

并结合模糊数学和层次分析法,避免了由于单一数字所造成主观判断的较大误差,该模型将待选方案的固有信息所得的评价权值以及决策者主观判断权值相结合,是一种主客观相结合的有效方法。模型结果将方案进行优先排序,为管理者进行项目管理及项目实施决策提供了科学的依据。

References:

- [1] Zhong W, Jia N, Zhou D, et al. Process of new drug R & D and essential tactics [J]. *Chin New Drugs J* (中国新药杂志), 2001, 10(12): 881-883.
- [2] Xu X S. Vague comprehensive evaluation model to risk investigation in innovation technology project [J]. *Res Mathemat Econom Technol Econom* (数量经济技术经济研究), 2001, (1): 34-36.

基于遗传神经网络的中药葛根及同属植物的质量评价研究

王 泽¹, 顾勤兰¹, 王志祥²

(1. 中国药科大学高等职业技术学院, 江苏 镇江 212003; 2. 中国药科大学药学院, 江苏 南京 210009)

摘要:遗传神经网络模型应用遗传算法对神经网络权值和学习进行优化,从而克服了BP神经网络训练速度慢、易陷入局域极小和全局搜索能力弱等缺点。以中药葛根同属植物所含的化学成分数据为考察对象,用该模型对中药葛根及同属植物进行质量评价,通过验证,识别成功率达100%,为葛根类药材的质量评价提供了一种新方法。

关键词:遗传算法;神经网络;葛根属植物;质量评价

中图分类号:R282.71

文献标识码:A

文章编号:0253-2670(2005)10-1580-03

Quality evaluation of plants in *Pueraria* DC. based on GA-BP neural network

WANG Ze¹, GU Qin-lan¹, WANG Zhi-xiang²

(1. Higher Vocational School, China Pharmaceutical University, Zhenjiang 212003, China;

2. College of Pharmacy, China Pharmaceutical University, Nanjing 210009, China)

Key words: genetic algorithm; neural network; plants in *Pueraria* DC.; quality evaluation

葛根为豆科植物野葛 *Pueraria lobata* (Willd.) Ohwi 的干燥根,是一味常用中药,具有解表退热、生津止渴和止泻之功效。我国葛根植物有11种,全国不同产地的葛根药材质量差异明显;同时葛根化学成分复杂,含有多种黄酮及三萜类等成分,而这些成分间往往是协同发挥作用的^[1]。以往对葛根的质量主要通过外部特征,如色泽、气味等进行评估,其经验性和局限性很大,而这类复杂的多变量非线性关系特别适合用人工神经网络加以解决^[2]。本文针对葛根的这一特点,建立了葛根及同属植物的遗传神经网络质量评价模型,并用该模型进行了质量评价,结果表明,将遗传神经网络技术应用于中药葛根及同属植物的质量评价,具有网络训练速度快、识别成功率高等优点。

1 网络模型的建立

遗传算法不要求目标函数连续、可微,而仅要求目标函数可计算且具有全局最优解,因而可解决BP神经网络训练速度慢、易陷入局域极小和全局搜索能力弱等缺点。本文应用遗传算法对神经网络权值和学习进行优化^[3,4],建立了遗传神经网络模型,并用该模型对葛根及同属植物进行了质量评价。

基于遗传神经网络的建模思路是先将试验数据模糊化,再作为前馈神经网络的输入和输出^[5]。同时利用遗传算法配合BP算法对网络进行训练,最后得到原始数据与有关指标之间的非线性映射模型^[6,7]。通过对已知样本的学习,调整好内部参数,即可用遗传神经网络模型对未知情况做出预测。

2 网络模型的求解方法

2.1 数据的预处理:从文献[8]可知,将收集到的17种葛属中药材,用UV法测定总黄酮及多糖的质量分数,HPLC法测定葛根素和其他3种异黄酮类成分的质量分数,并根据体外抗内毒素活性强度,将葛属中药材分为2类,第1类为有弱抗内毒素活性的葛属中药材,第2类为有强抗内毒素活性的葛属中药材,17种样本成分质量分数及已知类别见表1。

本研究以葛属植物的6种化学成分质量分数为网络输入参数,由表1可知葛属植物化学成分数据十分复杂,且不能直接用遗传神经网络模型进行学习。归一化处理可以解决遗传神经网络读取不出问题,并可保证收敛。因此,对表1中的原始数据进行归一化处理。规定遗传神经网络的各个输入量的区间范围为[0,1],某一输入量中最大值取1,最小值取0,其他输入量取0到1之间的某一值。具体取值按照下述

收稿日期:2005-03-10

作者简介:王 泽(1968—),男,上海人,中国药科大学实训中心副主任,讲师,硕士研究生,从事人工神经网络及遗传算法在药学领域的应用研究。 Tel:(0511)4432926 Fax:(0511)4422151 E-mail:wangze20041968@yahoo.com.cn

表1 17种样本成分测定结果及已知类别

Table 1 Determination of 17 sample components and their known species

编号	样本	采集地	总黄酮/%	葛根素/%	大豆苷/%	大豆苷元/%	3'-甲氧基葛根素/%	多糖/%	类别
1	野葛	山东泰安	8.512 5	4.556 3	0.088 9	0.021 9	1.129 1	4.467 1	2
2	野葛	辽宁旅顺	21.675 1	11.504 5	3.954 6	0.086 7	2.333 8	3.591 9	2
3	野葛	陕西西安	10.798 3	4.038 8	1.902 4	0.180 5	1.197 7	8.985 9	2
4 [#]	野葛	河北井泾	12.817 3	4.538 2	0.886 3	0.141 0	1.124 6	2.493 2	2
5	野葛	江西南昌	9.735 3	2.185 6	1.589 6	0.090 8	0.364 3	10.262 9	2
6	野葛	浙江临安	6.750 2	3.227 8	0.533 4	0.128 0	0.000 0	7.409 7	2
7 [#]	野葛	安徽金寨	5.442 2	2.502 5	0.408 8	0.051 4	0.165 1	13.091 0	1
8 [#]	野葛	四川峨眉	20.175 8	7.275 2	4.048 5	0.117 6	0.287 2	6.330 8	2
9	野葛	贵州岑巩	10.210 9	4.381 9	0.710 4	0.130 2	0.000 0	11.866 2	1
10	粉葛	云南景洪	4.107 3	2.772 2	0.274 7	0.222 0	0.000 0	10.753 7	1
11	粉葛	云南丽江	2.538 8	0.842 0	0.208 9	0.029 5	0.000 0	20.192 8	2
12	峨眉葛	四川峨眉	2.443 3	0.974 7	0.248 2	0.040 3	0.000 0	4.655 6	2
13	蜜花葛	云南景洪	0.930 8	0.024 7	0.025 3	0.001 0	0.000 0	17.831 4	1
14 [#]	食用葛	云南巍山	1.882 4	0.004 3	0.223 7	0.032 3	0.000 0	18.501 6	1
15	苦葛	云南巍山	1.402 4	0.000 0	0.103 6	0.016 8	0.002 1	5.182 3	2
16 [#]	三裂叶葛	广东广州	1.394 4	0.026 2	0.030 5	0.021 7	0.025 3	6.347 7	2
17	山葛	广东博罗	1.093 6	0.000 0	0.007 3	0.020 9	0.000 0	8.481 2	2

公式进行。

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

式中X表示葛根化学成分质量分数数据模糊化后某一输入量;X_{min}表示输入量中的最小值;X_{max}表示输入量中的最大值

输出量的区间范围也为[0,1]。根据葛根的体外抗毒素活性强度,将葛根分为2类,其中第1类的期望输出参数值为0.9、0.1,第2类的期望输出参数值为0.1、0.9。

2.2 网络结构的设置:在确定本模型BP网络的输入层节点数为6,输出层节点数为2之后,需确定隐含层个数以及每一隐含层的节点数。由于用3层BP神经网络即可实现任意精度的逼近,因此为提高网络运行效率,本模型选择3层BP网络。初选隐含层个数可按下式选取。初选隐含层为10,经过学习后最终确定隐含层为12。

$$k = \sqrt{n+m+l} \quad (2)$$

式中k表示隐节点数;n表示输入节点数;m表示输出节点数;l表示1~10的整数

用遗传算法来进化神经网络结构的步骤如下:(1)随机产生n个结构,对每个结构编码,每个编码个体对应一个结构;(2)用许多不同的初始权值分布对个体集中的结构进行训练;(3)根据训练的结果或其他策略确定每个个体的适应度;(4)根据若干适应度值最大的个体,直接继承给下一代;(5)对当前一代群体进行交叉和变异等遗传操作,以产生下一代群体;(6)重复步骤(2)~(5),直至当前一代群体中的某个个体(对应着一个网络结构)能满足要求为止。

2.3 遗传算法对数据的训练:对于神经网络结构,选定S型函数 $f(x) = \frac{1}{1+e^{-x}}$ 为激活函数,并按式(2)确定初始隐含层为10。遗传算法对数据的训练步骤如下。

2.3.1 采用二进制编码方案,0/1串编码,包括粒度、连接性质和权值3部分。粒度即编码长度,连接性质决定对应的两节点间有无权值相连,权值说明两节点间的联系程度。不同的初始粒度使0/1串不同。若网络的两节点之间没有连接,则在编码时,连接性质为“0”,而连接权可以随机设置。

2.3.2 适应度函数有多种选择方法,选择4种说明: $F = C - e$ 、 $F = 1/e$ 、 $F = C - E$ 、 $F = 1/E$ 。式中C表示常数;e表示误差, $e = \sum_m \sum_k (Y_{mk} - \bar{Y})$;E表示网络能量函数, $E = \sum_m \sum_k (Y_{mk} - \bar{Y}_{mk})^2$;Y_{mk}、 $\bar{Y}_{mk}$ 表示分别为第m个训练样本的第k个输出节点的期望输出与实际输出。

2.3.3 交叉:采用点交叉,将整个训练的0/1串分为粒度和结构两部分。交叉时分别对这两部分进行。

2.3.4 变异:分别对粒度、连接性质和权值3部分进行,赋予不同的变异概率,即粒度变异概率P_{mg}、连接变异概率P_{mr}以及权值变异概率P_{mw}当粒度发生变化时,其后的权值串也要相应地进行转换。

2.3.5 另外设计一个n元遗传选择算子NGA,该算子针对n个编码个体进行。设n个个体为x₁,x₂,...,x_n,其适应度值分别为F(x₁),F(x₂),...,F(x_n),并且使F(x₁)≥F(x₂)≥...≥F(x_n),另外设NGA的结果为x_{n-1}。

2.3.6 从当前父代和子代的各种取值中重新排序选择出n个适应度较大的个体作为下一代的样本,返回至步骤2.3.2中重新进行训练,终止条件为群体适应度趋于稳定或误差e小于给定值0.0001或达到预定的进化代数20000。

3 应用及结果

本研究采用的人机交互界面及遗传神经网络模型均采用VC++语言开发。软件在Windows环境下的微机上进

行。从上述网络模型建立过程可知, BP网络结构选用3层(6×12×2), 输入层节点数为6, 隐含层节点数12, 输出层节点数为2, 并由遗传算法对BP网络进行优化。网络模型建立后, 首先对遗传神经网络进行训练, 网络的训练过程实际上就是对样本的学习过程。以葛属植物的6种化学成分(见表1)为网络输入参数, 以第1类、第2类葛属植物的期望输出值(0.9、0.1; 0.1、0.9)为输出参数, 从原始数据中随机选择12个样本(表1中不带#的样本), 建立遗传神经网络训练数据库, 即输入参数(葛属植物的6种化学成分)与输出参数(葛属植物的第1类、第2类的期望输出值)的映射, 对遗传神经网络进行训练。然后, 以原始数据中的其他5个样本(表1中4、7、8、14、16号样)作为检验样本, 对遗传神经网络的识别效果进行检验, 识别结果见表2。可见该模型的识别成功率为100%, 同时训练速度明显优于一般的BP人工神经网络模型, 可用于葛根及同属植物的质量评价。

表2 检验样本识别结果

Table 2 Recognition result of unknown samples

样本	遗传神经网络输出值		已知分类	网络预测分类
4	0.099 99	0.900 79	2	2
7	0.899 98	0.100 03	1	1
8	0.108 54	0.974 18	2	2
14	0.900 35	0.099 16	1	1
16	0.098 42	0.913 91	2	2

4 结论

4.1 建立了葛根及同属植物质量评价的遗传神经网络模型, 该模型以随机选取原始数据样本的化学成分的质量分数为输入参数, 以质量等级分类的网络期望输出值为输出参数, 并通过检验样本对该网络的识别性能进行检验。

4.2 由于遗传算法对神经网络的权值和学习过程进行了优化, 因而解决了神经网络训练速度慢、易陷入局部极小和全局搜索能力弱等缺点。应用表明, 本研究建立的模型具有快捷、精确的优点, 可用于葛根及同属植物的质量评价。

References:

- [1] Zheng M, Zhang H M, Zheng S Q, et al. Studies on *Radix Pueraria* (*Pueraria lobata*) by chemical pattern recognition [J]. *Chin Tradit Herb Drugs* (中草药), 1998, 29(10): 652-654.
- [2] Hussian A S, Yu X, Johnson R D, et al. Application of neural computing in pharmaceutical product development [J]. *Pharm Res*, 1999, 16(1): 1-16.
- [3] Maniczzo V. Genetic evolution of the topology and weight distribution of neural networks [J]. *IEEE Trans Neural Networks*, 1994, 5(1): 39-53.
- [4] Harrald P G, Kamstra M. Evolving artificial neural networks to combine financial fore case [J]. *IEEE Trans Evolve Comp*, 1997, 1(1): 40-52.
- [5] Gatonicvic-Kustrin S, Beresford R. Basic concepts of artificial networks (ANNS) modeling and its application in pharmaceutical research [J]. *J Pharm Anal*, 2000, 22(4): 717-727.
- [6] Li X C, Zhu W X, Chen G, et al. Artificial neural networks based gears material selection hybrid intelligent system [J]. *Acta Metall Sin*, 2003, 16(6): 543-546.
- [7] Li X C, Chen G, Zhu W X, et al. Intelligent expert system used in gear mate rail selection and its heat treatment [J]. *Acta Metall Sin*, 2004, 40(10): 1051-1054.
- [8] Zhang H M, Zheng M, Zheng S Q, et al. Studies on root of *Pueraria* DC. to differentiate it from other species of same genus by pattern recognition (I) [J]. *Chin Tradit Herb Drugs* (中草药), 2001, 32(3): 253-254.

中药(复方)药动力学研究现状及研究方法探讨

贺福元^{1,3}, 罗杰英¹, 邓凯文², 刘西京³, 杨勇³

(1. 湖南中医学院药学院 药剂教研室, 湖南 长沙 410007; 2. 湖南中医学院第一附属医院, 湖南 长沙 410007;

3. 成都中医药大学药学院, 四川 成都 611730)

摘要: 根据目前中(西)药药动力学研究的方法, 分析中药学科的特点, 结合生物数学知识及多年单成分药动力学研究工作, 分析了中药药动力学研究现状及亟待解决的问题, 提出了中药多成分体系药动力学研究新方法——总量统计矩法, 该法既能满足中医宏观要求(宏观量化可算)又能满足现代科学微观可测(可按目前单个成分药动力学方法进行)的要求。适应中医药理论要求的新的中药(复方)药动力学研究方法可以在目前单个成分(指标)的药动力学研究方法的基础上建立和发展。

关键词: 中药复方; 药动力学; 总量统计矩; 数学模型; 谱效学; 药数学

中图分类号: R285.61

文献标识码: A

文章编号: 0253-2670(2005)10-1582-05

收稿日期: 2005-02-13

基金项目: 国家自然科学基金资助项目(30472199); 国家中医药管理局中医药科研基金资助项目(02-03ZP41)

作者简介: 贺福元(1965—), 湖南中医学院副教授, 现为成都中医药大学 2003 级中药药剂学博士研究生, 主要从事中药药剂提取工艺、生物药剂学及中医药信息数学分析, 共发表专业论文 40 余篇。 Tel: (0731) 5686164 E-mail: pharmsarking@tom.com