

文章编号: 1005-8893 (2005) 03-0016-03

基于近红外光谱的汽油辛烷值神经网络模型研究^{*}

李为民, 邬国英, 林西平

(江苏工业学院 化学工程系, 江苏 常州 213016)

摘要: 根据汽油组分的辛烷值与近红外光谱分析数据, 用人工神经网络 (ANN) 的反向传播算法建立了汽油组分辛烷值神经网络预测模型。检验表明, ANN 方法能准确地关联近红外光谱分析数据与汽油组分辛烷值的关系。马达法辛烷值预测平均误差为 0.20; 研究法辛烷值预测平均误差为 0.18。

关键词: 人工神经网络; 汽油; 辛烷值; 近红外光谱

中图分类号: TE 626.24; TP 18

文献标识码: A

汽油辛烷值通常用国际标准的 ASTM—CER 发动机来测量。该方法的缺点是其机器结构复杂, 操作和维护的费用高, 因而检测成本高。消耗样品量达数百毫升, 对于实验室来说尤为不便, 因为实验室的微反装置生产出的产物量往往不够满足马达法与研究法测定所需的用量。Kelly 等^[1]提出一种新的辛烷值在线快速检测光谱分析方法——近红外 (NIR) 光谱法, 文献 [2, 3] 对此也进行了研究。汽油的近红外光谱对应于烃类分子中 C—H 键伸展振动的倍频吸收, 由于光谱结构复杂, 谱图重叠严重, 不能用特征峰等简单分析方法, 需对整个谱区的光谱数据进行多元计算和统计分析, 误差较大^[2]。本研究采用金陵石化分公司炼油厂汽油组分近红外光谱吸收系数与辛烷值数据, 建立汽油辛烷值神经网络模型, 辛烷值的预测精度较高, 取得了满意的结果。

1 人工神经网络

所用的神经网络为前向型人工神经网络, 它由输入层、输出层及隐含层互连而成的 3 层网络, 其拓扑结构图见图 1。

神经网络的学习过程采用正向传播和反向传播法 (BP)^[4, 5]。用广义 Delta 规则算法来训练神经网络。

广义 Delta 规则是一种使平方误差最小的迭代梯度下降方法, 这种算法采用了一种称为动量的方法来加速训练, 动量是一种加到已调整的权重因子上的额外权重, 通过加速权重因子的变化提高训练速度。正向传播时, 将训练样本集中的任一样本置

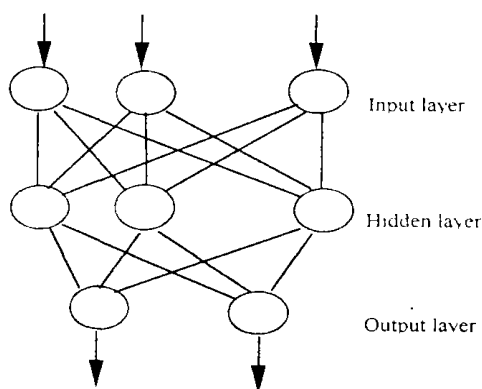


图 1 神经网络拓扑结构图

Fig 1 The topological structure of ANN

于网络的输入层, 经隐含层处理, 在输出层得到该样本对应的输出值。反向传播时, 则是根据输出值与期望值的误差调整网络中的各个权值, 以使其输出值与期望值误差在允许范围, 整个学习过程按正向传播和反向传播循环反复进行, 直至网络收敛。BP 的整个学习过程如下: ①送入一特定的输入,

* 收稿日期: 2005-06-29

基金资助: 江苏省高新技术研究 (农业) 项目 (BG2005325)

作者简介: 李为民 (1962—), 男, 浙江湖州人, 副教授, 从事化学工程与工艺的科研与教学工作。

测定其实际输出; ②将实际输出与期望输出进行比较, 计算误差; ③通过反复调整节点间的连接权重, 使误差(或均方根误差)达到最小。

BP 的学习过程(即求解权重的过程)实质上就是非线性函数的极值求解过程, 因而可以用求非线性函数总体极值的方法求解, 从而得到网络的全局最优解。

2 结果与讨论

2.1 神经网络拓扑结构的确定

神经网络的拓扑结构见图 1。选择汽油的近红外吸收系数 $11\ 400\text{ cm}^{-1}$ 、 $10\ 940\text{ cm}^{-1}$ 、 $10\ 760\text{ cm}^{-1}$ 作为神经网络的 3 个输入节点, 输出节点为研究法辛烷值与马达法辛烷值 2 个节点。

隐含节点数的选择是一个十分复杂的问题, 本研究主要通过检验来确定, 即由预测结果的好坏来确定隐含节点数目。利用已知辛烷值的 21 个汽油样本, 对输入、输出数据即近红外吸收系数、辛烷值作归一化处理, 采用交叉预测法对隐含节点的拓扑结构进行测试, 21 个样本的平均预测误差结果见表 1。

表 1 不同的拓扑结构隐含节点检验结果

Table 1 Effect of the number of hidden dots on the prediction results

Topological structure of ANN	Iteration	Trained error	Prediction error
3-9-2	4 200	0.01	0.010 8
3-8-2	3 700	0.01	0.011 0
3-7-2	2 200	0.01	0.010 4
3-6-2	2 900	0.01	0.011 6
3-5-2	2 800	0.01	0.012 5

Note: learning rate η is 0.2, and momentum factor α is 0.4.

根据上表的结果, 3-7-2 的拓扑结构, 其预测的平均误差较小, 而且迭代次数也不多, 而 3-8-2 及 3-9-2 的预测平均误差虽也较小, 但训练次数增加, 学习时间也相应变长, 因此选用 3-7-2 拓扑结构, 即 3 个输入节点、7 个隐含层节点, 2 个输出节点。

2.2 人工神经网络过拟合问题

神经网络的预报能力是神经网络优劣的主要因素, 用表 2 的 21 个实验样本进行交叉检验预测。学习速率对迭代次数影响较大, 根据经验与预研究取 0.2, 动量因子取 0.4。得到学习平均误差及预测平均误差与训练次数的关系见图 2。可以看出, 随着训练次数的增加, 学习误差不断下降, 但当训

练次数达到 2 000 次后, 尽管学习误差还在下降, 但预测误差变化不大, 当训练次数超过 4 000 次后, 如果继续训练下去, 预测误差开始变大, 所得的模型用于预测未知样本时, 其结果的可靠性下降, 即所谓的过拟合现象。因此, 在神经网络训练中, 以预测误差最小作为训练结束的依据, 本研究的训练次数为 2 000 次。

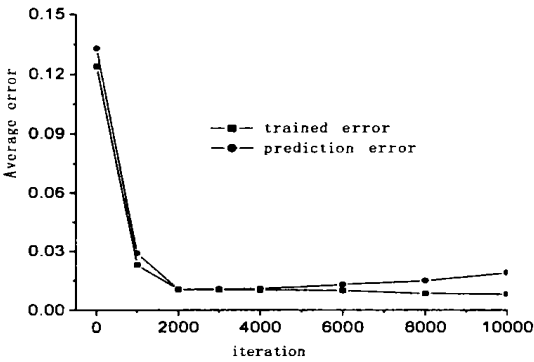


图 2 误差与迭代关系图

Fig. 2 Effect of iteration on the average error

2.3 交叉预测结果

根据所建立的汽油辛烷值神经网络模型, 由金陵石化分公司炼油厂直馏、催化、重整、焦化汽油组分及调合汽油等 21 个汽油样品的近红外吸收系数、汽油辛烷值数据, 得到交叉预测结果见表 2。

由表 2 预测结果可知, 人工神经网络法预测辛烷值的最大误差马达法为 0.37、研究法为 0.39, 平均误差马达法为 0.20、研究法为 0.18。而文献 [2] 的多元计算和统计分析辛烷值预估偏差 $\leq \pm 1.0$ 。基于近红外光谱的神经网络法不管是最大误差还是平均误差都有所降低, 预测精度得到了提高。

3 小 结

(1) 基于近红外光谱的神经网络方法预测汽油组分的辛烷值是可行的, 该神经网络模型合适的拓扑结构为 3-7-2, 学习速率 0.2, 动量因子 0.4。

(2) 神经网络模型预测汽油组分辛烷值最大误差马达法为 0.37、研究法为 0.39, 平均误差马达法为 0.20、研究法为 0.18。

表 2 辛烷值预报结果
Table 2 The prediction results of octane numbers

No.	Near infrared spectrum/ cm^{-1}			Motor octane number			Research octane number		
	11 400	10 940	10 760	Experiment data	Prediction results	Absolute error	Experiment data	Prediction results	Absolute error
1	0.075	0.205	0.118	81.7	81.46	0.24	95.0	95.17	-0.17
2	0.076	0.200	0.108	82.6	82.48	0.12	96.4	96.65	-0.25
3	0.079	0.200	0.106	83.3	83.58	-0.28	97.6	97.44	0.16
4	0.079	0.194	0.099	84.0	84.37	-0.37	97.6	97.49	0.11
5	0.084	0.191	0.096	85.0	85.01	-0.01	99.1	99.14	-0.04
6	0.073	0.204	0.115	81.8	81.95	-0.15	95.9	96.14	-0.24
7	0.070	0.201	0.110	82.1	82.24	-0.14	96.3	96.14	0.16
8	0.069	0.199	0.109	82.5	82.44	0.06	97.2	97.16	0.04
9	0.061	0.230	0.111	85.1	85.36	-0.26	97.4	97.40	0
10	0.051	0.229	0.106	85.9	85.81	0.09	98.0	98.14	-0.14
11	0.061	0.217	0.100	86.7	86.49	0.21	99.0	99.21	-0.21
12	0.058	0.233	0.105	86.0	86.28	-0.28	97.8	97.69	0.11
13	0.062	0.220	0.097	87.4	87.32	0.08	99.8	99.78	0.02
14	0.057	0.229	0.115	84.2	83.93	0.27	96.4	96.60	-0.20
15	0.026	0.282	0.213	44.9	45.00	0.10	45.8	45.45	0.35
16	0.026	0.276	0.201	52.4	52.75	-0.35	54.0	54.27	-0.27
17	0.030	0.269	0.189	59.2	59.38	-0.18	61.0	60.81	0.19
18	0.034	0.268	0.180	65.3	64.98	0.32	68.2	67.99	0.21
19	0.038	0.258	0.160	72.9	72.74	0.16	78.0	78.39	-0.39
20	0.044	0.246	0.143	77.0	77.13	-0.13	84.7	84.88	-0.18
21	0.036	0.261	0.167	70.3	70.01	0.29	74.0	74.24	-0.24

参考文献:

[1] Kelly J J, Callis J B, Jinguji T M, et al. Prediction of Gasoline Octane Number from Near-Infrared Spectral Features in the Range 660—1 215 nm [J]. Anal Chem, 1989, 61: 313—317.

[2] 曹动, 谭吉春, 陈哲, 等. 用近红外光谱分析法测定汽油辛烷值 [J]. 光谱学与光谱分析, 1999, 19 (3): 314—317.

[3] 史月华, 陆勇, 徐光明, 等. 主成分回归残差神经网络校正算法用于近红外光谱快速测定汽油辛烷值 [J]. 分析化学, 2001, 29 (1): 87—91.

[4] 李为民, 徐春明, 高国生, 等. 基于神经网络—遗传算法优化制氢工艺水碳比 [J]. 化工进展, 2004, 23 (9): 998—1 000.

[5] 李为民, 邬国英, 单玉华, 等. 汽油组分辛烷值预测改进 [J]. 江苏石油化工学院学报, 1998, 10 (3): 39—42.

Study of the Gasoline Octane Number Model Using ANN Method with near Infrared Spectra Analysis

LI Wei-min, WU Guo-ying, LIN Xi-ping

(Department of Chemical Engineering, Jiangsu Polytechnic University, Changzhou 213016, China)

Abstract: Based on the real world data of the gasoline octane number and the near infrared spectrum, an artificial neural network (ANN) was developed for predicting the gasoline octane number using BP (Back-Propagation) algorithm. It was shown that the ANN model can correlate and predict the gasoline octane number accurately after comparing the prediction results with the actual data. The average prediction error of motor octane number and research octane number were 0.2 and 0.18 respectively.

Key word: artificial neural network; gasoline; octane number; near infrared spectrum