

基于BP神经网络的临床路径优化

宋波, 辛文贤, 冯云霞

(青岛科技大学 信息科学技术学院, 山东 青岛 266061)

摘要: 针对临床路径的复杂性和模糊不确定性, 对其进行综合分析, 计算出临床路径诊疗单元的平均治疗天数以及临床路径的总治疗天数, 分别作为实验的样本数据和最终评价指标。在此基础上, 利用BP神经网络的方法对临床路径进行优化建模。并通过实验对比分析隐层神经元数量对临床路径优化结果的影响, 发现神经元数量增多时, 错误率明显下降, 迭代次数呈上下波动。最后选取7-9-1的BP神经网络结构, 以某地区三甲医院的糖尿病加高血压临床路径为例, 进行仿真实验。结果显示, 训练模型的输出值与期望值之间的相对误差范围在0-0.2%之内, 测试模型的输出值与期望值之间的相对误差范围在0-0.1%之内。说明该模型具有较强的自学习自适应能力, 能够有效地优化临床路径, 减少患者的治疗天数。

关键词: BP神经网络; 临床路径; 治疗天数; 诊疗单元; 优化

中图分类号: TP399

文献标识码: A

文章编号: 1673-629X(2020)04-0156-05

doi:10.3969/j.issn.1673-629X.2020.04.030

Clinical Path Optimization Based on BP Neural Network

SONG Bo, XIN Wen-xian, FENG Yun-xia

(School of Information Science and Technology, Qingdao University of Science and Technology,
Qingdao 266061, China)

Abstract: In view of the complexity and fuzzy uncertainty of the clinical pathway, a comprehensive analysis is carried out to calculate the average treatment days of the treatment unit and the total treatment days of the clinical pathway, respectively as the experimental sample data and the final evaluation indicators. On this basis, the BP neural network is used to optimize and model the clinical pathway. The effect of the number of hidden layer neurons on the optimization of the clinical pathway is compared and analyzed by experiment. It was found that when neurons increasing, the error rate decreases significantly and the number of iteration fluctuates up and down. Finally, the simulation experiment is performed by selecting the BP neural network structure of 7-9-1, with the clinical pathway of diabetes and hypertension in the top three hospitals in a certain area as an example. The results show that the relative error range between the output value and the expected value of the training model is within 0-0.2%, while that of the test model is within 0-0.1%. It indicates that the model has strong self-learning and self-adaptive, and can effectively optimize the clinical pathway and reduce the treatment days of patients.

Key words: BP neural network; clinical pathway; treatment days; treatment unit; optimization

0 引言

临床路径作为一套标准化的治疗模式和治疗程序, 在降低医疗费用、减少治疗天数、合理规范医务人员行为、有效利用医疗资源等方面有很大的改善^[1]。许多国内外研究者也在不断地优化临床路径治疗模式。其中, 国外研究者 Gomez Portilla 等以腹膜癌为例, 提出制定一个标准化临床路径管理计划, 以便严格评估护理计划中的发病率、死亡率和实施结果, 旨在能

够不断优化临床路径^[2]。Thomas BURKLE 基于工作流的方法开发和优化了腰神经压迫综合征临床路径系统, 以优化临床路径的执行过程为主, 但该系统并没有实际运用在电子病历中^[3]。Gang Du 等人为了提高模型的整体处理性能, 提出了基于随机协同分解粒子群优化和双变异机制的 T-S 模糊神经网络模型, 对骨肉瘤术前化疗肝肾中毒临床路径变异案例进行研究验证, 结果表明在稳定性、效率性、精确度和普遍性等均

收稿日期: 2019-03-26

修回日期: 2019-07-26

网络出版时间: 2019-12-18

基金项目: 国家自然科学基金(61572268, 61303193, 61402246); 山东省重点研发计划项目(2017GSF18110, 2018GGX101029)

作者简介: 宋波(1978-), 男, 工学博士、管理学博士后, 副教授, 研究方向为软件工程、医疗健康领域工程。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20191218.1112.022.html>

优于原来的模型^[4]。

在国内,李飞飞将调度算法和工作流技术应用到电子化临床路径管理当中,提高临床路径执行效率、保证临床路径的实施质量^[5]。赵艳丽对复杂病种的诊疗过程采用分层赋时着色 Petri 网建模,此模型对临床路径中病人状态变化、诊疗信息流转实现可视化监控^[6]。谭剑提出基于状态机工作流模型的临床路径优化方法,通过响应临床事件,降低临床路径的执行难度和增加变异的适应性^[7]。田宗梅利用聚类分析和概率相结合的方法,制定出符合当地的临床路径,提高了准确性和实施质量^[8]。

以上研究解决了临床路径优化的部分问题,但仍存在以下难点:(1)临床路径涉及检查检验、药品、护理等项目,数据量大且复杂,普通的数据挖掘技术难以发现隐藏数据关系;(2)大部分研究并没有对临床路径治疗天数或费用等进行实验分析,缺乏一定的可靠性。

在以前研究的基础之上,针对以上优化难点,文中采用 BP 神经网络与临床路径相结合的方法解决以上难点。以临床路径诊疗单元的平均治疗天数为网络的输入层,以临床路径的期望治疗天数为输出层,训练网络模型,通过修改相关参数对网络进行仿真实验,得到最优临床路径。

1 基于 BP 神经网络的临床路径整体优化结构

如图 1 所示,临床路径的优化结构分为三部分,第一部分是从医院信息管理系统中抽取临床路径数据,根据医嘱类型和治疗功效划分为诊疗单元,并计算诊疗单元的平均治疗天数;第二部分是建立 BP 神经网络临床路径优化模型,通过修改相关参数对网络进行训练;第三部分是分析输出值与期望值的相对误差,根据误差范围判断模型的有效性。

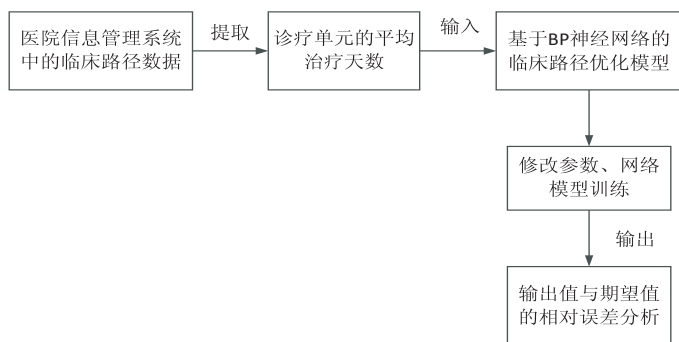


图1 基于BP神经网络的临床路径整体优化结构

2 BP 神经网络

临床路径的数据情况具有复杂性和模糊不确定性。BP 神经网络具有非线性映射^[9]、自学习自适应能力和并行分布式的存储方式^[10]以及很强的容错性,能够处理复杂的模糊数据,因此可以作为优化临床路径的方法。

2.1 BP 算法

1985 年,David Rumelhart 等科学家研究并发现了误差反向传播算法(error back propagation training,简称 BP 算法),BP 算法是由信号的正向传播和误差的反向传播两部分组成的^[11]。

BP 算法就是通过不断修改权系数对网络进行训练。基本原理如下^[12]:在正向传播过程中,训练样本通过网络的输入层传递给隐层,经隐层处理再传递给输出层,输出训练后的值。计算输出值与期望输出值是否一致,若值不一致,则进入反向传播阶段。反向传播过程中,首先计算输出层输出值与期望输出值之间的误差之和,接下来计算隐层每个神经元的输出值的

误差和权系数的误差,通过得到的误差对权系数进行更新,直至网络中的神经元的权系数都更新完毕。重新训练样本数据,若输出值与期望输出值还不一致,则继续重复以上步骤,一直训练到网络输出值与期望输出值的差在一定接受范围之内,或者是达到设定的最大迭代次数。

BP 算法的具体步骤如下^[13]:

步骤 1:初始化感知器的所有参数,包括权系数 w_i 、偏置因子 b_i 、最小值 m 和学习率参数 θ ($0 < \theta < 1$)。

步骤 2:确定输入信息 $x = (x_1, x_2, \dots, x_n)^T$,期望值 $d = (d_1, d_2, \dots, d_n)$ 。

(1)正向传播。计算隐层和输出层的各个神经元的输出值 u_j :

$$W_j = \sum_i w_{ji} u_i + b_i \quad (1)$$

$$u_j = f(W_j) \quad (2)$$

(2)反向传播。计算各隐层神经元的误差 δ_j :

$$\delta_j = u_j(1 - u_j) \sum_i w_{ij} \delta_i \quad (3)$$

其中, δ_i 是对第 $j+1$ 层的所有神经元误差信息

求和。

(3)修正权系数。

$$\Delta w_{ji} = \theta \delta_j u_i$$

(4)

(4)更新神经元的权系数。

$$w_{ji} = w_{ji} - \Delta w_{ji}$$

(5)

步骤 3:输入训练样本,重复进行第二步,直至所有的输出值与期望值之间的误差等于 0 或小于初始化

的最小值 m ,或者学习次数达到设定最大值。

2.2 BP 神经网络模型

BP 神经网络是一种多层前向神经网络模型,能够有效解决网络中隐层神经元之间的连接权系数,具有很强的自学习和自适应性、非线性映射能力,是目前应用范围最广的神经网络^[14]。BP 神经网络结构如图 2 所示。

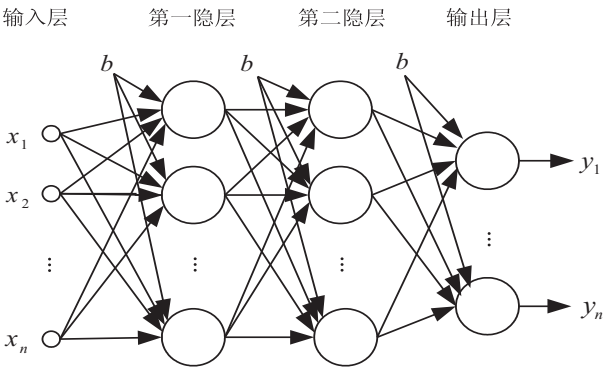


图 2 BP 神经网络模型结构

BP 神经网络包含输入层、输出层和一个或者多个隐层,图 2 中是一个包含两层隐层的 BP 神经网络。处理好的样本数据从输入层输入,经由隐层进行处理,最后通过输出层输出结果。

3 临床路径优化建模

以某地区三甲医院心内科近两年的糖尿病加高血压临床路径数据为实验数据,从医院信息管理系统中

抽取标准的糖尿病加高血压临床路径 60 例。首先计算每个诊疗单元的平均治疗天数,然后建立基于 BP 神经网络的临床路径优化模型,对网络模型进行仿真及结果分析,从而得到最优临床路径。

3.1 提取诊疗单元

诊疗单元是指按照治疗功效将药品、检查检验项目和各种护理方法组合成不同的集合^[15]。糖尿病加高血压临床路径的诊疗单元如图 3 所示。

	阶段第1日	阶段第2日	阶段第3日	阶段第4日
▶ 长期医嘱项目	内分泌科护理常规	===>	===>	===>
长期医嘱项目	长期口服药物	===>	===>	===>
长期医嘱项目	抗血小管聚集	===>	===>	===>
长期医嘱项目	长期胰岛素	===>	===>	===>
长期医嘱项目	降压药	===>	===>	===>
长期医嘱项目	改善循环	===>	===>	===>
长期医嘱项目	辅助用药	===>	===>	===>
长期医嘱项目	营养神经药物	===>	===>	===>
长期医嘱项目	改善微循环	===>	===>	===>
临时医嘱项目	常规检查			
临时医嘱项目	胰岛素取药			

图 3 糖尿病加高血压临床路径诊疗单元

根据诊疗单元的划分规则,首先将糖尿病加高血压的临床路径划分为内分泌科护理常规、长期口服药物、抗血小管聚集、长期胰岛素、降压药、改善循环和营养神经药物 7 个单元,然后计算每个诊疗单元的平均

治疗天数,取临床路径治疗天数作为最终指标,设 7 个诊疗单元为 $X_1, X_2, X_3, X_4, X_5, X_6, X_7$,治疗天数为 DAY。对数据进行提取、分析,最终得到的诊疗单元部分数据如表 1 所示。

表 1 诊疗单元

编号	X_1	X_2	X_3	X_4	X_5	X_6	X_7	DAY
1	10.6	8.6	8.5	3	3.3	2	3.6	12
2	12.3	9.6	11.1	5.4	3	2	13	14

续表 1

编号	X_1	X_2	X_3	X_4	X_5	X_6	X_7	DAY
3	13	13.8	13.4	2.6	6.2	2	12.3	20
4	3.6	3.1	2.7	3	2.9	1.3	10.6	14
5	5.4	4.8	5.2	5.5	4.7	6.9	8	10
6	3.6	5	3.6	2.3	5	3.6	5	10
7	7	10.6	13	13	10.9	9.5	3.6	14

3.2 数据归一化

BP 神经网络输入样本的数据中,不同维度的样本数据也有所不同,可能是不同维度之间的样本数据差距较大,也可能是同维度的样本数据之间相差范围较大,数据波动范围太大将会导致神经网络模型的训练时间增长、收敛效果不理想,因此在确定输入样本后,还需要对数据进行归一化处理,也就是将数据映射到区间[0,1]或者[-1,1]中。

本课题采用的归一化公式^[16]为:

$$y = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{6}$$

3.3 基于 BP 神经网络的临床路径优化模型仿真及结果分析

由于临床路径数据涉及治疗项目较多,因此建模前将临床路径划分为诊疗单元,并对诊疗单元数据进行归一化处理。将得到的 60 例样本数据中的 42 例(70%)作为训练样本,用于模型的训练,剩下的 18 例(30%)作为测试样本,检验模型的有效性。建立模型的步骤如下:

(1)BP 神经网络的输入。
糖尿病加高血压临床路径的 7 个诊疗单元即构成了 BP 神经网络的输入层。

(2)建立 BP 神经网络优化模型。
综合以上 BP 神经网络的研究以及临床路径数据的分析,建立本实验的 BP 神经网络模型结构如下:

第一层:输入层。输入的样本数据为内分泌科护理常规、长期口服药物、抗血小管聚集、长期胰岛素、降压药、改善循环和营养神经药物 7 个诊疗单元的平均治疗天数,因此输入层的神经元共有 7 个。

第二层:隐层。正向传播时,利用式(1)和式(2)计算隐层的总输出值,并将每个神经元的值传递给输出层;反向传播时,通过输出层传递的误差信号计算隐层各神经元的误差信号,并传递给输入层。因为隐层层数与神经元数量对 BP 神经网络有一定影响,因此分别选取 3、5、7、9 作为隐层的神经元数量进行实验,选取错误率最低的一个。

第三层:输出层。通过分析输出值与期望值的相对误差,就能判断模型是否具有有效性。输出值是临床路径的治疗天数,因此输出层神经元为 1 个。

基于 BP 神经网络的临床路径优化模型结构如图 4 所示。

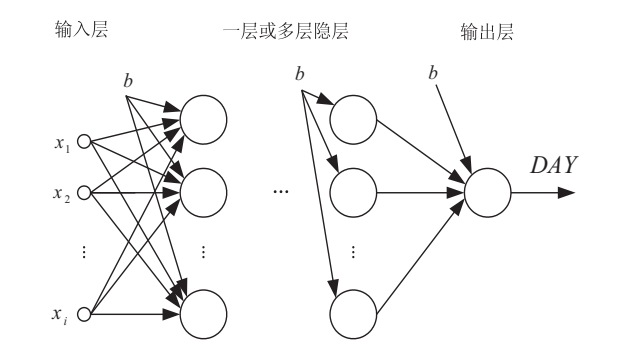


图 4 BP 神经网络模型结构

(3)确定隐层神经元数量。

选取 3、5、7、9 作为隐层的神经元数量进行分组对比实验,结果如图 5 和图 6 所示,分别反映了当神经元数量不同时,模型的错误率和迭代次数。

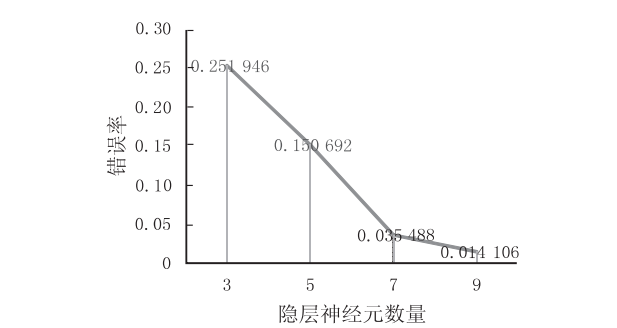


图 5 错误率

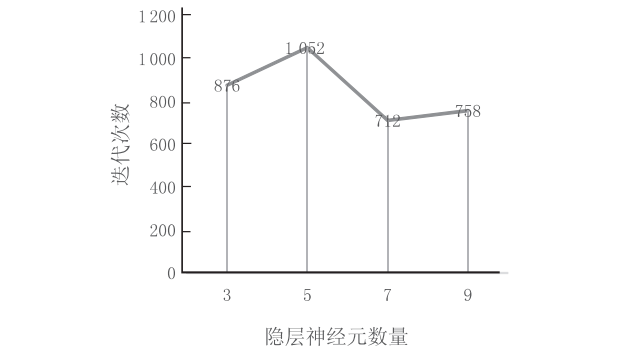


图 6 迭代次数

从图 5、图 6 中可以看出,当隐层神经元数量增加时,错误率随之降低,迭代次数出现波动,当神经元数量为 5 时迭代次数最高为 1 051 次。结合两图结果,选取 9 个神经元作为模型的隐层神经元数量。

(4) 优化模型仿真及结果分析。

对42例糖尿病加高血压临床路径的训练样本进行训练,计算输出值与期望值的相对误差,结果如图7所示。

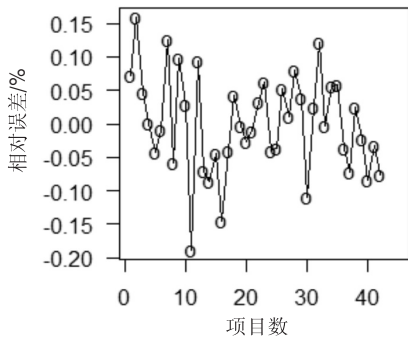


图7 训练输出相对误差

从图7中可以看出网络模型的输出值与期望输出值非常接近,相对误差的绝对值在0%~0.2%之间,这说明该网络模型具有较好的学习能力,能够通过训练来优化临床路径,有效降低了临床路径的治疗天数。

为了验证网络模型的有效性,将18条测试样本数据输入到网络模型中进行测试,输出结果的相对误差见图8。

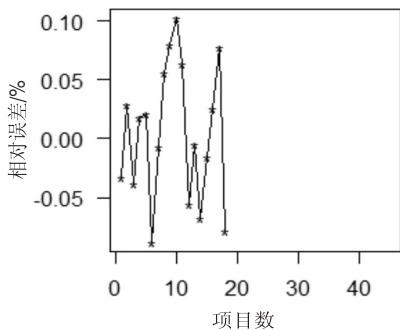


图8 测试输出相对误差

图8中显示,模型的输出值与期望值之间的相对误差的绝对值在0%~0.1%之间,测试样本的输出相对误差小于训练样本的输出相对误差,说明模型在训练过程中已经学习到了诊疗单元的平均治疗天数与临床路径总执行天数之间的关系,证明网络具有很好的自学习和自适应性,具有一定的有效性。

4 结束语

临床路径涉及检查检验、手术、药品、护理等项目,数据复杂,BP神经网络非常适合对临床路径进行优化建模研究。文中选用糖尿病加高血压临床路径作为实验样本,基于BP神经网络选择错误率最低的隐层神经元数量建立临床路径优化模型进行仿真实验。实验结果证明,该模型具有良好的自学习自适应性,并且具有一定的有效性,能够提高临床路径的治疗效果,减少

临床路径的治疗天数。因此,BP神经网络对临床路径优化方面具有一定的有效性。

参考文献:

- [1] OLSEN C A. Building critical pathways for a hospital-based home care program[J]. Outreach,1993,14(3):1-3.
- [2] GOMEZ P A, DERACO M, SUGARBAKER P H. Clinical pathway for peritoneal carcinomatosis from colon and rectal cancer:guidelines for current practice[J]. Tumori,1997,83(4):725-728.
- [3] BÜRKLE T,BAUR T,HÖSS N. Clinical pathways development and computer support in the EPR:lessons learned[J]. Studies in Health Technology and Informatics,2006,124:1025-1030.
- [4] DU G,JIANG Z,DIAO X,et al. Variances handling method of clinical pathways based on T-S fuzzy neural networks with novel hybrid learning algorithm[J]. Journal of Medical Systems, 2012,36(3):1283-1300.
- [5] 李飞飞. 基于任务调度和 workflow 技术的计算机辅助临床路径管理系统研究[D]. 上海:上海交通大学,2009.
- [6] 赵艳丽,江志斌,李娜. 基于分层赋时着色 Petri 网的临床路径建模[J]. 上海交通大学学报,2010,44(2):252-258.
- [7] 谭剑,郝立巍,程远雄,等. 基于状态机的急性加重慢性阻塞性肺疾病临床路径优化[J]. 南方医科大学学报, 2014,34(4):568-570.
- [8] 田宗梅,王莉,肖琳. 基于数据挖掘的临床路径抽取与应用[J]. 中国数字医学,2015(6):2-4.
- [9] 苏高利,邓芳萍. 论基于 MATLAB 语言的 BP 神经网络的改进算法[J]. 科技通报,2003,19(2):130-135.
- [10] 李如平,朱炼,吴房胜,等. BP 神经网络算法改进及应用研究[J]. 菏泽学院学报,2016,38(2):13-17.
- [11] 潘庆先,董红斌,韩启龙,等. 一种基于 BP 神经网络的属性重要性计算方法[J]. 中国科学技术大学学报,2017,47(1):18-25.
- [12] LIN Y C,CHEN D D,CHEN M S,et al. A precise BP neural network-based online model predictive control strategy for die forging hydraulic press machine[J]. Neural Computing and Applications,2016,29(9):585-596.
- [13] 吕晓玲,宋捷. 大数据挖掘与统计机器学习[M]. 北京:中国人民大学出版社,2016.
- [14] 孙妮妮. BP 神经网络的算法改进及应用研究[D]. 重庆:重庆大学,2009.
- [15] 张千竹. 基于大数据的临床路径优化诊疗的研究[D]. 青岛:青岛科技大学,2018.
- [16] 陈小前,罗世彬,王振国,等. BP 神经网络应用中的前后处理过程研究[J]. 系统工程理论与实践,2002,22(1):65-70.