

遗传算法结合反向传播算法神经网络优化 党参多糖的提取工艺

张 丽, 田 密, 李 凯*

(北京振东光明药物研究院有限公司, 北京 100089)

摘 要: **目的** 通过遗传算法结合反向传播算法(back propagation, BP)神经网络, 与正交实验结果对比, 优化党参多糖的提取工艺。**方法** 以党参多糖的提取率为指标, 采用三因素(提取次数、提取时间、料液比)水平对 BP 神经网络模型参数进行优化, 建立网络模型, 并采用遗传算法对 BP 神经网络进行目标寻优, 得到党参多糖的最佳提取工艺。**结果** 得到的最优提取工艺为提取次数 3 次, 提取时间 2 h, 料液比为 1:10(*m:m*), 在此条件下党参多糖得率预测值为 55.29 mg/g, 与实际测量值的相对误差仅为 1.10%, 具有较好的网络预测性。**结论** 利用遗传算法结合 BP 神经网络对党参多糖的提取工艺进行优化具有较好的预测性, 可以为党参的进一步研究开发提供新的思路。

关键词: 党参; 遗传算法; 反向传播算法神经网络; 正交实验; 多糖

Genetic algorithm combined with back propagation neural network to optimize the extraction process of *Codonopsis pilosula* polysaccharide

ZHANG Li, TIAN Mi, LI Kai*

(Beijing Zhendong Guangming Pharmaceutical Research Institute Co., Ltd., Beijing 100089, China)

ABSTRACT: Objective To optimize the extraction process of polysaccharides from *Codonopsis pilosula* by genetic algorithm combined with back propagation(BP)-based neural network and orthogonal experiments. **Methods** Taking the extraction amount of polysaccharide as the index, the parameters of BP neural network model were optimized by 3 factors (extraction times, extraction time, ratio of material to liquid), and the network model was established. Then the genetic algorithm was used to optimize the BP neural network and obtain the optimal extraction process of the *Codonopsis pilosula* polysaccharide. **Results** The optimal process conditions were as follows: extraction times three times, extraction time 2 h, the ratio of material to liquid is 1:10(*m:m*), the predicted value of this condition was 55.29 mg/g, and the relative error of the actual measured value was only 1.10%, which had good network prediction. **Conclusion** Using genetic algorithm combined with BP neural network to optimize the extraction process of *Codonopsis pilosula* polysaccharide has good predictability, which can provide new ideas for further research and development of *Codonopsis pilosula*.

KEY WORDS: *Codonopsis pilosula*; genetic algorithm; back propagation neural network; orthogonal test; polysaccharide

*通讯作者: 李凯, 主要研究方向为保健食品开发与应用。E-mail: 308218025@qq.com

*Corresponding author: LI Kai, Beijing Zhendong Guangming Pharmaceutical Research Institute Co., Ltd., Beijing 100089, China. E-mail: 308218025@qq.com

1 引言

党参为中国传统补益药,古代以山西上党地区出产的党参为上品,故称为党参^[1]。党参效近人参而为较弱,甘平。归脾、肺经,健脾益肺,养血生津。用于脾肺气虚,食少倦怠,咳嗽虚喘,气血不足,面色萎黄,心悸气短,津伤口渴,内热消渴^[1]。近年来,国内外一系列研究表明,党参具有调节胃肠运动、抗溃疡、抗辐射、提高缺氧耐受力、增强免疫、延缓衰老等功能^[2]。党参是一种常用的保健药材,多见于食疗方面的使用,近年来也越来越多的用于保健食品的研制开发,发展前景广阔。

党参中含有糖类、苷类、氨基酸类、三萜类等成分^[2],其中多糖是组成党参糖类物质的主要成分,也是党参的主要活性成分。目前多数学者采用正交实验法对党参多糖提取工艺进行优化^[3,4],但是由于党参多糖的提取控制是一个非常复杂的非线性系统、各因素的耦合性较强,且易受到外界因素的干扰,导致党参多糖提取率较低、杂质较多、质量不可控。为了获得纯度高、杂质少、质量稳定的党参多糖,提高党参的利用价值,建立一种高效的党参多糖的提取方法非常必要。

反向传播算法(back propagation, BP)神经网络是一种按误差逆传播算法训练的多层前馈网络,是迄今为止最成功的神经网络学习算法^[5-7]。BP网络能学习和存贮大量的输入-输出模式映射关系,且无需预先揭示这种映射关系的数学方程。它的学习规则是使用最速下降法,通过反向传播来不断调整网络的权值和阈值,使网络的误差平方和最小,广泛应用于非线性建模。但是神经网络模型是通过神经元间的连接权值与阈值来实现的,所以在解决最优化问题时,无法使用传统的优化方法来实现。遗传算法(genetic algorithm)是借鉴生物界的进化规律演化而来的随机搜索方法^[8,9]。遗传算法从问题解的串集开始搜索,覆盖面大,利于全局择优;遗传算法可以同时处理群体中的多个个体,减少了陷入局部最优解的风险,同时算法本身易于实现并行化。

与正交实验法相比,BP神经网络具有学习规则、自身训练的能力,不仅能在给定区域优化,还可以在预测网络下进行优化,使得实验结果更加可靠、稳定,并且神经网络结合遗传算法寻优,不需要目标函数具备明确的数学表达式,就可以快速得到最优的提取条件,此方法在细脚拟青霉多糖^[10]、密花石斛多糖^[11]、香菇多糖^[12]、黄芪多糖^[13]、合欢皮多糖^[14]的提取工艺优化中均取得较好效果。

本研究以党参多糖的提取率为指标,采用三因素(提取次数、提取时间、料液比)水平对BP神经网络模型参数进行优化,建立网络模型,并采用遗传算法对BP神经网络进行目标寻优,得到党参多糖的最佳提取工艺,以期对党

参的进一步研究开发提供新的思路。

2 材料与方法

2.1 实验仪器

TLE204 电子天平(瑞士梅特勒托利多仪器有限公司); T10CS 紫外分光光度计(北京普析通用仪器有限责任公司); G560E 旋涡仪(太仓市华利达实验设备有限公司); KH20R 多功能离心机(湖南凯达科学仪器有限公司)。

2.2 试剂与样品

苯酚(分析纯,东莞市聚鹏化学有限公司); 浓硫酸(分析纯,国药集团化学试剂有限公司); 磷酸氢二钠、无水乙醇、磷酸二氢钠(分析纯,天津市福晨化学试剂厂); 实验水为超纯水; 对照品: D-无水葡萄糖(110833-201707, 中国食品药品检定研究院)。

党参饮片(批号: 2018061504, 山西振东制药有限公司)。

2.3 实验方法

2.3.1 党参粗多糖测定

(1)标准曲线的绘制

取 D-无水葡萄糖对照品 0.5 g, 精密称定后加水溶解并定容至 50 mL, 临用前将此溶液稀释 100 倍为使用液(0.1 mg/mL)。精密吸取 D-无水葡萄糖标准使用液 0、0.10、0.20、0.40、0.60、0.80、1.00 mL 于 25 mL 比色管中, 补加水至 2.0 mL, 精密加入 5% 的苯酚溶液 1.0 mL, 漩涡混匀, 加入浓硫酸 10 mL, 漩涡混匀, 于沸水浴中加热 2 min, 冷却至室温, 用紫外分光光度计于 485 nm 处, 试剂为空白, 1 cm 比色皿测定吸光度值。以 D-无水葡萄糖质量浓度(mg/mL)为横坐标, 吸光度值为纵坐标, 绘制标准曲线。

(2)供试品溶液的制备

精密称取党参药材粉末 2.0 g, 转移至 100 mL 容量瓶中, 加水 80 mL 左右, 于沸水浴中加热 1 h, 冷却至室温后补加水至刻度, 混匀后过滤, 弃去初滤液, 收集剩下的滤液, 作为样品提取液。取 50 mL 样品提取液置于 100 mL 具塞锥形瓶中, 冷却至 60 °C 以下, 加 0.5 mL 0.2 mol/L 磷酸盐缓冲液, 再加适量的糖化酶(约为样液体积的 1%)于 60 °C 以下水解 60 min 后取出(用碘液检验是否水解完全, 如不完全可延长水解时间至酶解液加碘液不变蓝色为止), 于电炉上小心加热至沸(灭酶), 冷却, 定容, 过滤, 滤液沉淀粗多糖。精密吸取上述滤液 5.0 mL 于 50 mL 离心管中, 加入无水乙醇 20 mL, 混匀, 4 °C 静置 4 h 以上, 4000 r/min 离心 5 min, 上清液弃去, 残渣用 80% 乙醇溶液洗涤, 继续离心, 弃去上清液, 反复操作 3 次。残渣用水溶解并定容至 20 mL。准确吸取供试品溶液 0.5 mL, 置于 25 mL 比色管中, 补加水至 2.0 mL, 测定吸光度值。从标准曲线上读出 D-无水葡萄糖含量, 计算样品中粗多糖含量。

2.3.2 正交设计

采用热水浸提法对党参多糖进行提取。确定影响党参多糖提取率的 3 个因素: 提取次数(A)、提取时间(B)、料液比(C), 分别取 3 个水平, 以党参多糖的提取率为指标, 考察各因素对党参多糖提取率的影响。按照 $L_9(3^4)$ 正交表, 因素与水平设计见表 1。

表 1 党参多糖提取水平因素表
Table 1 Table of factors for extracting polysaccharides from *Codonopsis pilosula*

水平	提取次数 (A)/n	提取时间 (B)/h	料液比 (C)/(g:g)	空白(D)
1	1	1.0	1:6	1
2	2	1.5	1:8	2
3	3	2.0	1:10	3

2.3.3 BP 神经网络模型设计

(1)BP 神经网络模型

本研究以党参多糖的提取中影响较大的 3 个因素作为 BP 神经网络的输入, 以党参多糖的提取率作为 BP 神经网络的输出, 采用 3 层网络。模型拓扑结构为 4-n-1, 其中 4 代表输入的因素个数、1 代表输出的响应值个数, n 代表隐层节点数, 隐层节点数的选择一般是通过经验公式和试错法确定^[3]。BP 神经网络模型建模和遗传算法寻优均在 Matlab R2012a 软件中进行。

(2)神经网络训练样品选择

本研究采用党参多糖提取正交实验得到 9 组样本数据, 但为了获得更加可信的 BP 神经网络, 需要增加样本数据量进行训练, 因此增加了虚拟样本量, 在实际生产中, 由于原料产地、批次差异、操作条件不同, 检测仪器本身的限制等, 输入的变量和实际值会存在微小的误差, 这种误差是不可避免的, 那么在此误差范围内的输入值与输出值应该是对应的。虚拟样本的生成方法是在每个实际样本的数据上增加一个 $\pm \Delta i$ 值, 本研究 Δi 取 $\pm 0.2\%$, 根据 $L_9(3^4)$ 正交设计表, 使每个样本产生 8 个虚拟样本, 由此得到 72 个虚拟样本, 增加了训练样品的代表性, 参加神经网络训练的样本一共 81 个, 增加了训练空间内的样本密度。

(3)BP 神经网络程序代码

将原始样本数据导入 Matlab R2012a 软件中, 随机分为训练样本(70 个)和检验样本(11 个)。测试预测值与实际值在不同隐层节点数下的误差平方和, 由此确定隐层节点数。

2.3.4 遗传算法的寻优设计

根据 2.3.3 已经建立好的 BP 神经网络模型, 采用遗传算法对党参多糖提取率进行寻优, 并对提取因素的各水平进行预测。在 Matlab R2012a 中输入 optimtool 打开优化工具箱, 求解器选择 GA-Genetic Algorithm, 变量数选择 4,

变量上下限分别为 [1 1 6 1] 和 [2 3 10 3], 初始种群 20, 种群类型选双精度, 种群选择方法 Roulette, 其他选择默认, 适应度函数为 @fit。

3 结果与分析

3.1 党参多糖标准曲线的绘制

吸光度 A 与 D-无水葡萄糖对照品溶液浓度(mg/mL)的线性方程 $Y=4.799X+0.0055(r^2=0.9994)$ 。结果表明, D-无水葡萄糖对照品在质量浓度 0.01~0.10 mg/mL 内呈良好的线性关系。

3.2 正交实验结果分析

按 2.3.2 项下的正交设计进行正交实验, 并分别检测在不同提取条件下所得的党参多糖的提取率, 结果见表 2, 方差分析结果见表 3。

表 2 党参多糖的正交实验结果
Table 2 Orthogonal test results of crude polysaccharides from *Codonopsis pilosula*

序号	提取次数	提取时间	料液比	空白	党参多糖提取量/(mg/g)
1	1	1	1	1	32.81
2	1	2	2	2	46.84
3	1	3	3	3	41.08
4	2	1	2	3	41.58
5	2	2	3	1	50.35
6	2	3	1	2	43.33
7	3	1	3	2	48.60
8	3	2	1	3	50.35
9	3	3	2	1	52.11
K_1	40.24	41.00	42.16	45.09	
K_2	45.09	49.18	46.84	46.26	
K_3	50.35	45.51	46.68	44.34	
R	10.11	8.18	4.69	1.92	

表 3 党参多糖正交实验方差分析结果
Table 3 Analysis of variance in orthogonal test of crude polysaccharides from *Codonopsis pilosula*

因素	偏差平方和	自由度	F 比	F 临界值	显著性
A	153.408	2	27.321	19.000	*
B	100.800	2	17.952	19.000	
C	42.300	2	7.533	19.000	
D	5.615	2	1.000	19.000	
误差	5.62	2			

注: *表示有显著性差异($P<0.05$)。

从表 3 分析可知, A、B、C3 个因素对党参多糖的提取率影响的权重为提取次数>提取时间>料液比。即提取次数对党参多糖提取率影响最大, 其次为提取时间, 最后为料液比。从表 4 的方差分析结果可以看出, 提取次数对党参多糖提取率的影响具有显著的统计学意义($P<0.05$)。正交试验优化得到的最佳提取工艺为 $A_3B_2C_2D_2$, 即在提取 3 次, 提取时间为 1.5 h, 料液比为 1:8 的条件下, 能达到党参多糖的最大提取率。经验证, 在此提取条件下, 党参多糖提取量的平均值为 51.87 mg/g, 与正交表中的最大提取量 52.11 mg/g 相差不多。

3.3 BP 神经网络模型的建立与检验

本试验采用提取次数、提取时间、料液比和空白列作为输入变量, 变量的选取范围见表 1, 采用 3 层 BP 神经网络建模, 输入层传递函数为 *tansig*, 输出层传递函数为 *purelin*, 训练函数为 *traingdx*^[15]。因为确定了输入层及输出层节点数、传递函数和层数等, 为了获得更好的网络性能, 依次选择 1~20 个隐层节点数, 经过 2000 次训练后, 得到了相应的误差平方和, 结果如图 1 所示。结果显示当隐层节点数为 12 时, 误差平方和最小。因此选择 12 个隐层节点数对 BP 神经网络进行 2 万次的训练, 训练目标为均方误差达到 10^{-5} , 学习步长为 0.05。网络经过 19994 次迭代后均方误差达到目标值, 网络性能稳定, 训练曲线见图 2。

为了检验模型的准确性, 将检验样本的自变量输入模型, 经过模型预测出预测值后, 比较预测值与实际值之间的差别, 结果见图 3。从图 3 可知, 预测值与实际值之间的误差很小, 正负误差在 1.0% 以内, 吻合度很高。表明所建立的 BP 神经网络模型稳定且泛化能力强, 可以用来估算不同提取条件下的党参多糖提取率。

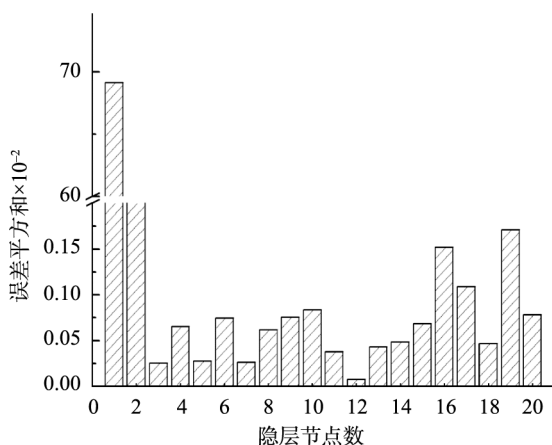


图 1 不同隐层节点数下的 BP 神经网络的误差平方和
Fig.1 Error square sum of BP neural network with different numbers of hidden layer nodes

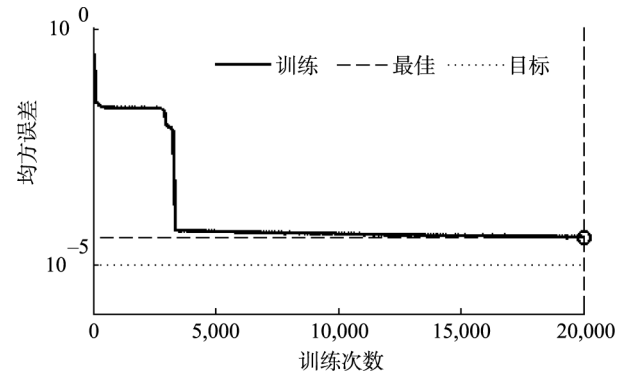


图 2 BP 网络训练曲线
Fig.2 Training curve of BP network

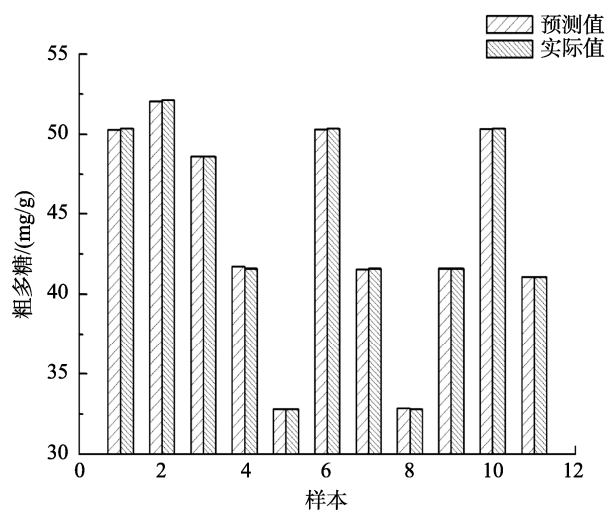


图 3 检验样本预测值与实际值的比较
Fig.3 Comparison of predictions and actual values of test sample

3.4 BP 神经网络结合遗传算法优化党参多糖的提取条件

BP 神经网络建模成功后, 利用 BP 网络捕获的输入与输出之间的映射关系, 作为遗传算法的适应度函数^[15]。再根据训练数据样本确定各个提取条件的变化范围, 在搜索空间内随机选择 20 种组合为初始群, 经过不断的选择、交叉、变异, 直至达到终止条件。最终, 遗传算法给出的最佳提取条件为: 提取次数为 2.986, 提取时间为 1.999 h, 料液比为 1:9.6(m:m)时, 获得党参多糖的最大提取预测值为 55.29 mg/g。为了便于实际操作, 将提取条件修约为提取次数 3 次, 提取时间 2 h, 料液比为 1:10。与正交结果相比(料液比为 1:8, 提取时间为 1.5 h, 提取 3 次), 提取次数 3 次没有变化, 提取时间由原来的 1.5 h 增加到优化后的 2 h, 料液比由原来的 1:8 增加到优化后的 1:10。

经验证, 在遗传算法给出的最佳条件下, 党参粗多糖的提取率为 54.68 mg/g, 与预测值的相对误差为 1.10%, 比正交实验所得的最佳提取条件下的党参多糖提取率

(51.87 mg/g)提高了 5.42%。因此, BP 神经网络模型可以很好的模拟提取因素与党参多糖提取率的关系, 与遗传算法相结合, 无需复杂的实验过程, 就可以快速得到最优的提取条件, 且所得的党参多糖提取率较正交实验最优工艺所得的党参多糖提取率高出 5.42%。

4 结 论

本研究采用拓扑结构为 4-12-1 的 BP 神经网络建立了影响党参多糖的提取因素与党参多糖提取率的网络模型, 很好地模拟了它们之间的关系, 与遗传算法相结合, 快速的得到了党参多糖的最优提取工艺, 得到的最优提取工艺为提取次数 3 次, 提取时间 2 h, 料液比为 1:10(m:m), 在此条件下党参多糖得率预测值为 55.29 mg/g。经验证, 党参多糖提取率较正交试验最优工艺所得的党参多糖提取率高出 5.42%。因此, BP 神经网络结合遗传算法用于党参多糖提取工艺的优化, 是一种更高效、快速、简便的新方法, 可以为党参的进一步研究开发提供新的思路。

参考文献

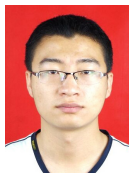
- [1] 刘美霞, 戚进, 余伯阳. 党参药理作用研究进展[J]. 海峡药学, 2018, 30(11): 36-38.
Liu MX, Qi J, Yu BY. Progress in research on pharmacological action of *Codonopsis* [J]. Strait Pharm, 2018, 30(11): 36-38.
- [2] 孙政华, 邵晶, 郭政. 党参化学成分及药理作用研究进展[J]. 安徽农业科学, 2015, 43(33): 174-176.
Sun ZH, Shao J, Guo M. Research progress on chemical constituents and pharmacological action of *Codonopsis* [J]. J Anhui Agric Sci, 2015, 43(33): 174-176.
- [3] 李启艳, 胡德福, 张雪梅, 等. 党参多糖提取纯化工艺优化及其组成研究[J]. 中草药, 2016, 47(15): 2663-2667.
Li QY, Hu DF, Zhang XM, et al. Study on optimization of extraction and purification process of *Codonopsis* polysaccharides and its composition [J]. Chin Tradit Herbal Drug, 2016, 47(15): 2663-2667.
- [4] 范济民, 蒋小丽, 赵志焕. 山西党参多糖提取工艺的优化[J]. 化学与生物工程, 2012, 29(9): 47-50.
Fan JM, Jiang XL, Zhao ZH. Optimization of extraction technology of polysaccharide from *Codonopsis pilosula* L [J]. Chem Biol Eng, 2012, 29(9): 47-50.
- [5] Wang Y, Lu CJ, Zuo CP. Coal mine safety production forewarning based on improved BP neural network [J]. Int J Mining Sci Technol, 2015, (25): 319-324.
- [6] Hu FY, Wang L, Wang SS, et al. A Human body posture recognition algorithm based on BP neural network for wireless body area networks [J]. China Commun, 2016, (8): 198-208.
- [7] Andre E, Brett K, Roberto A, et al. Dermatologist-level classification of skin cancer with deep neural networks [J]. Nature, 2017, (542): 115-118.
- [8] Martowibowo SY, Kaswadi A, Engineering MA, et al. Optimization and simulation of plastic injection process using genetic algorithm and moldflow [J]. Chin J Mach Eng, 2017, (30): 398-406.
- [9] Zhong PS, Li TT, Liu M, et al. Robot path planning method based on improved genetic algorithm and morphin algorithm [J]. Mach Hydraul, 2019, 24(47): 33-38.
- [10] 李珏, 董媛, 李雅雅, 等. 神经网络-遗传算法优化细脚拟青霉多糖超声提取工艺[J]. 吉林医药学院学报, 2019, 40(1): 15-17.
Li J, Dong Y, Li YY, et al. Neural network-genetic algorithm to optimize the ultrasonic extraction process of polysaccharides from *Penicillium spinellum* [J]. J Jilin Med Univ, 2019, 40(1): 15-17.
- [11] 王正明, 洪胜, 罗建平. 人工神经网络-遗传算法优化密花石斛多糖超声辅助提取工艺[J]. 安徽农业科学, 2013, 41(13): 5710-5712, 5715.
Wang ZM, Hong S, Luo JP. Optimization of ultrasonic assisted extraction of *Dendrobium densiflorum* polysaccharide by artificial neural network and genetic algorithm [J]. J Anhui Agric Sci, 2013, 41(13): 5710-5712, 5715.
- [12] 朱俊访, 李博, 聂阳. 多元非线性回归与 BP 神经网络在香菇多糖提取工艺研究中的应用[J]. 海峡药学, 2014, 26(2): 15-17.
Zhu JF, Li B, Nie Y. Application of multiple nonlinear regression and BP neural network in the extraction process of *Lentinus edodes* polysaccharide [J]. Strait Pharm, 2014, 26(2): 15-17.
- [13] 宋金军, 陈冰, 周静, 等. R 语言与正交试验对黄芪中黄芪多糖提取工艺的优化比较[J]. 中国中医急症, 2019, 28(7): 1129-1132.
Song JJ, Chen B, Zhou J, et al. Optimization of extraction process of *Astragalus* polysaccharide from *Astragalus membranaceus* by R language and orthogonal experiment [J]. J Emerg Tradit Chin Med, 2019, 28(7): 1129-1132.
- [14] 韩伟, 孙晓海, 罗文峰. 合欢皮多糖提取工艺优化及 BP 神经网络模型[J]. 南京工业大学学报(自然科学版), 2013, 35(5): 57-61.
Han W, Sun XH, Luo WF. Optimization of extraction process and BP neural network model of *Cortex albiziae* polysaccharide [J]. J Nanjing Univ Technol (Nat Sci Edit), 2013, 35(5): 57-61.
- [15] 王强, 冯玲然, 余晓斌. 基于 BP 神经网络和遗传算法优化番茄红素发酵培养基[J]. 食品与生物技术学报, 2019, 38(2): 111-119.
Wang Q, Feng LR, Yu XB. Optimization of lycopene fermentation medium based on BP neural network and genetic algorithm [J]. Food Sci Biotechnol, 2019, 38(2): 111-119.

(责任编辑: 李磅礴)

作者简介



张 丽, 主要研究方向为保健食品及功能食品研发。
E-mail: 314382566@qq.com



李 凯, 主要研究方向为保健食品开发与应用。
E-mail: 308218025@qq.com