

人工神经网络分类鉴别苦丁茶红外光谱

庞涛涛, 姚建斌, 杜黎明*

山西师范大学分析测试中心, 山西 临汾 041004

摘要 为了分类鉴别苦丁茶, 采用竞争神经网络 (CNN) 和反向传播人工神经网络 (BP 网络) 两种模式的人工神经网络 (ANN) 分别分析了各种苦丁茶的红外谱图。作者采用 25 个样本作训练集, 11 个样本作检验集, 用两种网络进行了训练。结果表明, CNN 网络和 BP 网络均能够有效地实现苦丁茶产地的鉴别, 但 CNN 网络能够进一步地区分苦丁茶的级别。实验表明, CNN 速度快, 预测结果准确, 可望用竞争神经网络 (CNN) 和红外光谱法结合分类鉴别苦丁茶。

关键词 苦丁茶; 人工神经网络; 竞争神经网络; 反向传播人工神经网络; 红外光谱

中图分类号: O657.3 **文献标识码:** A **文章编号:** 1000-0593(2007)07-1336-04

引言

人工神经网络是在生物学中神经网络理论上发展起来的多学科交叉、共同发展的前沿学科。它的研究, 使诸如生物学、认知科学、非线性科学等学科与计算机、电子学、人工智能、微电子、信息处理、模式识别等工程学科有机地结合起来, 从而具有广泛的应用前景。

红外光谱近年来, 在许多学科领域得到广泛应用^[1-9]。人工神经网络在红外光谱测定中的应用已有较多文献报道^[10-15], 运用人工神经网络对药物制剂也进行了研究^[16-19], 本文首次应用竞争神经网络 (CNN) 和反向传播人工神经网络 (BP 网络) 对苦丁茶进行了分类鉴别。结果表明, 无损药材苦丁茶红外光谱的重复性好, 从光谱中提取的用于两种人工神经网络计算的特征数据规律明显, 无监督学习的竞争神经网络 (CNN) 的分类能力要比有监督学习的 BP 网络的分类能力强, 此方法具有简便、快速、准确等特点。实验结果表明采用竞争神经网络技术可分类鉴别不同产地和不同级别的苦丁茶。

1 基本理论

BP 网络是一种具有三层或三层以上的多层神经元网络, 它的左右各层之间的各个神经元可以实现连结, 即左层的每一神经元与右层的每个神经元都有连接, 而上下层各神经元之间无连接。本文采用三层 BP 网络, 该网络由一个输入层,

一个中间层, 一个输出层组成。

BP 网络的传递函数是处处可微的, 经常使用的有 S 型的对数式的正切传递函数和线性函数, 本文采用对数 S (Sigmoid) 型传递函数: $\log \text{sig}(x) = \frac{1}{1 + e^{-x}}$, S 型函数有非线性放大系数功能, 它可以把输入从负无穷大到正无穷大的信号变换为 -1 到 1 之间输出, 对较大的输入信号, 放大系数较小, 而对较小的输入信号, 放大系数则较大, 所以采用 S 型传递函数可以处理和逼近非线性的输入/输出关系, 输出采用线性函数, 输出层节点 i 的输出值 a_i 和期望值 t_i 之间进行比较, 其误差函数为 $e = \frac{1}{2} \sum (t_i - a_i)^2$ 。

竞争神经网络的显著特点是它的输出神经元相互竞争以确定胜者, 胜者指出哪一种原型模式最能代表输入模式。竞争神经网络一般是以无监督学习 (Unsupervised learning) 或称为自组织学习 (Self-organized learning) 方式进行学习的。自组织学习算法的目的是在无导师指导下寻找输入数据空间中的重要模式特征。自组织学习的网络结构模型, 比有监督学习的网络结构模型更接近生物神经系统, 因为网络的自组织过程正是大脑组织的一个基本现象。

竞争型神经网络有很多具体形式和不同的学习算法, 本文只介绍一种比较简单的网络结构和学习算法。网络结构如图 1 所示。

竞争型网络可分为输入层和竞争层。假定输入层由 N 个神经元构成, 竞争层有 M 个神经元。网络的连接权值为

收稿日期: 2006-05-10, 修订日期: 2006-08-20

基金项目: 山西省自然科学基金项目 (2007011027) 资助

作者简介: 庞涛涛, 1981 年生, 山西师范大学化学与材料科学学院硕士研究生 * 通讯联系人 e-mail: lmd@dns.sxnu.edu.cn

w_{ij} , $i = 1, 2, \dots, N$, $j = 1, 2, \dots, M$, 且满足约束条件 $\sum_{i=1}^N w_{ij} = 1$ 。在竞争层中, 神经元之间相互竞争, 最终只有一个或者几个神经元获胜, 以应当前的输入样本。竞争胜利 的神经元就代表着当前输入样本的分类模式。

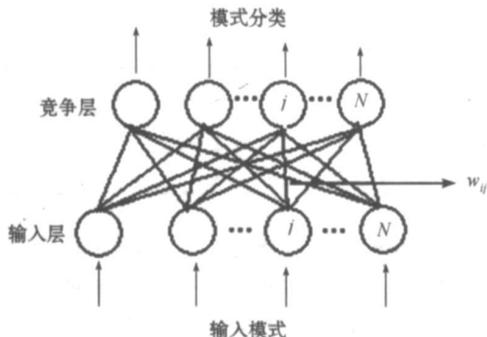


Fig 1 The structure of competitive neural

2 实验部分

2.1 仪器设备和参数设置

FTIR 410 型傅里叶变换红外光谱仪(美国 NICOLET 公司), DTGS 检测器, 光谱范围为 $4\ 000 \sim 400\ \text{cm}^{-1}$, 分辨率为 $4\ \text{cm}^{-1}$, 扫描累加次数为 32 次, 扫描时应扣除水和二氧化碳的干扰。

2.2 样品来源和制备

样品来源于海南、四川、广东、广西、云南、浙江 6 个产地共 36 种苦丁茶。将苦丁茶粉碎, 过 100 目筛, 取其筛后粉末少量, 与 KBr 粉末混合研磨均匀后压片, 直接放入傅里叶变换红外光谱仪中测定。

2.3 神经网络系统的设计依据

采用 Mathworks 公司推出的一套高性能可视化数值计算软件 MATLAB, 基于 MATLAB6.5 神经网络工具箱, 设计出了苦丁茶的神经网络鉴别系统。

2.4 光谱吸收波段的选择

11 种不同产地和不同级别的苦丁茶样本的红外光谱如图 2 所示。从 36 种苦丁茶中抽取 11 种作为预测集, 其余作为训练集。选择红外光谱的 $1\ 750 \sim 1\ 255\ \text{cm}^{-1}$ 波段作为多组分分析的波数范围, 这一范围的选择是为了去除大部分二氧化碳的光谱带以及水的光谱带的贡献, 对每一个模拟样本, 在 $1\ 750 \sim 1\ 255\ \text{cm}^{-1}$ 波段内, 从 $1\ 750\ \text{cm}^{-1}$ 开始, 每隔 $5\ \text{cm}^{-1}$ 取一个点, 共取 100 个点的吸光度, 作为人工神经网络的输入值。输出值为不同种类苦丁茶编码, 如广东苦丁茶编码为 010000; 此代码在训练集中为目标值, 在预测集中为样品种类的编码。

3 结果与讨论

3.1 BP 神经网络对苦丁茶红外光谱的研究

3.1.1 隐含层的节点数

从理论上讲, 隐含层的节点数不大于输入信号的个数, 若隐含层节点数过多, 则网络更易于区分各样本之间的微小差别, 但同时网络的复杂程度增加, 收敛速度减慢。在本研究中, 通过将隐含层节点数分别设定为从 20 到 30 之间的值, 通过比较各不同节点数时, 神经网络的预测误差来确定的。图 3 表示当取不同节点数时, 所对应的苦丁茶的标准预测误差(SEP%)。从图 3 中可以看出, 当隐含层节点数为 21 时, 神经网络的标准预测误差为最小。故在本研究中, 取隐含层节点数为 21。

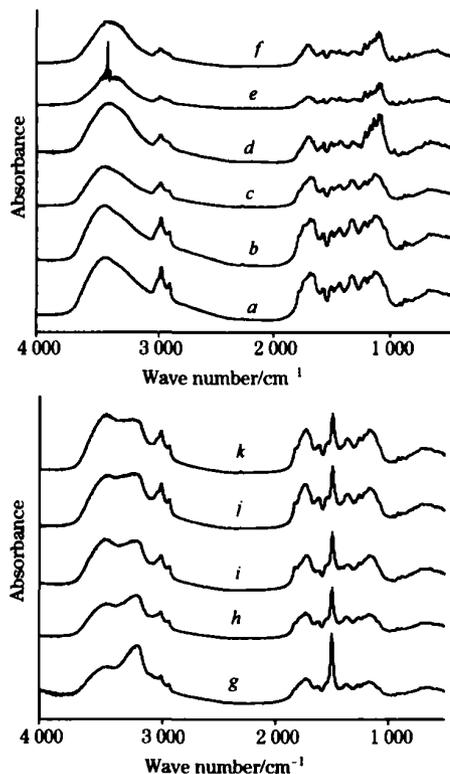


Fig 2 FTIR spectra of Ilex Kudingcha

a: 海南(特级); b: 海南(一级); c: 海南(二级); d: 四川(特级); e: 四川(一级); f: 四川(二级); g: 浙江; h: 云南; i: 广东; j: 广西(特级); k: 广西(一级)

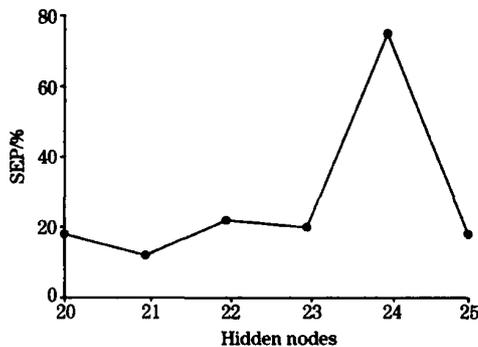


Fig 3 Relationship between the number of hidden nodes and SEP% of Ilex Kudingcha

3.1.2 BP 神经网络对苦丁茶的预测结果讨论

BP 神经网络对苦丁茶的预测结果如表 1 所示。

Table 1 BP Neural networks for the identification of *Ilex kudingcha*

Samples	Predicted result											
1	0.022	1.0	0.014	9.0	0.000	3.0	0.007	1.0	0.011	7.1	0.027	0.0
2	0.006	1.0	0.040	6.0	0.007	0.0	0.001	1.0	0.007	0.1	0.024	7.0
3	0.027	6.0	0.007	9.0	0.011	2.0	0.003	7.0	0.003	5.0	0.996	7.0
4	0.001	6.0	0.007	8.0	0.001	8.0	0.001	1.0	0.997	1.0	0.009	8.0
5	0.009	5.0	0.022	6.0	0.017	0.0	0.002	1.1	0.014	5.0	0.020	2.0
6	0.018	8.0	0.068	9.0	0.006	6.0	0.004	5.0	0.995	2.0	0.038	3.0
7	0.002	7.0	0.061	5.0	0.005	7.0	0.991	9.0	0.007	2.0	0.042	7.0
8	0.002	5.0	0.042	0.990	9.0	0.002	5.0	0.007	7.0	0.032	8.0	8.0
9	0.025	2.0	0.857	3.0	0.014	4.0	0.009	7.0	0.018	5.0	0.074	7.0
10	1.022	0.0	0.088	6.0	0.005	0.0	0.004	5.0	0.000	4.0	0.042	3.0
11	0.969	6.0	0.033	3.0	0.011	8.0	0.001	8.0	0.011	4.0	0.012	4.0

从表 1 可以看出: 样本 1, 2, 3 是同一类, 为海南苦丁茶, 4, 5, 6 是同一类, 为四川苦丁茶, 7 是浙江苦丁茶, 8 是云南苦丁茶, 9 是广东苦丁茶, 10, 11 是广西苦丁茶。此分类结果和实际情况相吻合。说明应用 BP 神经网络对苦丁茶产地和级别进行预测, 产地结果比较满意, 但并不能很好的对级别进行预测。

3.2 竞争神经网络对苦丁茶红外光谱的研究

竞争神经网络对 11 种样本的预测结果如下表 2 所示。

对结果进行分析可得, 当训练步数为 10 时, 样本序号为 1, 2, 3 的分为一类, 为海南苦丁茶; 样本序号为 4, 5, 6 的分为一类, 为四川苦丁茶; 样本序号为 7 的为一类, 为浙江苦丁茶; 样本序号为 8 的为一类, 为云南苦丁茶; 样本序号

为 9 的为一类, 为广东苦丁茶; 样本序号为 10, 11 的分为一类, 为广西苦丁茶。由此可见, 网络已经对样本进行了初步的分类, 这种分类虽然准确但不够精确, 即对苦丁茶的产地预测结果满意, 但对其级别还是不能够很好的预测。

Table 2 Competitive neural network for the identification of 11 samples

Training step	Predicted result										
10	3	3	3	7	7	7	5	1	2	6	6
100	3	7	3	8	4	4	5	1	2	6	6
1000	2	9	10	5	3	6	8	7	1	4	11

当训练步数为 100 时, 样本 2 和 4 分别从海南苦丁茶和四川苦丁茶中划分出来, 是由于生产工艺和地域差别导致的级别不同而产生的结果, 则可知分类结果进一步的细化了。

当训练步数为 1000 时, 每一个样本都被划分为一类, 这和实际情况是吻合的, 即在产地分类的基础上, 对其不同的级别也进行了很好的预测。此时如果再提高训练步数, 已经没有实际意义了。

4 结 论

本研究成功地将人工神经网络应用到苦丁茶分类体系中, 人工神经网络作为一种新兴的化学计量学方法, 显示了其他经典方法无可比拟的特点。采用两种不同模式的人工神经网络对苦丁茶进行了分类, 从预测集分类的结果来看, 无监督学习的竞争神经网络(CNN)的分类能力要比有监督学习的 BP 网络的分类能力强, 该法具有简便、快速、准确等特点, 仍不失为对多种样本同时分类的一种首选方法。

参 考 文 献

- [1] WANG Zhao, SUN Su-qin, LI Xiao-bo, et al(王 钊, 孙素琴, 李晓波, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2001, 21(3): 311.
- [2] GUO Ping, XIONG Ping, YUAN Ya-li(郭 萍, 熊 平, 袁亚莉). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2002, 22(4): 603.
- [3] LIU Guo-lin, CAI Jin-na, LI Wei, et al(刘国林, 蔡金娜, 李 伟, 等). Computers and Applied Chemistry(计算机与应用化学), 2000, 17(2): 109.
- [4] CHENG Cun-gui, RUAN Yong-ming, LI Bing-lan(程存归, 阮永明, 李冰岚). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(11): 1355.
- [5] PENG Yong, SUN Su-qin, ZHAO Zhong-zhen, et al(彭 勇, 孙素琴, 赵中振, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(6): 679.
- [6] Samanidou V F, Hapeshi E A, Papadoyannis I N. J. Chromatogr B., 2003, 788(1): 147.
- [7] Masaaki Kai, Hiromi Kinoshita, Kazuko Ohta, et al. J. Pharm. Biomed. Anal., 2003, 30(6): 1765.
- [8] Kenneth K W To, Yee-Ping Ho, Steve C F, et al. J. Chromatogr A., 2002, 947(2): 319.
- [9] ZOU Hua-bin, YUAN Jia-rong, YUAN Hao(邹华彬, 袁久荣, 袁 浩). Chinese Traditional Patent Medicine(中成药), 2003, 25(4): 261.
- [10] XU Yong-qun, HUANG Hao, ZHOU Qun, et al(徐永群, 黄 昊, 周 群, 等). Chinese J. Anal. Chem.(分析化学), 2003, 31(1): 5.
- [11] LIN Sheng-ling, XU Shao-fen, XIE Chun-sheng, et al(林生岭, 徐绍芬, 谢春生, 等). Chinese J. Anal. Chem.(分析化学), 2004, 32(11): 1421.
- [12] GUO Ye, GOU Ya-hui, TANG Zhen, et al(郭 晔, 苟玉慧, 汤 真, 等). Chinese J. Anal. Chem.(分析化学), 2001, 29(1): 121.

- [13] YANG Nan-lin, CHENG Yi-yu, ZHAI Hai-bin(杨南林, 程翼宇, 翟海斌). Chinese J. Anal. Chem. (分析化学), 2003, 31(6): 664.
- [14] XU Yong-qun, SUN Su-qin, ZHOU Qun, et al(徐永群, 孙素琴, 周群, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2002, 22(6): 945.
- [15] XU Yong-qun, SUN Su-qing, FENG Xue-feng, et al(徐永群, 孙素琴, 冯学峰, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2003, 23(3): 502.
- [16] Lo Shi-Chung B, Heang-Ping Chan, Jyh-Shyan Lin, et al. Neural Networks, 1995, 8(7-8): 1201.
- [17] Ohlsson Mattias. Artificial Intelligence in Medicine, 2004, 30(1): 49.
- [18] Abbas Hussein A. Artificial Intelligence in Medicine, 2002, 25(3): 265.
- [19] Zur R M, Jiang Y, Metz C E. International Congress Series, 2004, 1268(6): 886.

Artificial Neural Networks for the Identification of Infrared Spectra of Ilex Kudingcha

PANG Tao-tao, YAO Jian-bin, DU Li-ming*

Center of Analysis and Test, Shanxi Normal University, Linfen 041004, China

Abstract In order to identify Ilex Kudingcha, two kinds of models of artificial neural networks (ANN), i. e. competitive neural network and back propagation neural network, were used to analyze their infrared spectra. Ilex Kudingcha samples were collected by Fourier transform infrared (FTIR) spectra. Twenty five samples were gathered as a train set, and 11 samples as a test set, then their training was performed using two networks each. The results show that the identification of Ilex Kudingcha from different areas can be effectively performed with the competitive neural network and BP network, but the competitive neural network is used in the identification of different grades of Ilex Kudingcha. The results were better in training speed and accuracy with the competitive neural network. In conclusion, the competitive neural network combined with FTIR spectroscopy is a good method for the identification of Ilex Kudingcha.

Keywords Ilex Kudingcha; Artificial neural networks (ANN); Competitive neural network (CNN); Back propagation neural network; Infrared spectra

(Received May 10, 2006; accepted Aug. 20, 2006)

* Corresponding author