF-BP 将原有连接 RBF 网络隐层与输出层的线性权值连接更改为非线性的 BP 网络,因此其对复杂问题具有更强的适应能力. 因此,HRBF-BP 可以将 RBF 网络隐节点处的局部非线性映射能力与 BP 网络的全局非线性分类能力有效地结合起来,从而有效改善单一结构 RBF 神经网络及 BP 神经网络的不足.

图 1 为本文所构建的 HRBF-BP 神经网络模型. 该模型由输入层、RBF 隐藏层、BP 隐藏层以及输出层 4 个部分组成,其中 RBF 隐藏层由一组不同参数的高斯核函数组成. 设 RBF 隐藏层中高斯核的个数为 K ,当输入的样本为 \mathbf{x} 时,通过隐藏层的高斯核的映射可表示为:

$$\begin{aligned} \varphi_j(\mathbf{x}) &= \exp\left(-\frac{1}{2\sigma_j^2}\|\mathbf{x} - \mu_j\|^2\right), \\ j &= 1, 2, \dots, K. \end{aligned} \quad (1)$$

式中: μ_j 为高斯核函数的中心,该值利用 k 均值聚类算法对训练样本集合进行计算来获取;

σ_j 为核宽, $\sigma = \frac{d_{\max}}{\sqrt{2K}}$, d_{\max} 为各聚类中心的最大值.

利用 RBF 网络隐节点对原始样本进行核映射后,为将所得的核映射值作为级联的 BP 网络的输入,需要对核映射值进行双极化处理. 双极化处理的公式为:

$$g_j(\mathbf{x}) = 2 \cdot \varphi_j(\mathbf{x}) - 1. \quad (2)$$

在 HRBF-BP 网络中,BP 隐藏层由 RBF 隐层到 RBF 输出层之间的各层节点组成. 由于 BP 隐藏层的

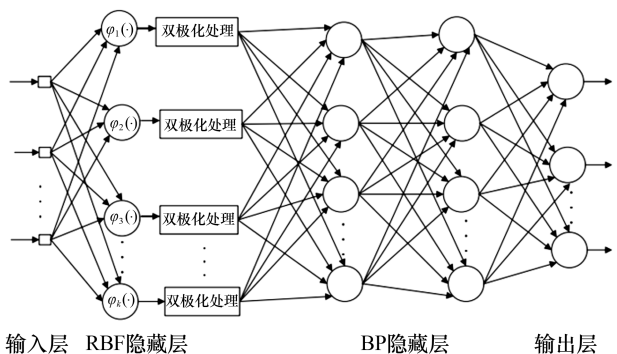


图 1 HRBF-BP 网络模型图

sigmoid 函数为双曲正切函数,因此第 1 个 BP 隐藏层中的节点 j 的输出信号可表示为:

$$y_j^{(l)} = \phi_j(v_j) = a \tanh(bv_j), \quad (3)$$

其中 a 和 b 为常数.

将 RBF 隐藏层双极化后的输出 $g_j(\mathbf{x})$ 作为 BP 网络隐藏层的输入. 当 $l=0$ 时, $g_j(\mathbf{x})$ 为 BP 网络输入层的第 j 个节点,可表示为 $y_j^{(0)} = g_j(\mathbf{x})$. 第 l 个 BP 隐藏层的第 j 个节点的诱导局部域可表示为 $v_j^{(l)} = \sum_i \omega_{ji}^{(l)} y_i^{(l-1)}$, 其中 $\omega_{ji}^{(l)}$ 是从第 $l-1$ 层的节点 i 指向第 l 层的节点 j 的权值, $y_i^{(l-1)}$ 为第 $l-1$ 层节点 i 的输出信号. 输出层中节点 k 处的输出信号可表示为 $o_k(n) = y_j^{(L)}$, 其中 L 为 BP 网络的部分输入层、隐藏层以及输出层层数的和.

2 HRBF-BP 网络算法的实现

HRBF-BP 网络算法的实现分为两个阶段:一是优化 RBF 网络隐节点处各高斯核的参数,二是优化 BP 网络中各层的权值. 图 2 为 HRBF-BP 网络算法的具体实现过程.

图 2 中, BP 网络的整体均方误差的计算公式为:

$$J(\omega) = \frac{1}{2} \sum_{j=1}^c e_j^2 = \frac{1}{2} \sum_{j=1}^c (d_j - o_j)^2,$$

其中 d_j 为 HRBF-BP 网络的目标输出, o_j 为网络的实际输出, c 为输出节点的个数.

BP 网络的反向计算即为局域梯度的更新过程,可表示为:

$$\delta_j^{(l)} = \begin{cases} e_j^{(L)} \phi_j'(v_j^{(L)}), & \text{对输出层 } L \text{ 的神经元 } j; \\ \phi_j'(v_j^{(L)}) \sum \delta_k^{(l+1)} \omega_{kj}^{(l+1)}, & \text{对隐藏层 } l \text{ 的神经元 } j. \end{cases}$$

BP 网络中第 l 层权值的更新过程为

$$\omega_{ji}^{(l)}(m+1) = \omega_{ji}^{(l)}(m) + \eta \delta_j^{(l)}(m) y_i^{(l-1)}(m),$$

其中 m 为迭代步长, η 为学习率.

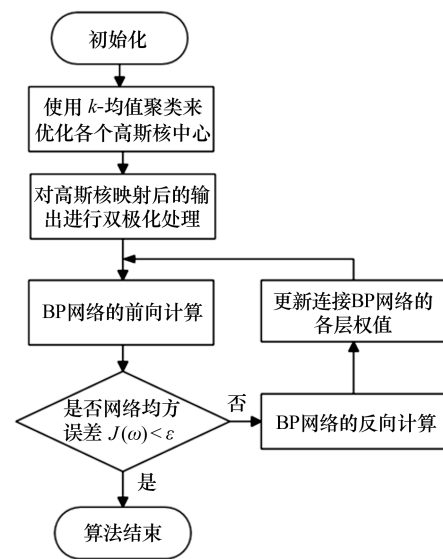


图 2 HRBF-BP 网络学习算法

3 实验与分析

为了验证本文算法的性能,利用 UCI 机器学习数据库中的 3 个不同的医学数据(见表 1)对本文算法、SGBP 算法、KMRBF 算法、PFRBF 算法的性能进行实验对比. 实验中:各数据样本全部归一化到 $[-1, 1]$; RBF 隐藏层中的核个数根据样本空间的分布进行调整(手动);分别设定 BP 隐层的层数为 1 层和 2 层,隐藏层节点的个数设置为 2~9,网络学习率 η 采用模拟退火算法进行迭代调整, sigmoid 核参数取 $a = 1.716$, $b = 0.667$. 实验运行环境为 Intel(R) Core(TM) i5-7500, 3.4 GHz CPU, 8 G RAM, MATLAB 2013, 每个实验重复 10 次.

表 1 不同分类数据集的信息

数据集	类别	特征维数	训练集样本数	验证集样本数	测试集样本数
Heart Disease (HD)	2	13	151	76	76
Diabetes	2	8	576	100	92
Breast Cancer (BC)	2	30	284	143	132

表 1 中, HD 数据集用于诊断 303 个患者是否患有心脏病. Diabetes 数据集用于诊断患者是否患有糖尿病, BC 数据集用于诊断患者是否患有乳腺癌.

3.1 基准医学分类数据集下不同算法的性能对比

在基准医学数据集下,本文算法与其他算法的性能对比结果见表 2—表 4. 由表 2—表 4 可以看出: 本文算法在学习过程中虽然比 KMRBF 和 PFRBF 算法略增加了一些训练时间,但其训练精度和测试精度均优于其他算法,这表明本文算法对训练样本空间具有更好的学习效果. 另外,由表 2—表 4 还可以看出,在本文算法中 2 个 BP 隐藏层数所需的训练时间明显少于 1 个 BP 隐藏层数所需的训练时间,且训练精度和测试精度也得到进一步提高.

表 2 HD 医学数据集下不同算法的性能

算法	隐节点个数	训练时间/s	训练精度/%	测试精度/%
SGBP (1 个隐层)	7	0.72	75.54	62.82
SGBP (2 个隐层)	5,4	0.75	78.53	64.73
KMRBF	6	0.65	79.64	75.36
PFRBF	10	0.74	80.58	77.14
本文算法(1 个 BP 隐藏层)	6,3	0.73	80.93	78.93
本文算法(2 个 BP 隐藏层)	5,3,2	0.71	81.61	79.65

表 3 Diabetes 医学数据集下不同算法的性能

算法	隐节点个数	训练时间/s	训练精度/%	测试精度/%
SGBP (1 个隐层)	7	2.35	65.80	53.82
SGBP (2 个隐层)	5,4	1.96	71.56	62.78
KMRBF	20	2.47	78.83	76.25
PFRBF	35	1.32	79.72	78.56
本文算法(1 个 BP 隐藏层)	18,7	2.18	80.49	79.31
本文算法(2 个 BP 隐藏层)	18,6,4	1.64	80.78	79.73

表 4 BC 医学数据集下不同算法的性能

算法	隐节点个数	训练时间/s	训练精度/%	测试精度/%
SGBP (1 个隐层)	7	1.86	84.38	83.62
SGBP (2 个隐层)	7,4	1.82	85.22	84.29
KMRBF	18	2.17	86.65	85.27
PFRBF	49	1.69	87.41	86.13
本文算法(1 个 BP 隐藏层)	18,7	1.75	88.52	87.72
本文算法(2 个 BP 隐藏层)	18,7,5	1.64	88.67	87.94

3.2 参数分析

以 Diabetes 医学数据集为例进行参数分析. 首先固定 BP 隐藏层层数和 BP 隐层节点个数,然后通过调节 RBF 隐层节点个数来对比 HRBF-BP 网络和 KMRBF 算法的性能,结果如图 3 所示. 由图 3 可以看出,当 RBF 隐层节点个数为 16~30, RBF 隐层节点个数为 14~30 时,其所对应的 KMRBF 算法和 HRBF-BP 算法的网络分类精度虽均可维持在一个稳定的精度范围内,但 HRBF-BP 的网络分类精度显著优于 KMRBF 算法. 该结果进一步表明,HRBF-BP 网络比 KMRBF 算法具有更好的学习能力.

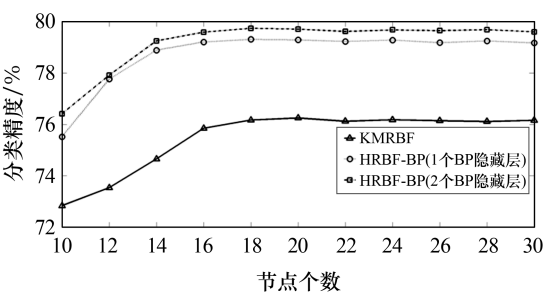


图 3 RBF 隐藏层节点个数变化时 HRBF-BP 与 KMRBF 算法的分类性能

在 Diabetes 医学数据集中,不同的 RBF 隐层节点个数、BP 隐藏层层数以及 BP 隐层节点个数对 HRBF-BP 算法性能的影响见表 5. 由表 5 可以看出,HRBF-BP 算法的训练精度和测试精度相对较为稳定,且均优于 KMRBF 算法. 这表明,HRBF-BP 算法能够有效降低对参数(RBF 隐层节点个数、BP 隐藏层层数以及 BP 隐层节点个数)选择的依赖.

表 5 不同参数下 HRBF-BP 算法的性能

BP 隐藏层 个数	隐节点 个数	训练 时间/s	训练 精度/%	测试 精度/%	BP 隐藏层 个数	隐节点 个数	训练 时间/s	训练 精度/%	测试 精度/%
1	14,4	2.35	79.85	78.82	2	14,5,7	1.83	80.38	79.28
1	14,6	2.26	80.13	78.89	2	14,7,3	1.91	80.29	79.22
1	16,5	2.28	80.27	79.18	2	16,6,3	1.75	80.49	79.56
1	18,7	2.19	80.49	79.31	2	16,6,5	1.69	80.60	79.68
1	18,8	2.21	80.39	79.27	2	18,8,4	1.95	80.71	79.62
1	20,5	2.39	80.43	79.22	2	18,6,3	1.68	80.78	79.71
1	20,7	2.28	80.26	79.14	2	20,7,3	1.89	80.67	79.69
1	22,5	2.38	80.32	79.27	2	20,9,4	1.96	80.72	79.63
1	22,7	2.41	80.35	79.18	2	22,5,7	1.98	80.78	79.69
1	24,6	2.47	80.21	79.15	2	22,7,4	1.94	80.73	79.53

4 结论

研究表明,本文提出的 HRBF-BP 算法可将 RBF 网络稳定性好与 BP 网络泛化能力强的优点融合在一起,且其训练精度和分类精度均优于 SGBP、KMRBF、PFRBF 算法;因此,本文算法可有效改善单一 RBF 网络及 BP 网络的分类性能和提高网络学习的性能. 在研究中,本文的训练样本采用的是批学习方式,未能考虑序列样本的学习方式,因此在今后的研究中我们将引入基于序列样本的 HRBF-BP 算法,以完善本文算法.

参考文献:

- [1] DING S F, SU C Y, YU J Z. An optimizing BP neural network algorithm based on genetic algorithm[J]. Artificial Intelligence Review, 2011,36(2):153-162.
- [2] VETELA J E, REIFMAN J. Premature saturation in back-propagation networks: mechanism and necessary conditions[J]. Neural Networks, 1997,10(4):721-735.
- [3] BHAYA A, KASZKUREWICZ E. Steepest descent with momentum for quadratic functions is a version of the conjugate gradient method[J]. Neural Networks, 2004,17:65-71.
- [4] RIMER M, MARTINEZ T. CB3: An adaptive error function for back propagation training[J]. Neural Processing Letters, 2006,24(1):81-92.
- [5] CHEN C H, YAO T K, KUO C M, et al. Evolutionary design of constructive multilayer feedforward neural network[J]. Journal of Vibration and Control, 2013,19(16):2413-2420.
- [6] MOODY J, DARKEN C J. Fast learning in networks of locally-tuned processing[J]. Neural Computation, 1989,1(2):281-294.
- [7] WU Q, WANG X J, SHEN Q H. Research on dynamic modeling and simulation of axial-flow pumping system based on RBF neural network[J]. Neurocomputing, 2016,186:200-206.
- [8] 张爱科,符保龙,李辉. 基于改进的模糊聚类 RBF 网络集成的文本分类方法[J]. 四川大学学报(自然科学版), 2012,49(6):1235-1239.
- [9] 韩红桂,乔俊飞,薄迎春. 基于信息强度的 RBF 神经网络结构设计研究[J]. 自动化学报,2012,38(7):1083-1090.
- [10] 黄朝辉,闻辉,车艳. 基于势函数聚类的改进 RBF 网络算法研究[J]. 延边大学学报(自然科学版), 2020,46(2):145-149.
- [11] BLAKE C, MERZ C. UCI repository of machine learning databases[EB/OL]. [2020-10-13]. <http://archive.ics.uci.edu/ml/35>.