

# 基于神经网络的自动售货机销量预测

## Vending Machine Sales Forecasting Based on Neural Network

钱永渭 余世明 (浙江工业大学信息工程学院, 浙江 杭州 310023)

### 摘要

采用神经网络理论对罐装饮料自动售货机销售额的预测进行了数学建模, 采用BP神经网络实现了罐装饮料自动售货机销售额的预测, 并通过隐层神经元和输入层神经元个数的确定以及训练算法性能对比验证了所设计模型和方法的可行性和有效性, 为罐装饮料自动售货机的管理工作提供了一种新的方法。

关键词: 自动售货机, 神经网络, 销量预测

### Abstract

In this paper, a mathematical model of the vending machine sales forecasting is build based on neural network theory, which achieved vending machine sales forecasting using BP neural network. At the same time, hidden layer neurons and input layer neurons are determined and training algorithm performance is compared to test and verify the feasibility and effectiveness of the model, which gave a new method for the vending machine management.

Keywords: vending machine, neural network, sales forecasting

自动售货机运营过程中工作人员对每台售货机定期处理的周期间隔难以确定, 例如夏天由于天热, 人们需要补充更多水分, 销售量较大, 而冬天的情况刚好相反。因此有必要对罐装饮料自动售货机的销量的规律进行总结和预测, 合理安排工作人员对每台饮料自动售货机操作的工作周期间隔, 避免周期间隔过短造成的不必要的人力浪费和重复工作, 也避免周期间隔过长造成的饮料自动售货机缺少货物。本文选用神经网络模型建立了罐装饮料自动售货机销量的预测模型, 通过神经网络算法、网络结构等方面的研究, 确定了合适的网络结构和算法。

### 1 预测模型建立

基于神经网络的数据预测问题在不同工程背景和应用领域中得到了广泛的关注, 并且取得了大量的算法成果<sup>[1-5]</sup>。本文使用BP神经网络进行校园自动售货机的销售额预测, BP神经网络一般是指反向传播算法的多层前馈型神经网络, 输入的数据从输入层传递至隐含层, 再从隐含层传递至输出层。各层神经元之间存在的连接强度, 用权值表示强度的强弱并反映神经元间连接的紧密度。如图1是本文罐装饮料自动售货机销售额预测神经网络模型, 其中输入层的 $x_1, x_2, x_m$ 分别是预测依据的历史数据, 输出层的 $y_1, y_2, y_n$ 分别是需要预测的对象, 这个BP神经网络建立了m维空间到n维空间的映射关系实现预测的功能。

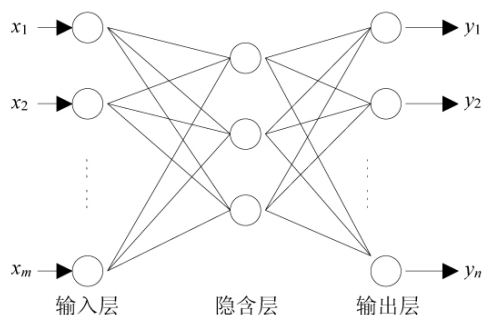


图1 罐装饮料自动售货机销售额预测神经网络模型

### 2 实验数据来源与处理

#### 2.1 实验数据来源

本文的数据主要源于校企合作的罐装饮料自动售货机在学

校的销售情况, 图2为自2009年7月至2011年6月共计104周罐装饮料自动售货机每周销售额的曲线图。

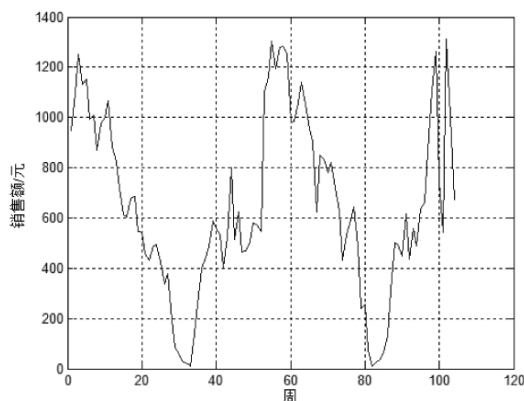


图2 罐装饮料自动售货机周销售额曲线

#### 2.2 数据的归一化处理

在建立神经网络预测模型之前, 首先要将数据进行归一化处理, 使所有数据的范围统一, 保证在 $[0, 1]$ 的区间内, 以适应神经网络中的S函数。

若使用 $x, y$ 分别表示转换前和转换后的数值, 归一化方法的计算公式如下:

$$y = (x - \text{MinValue}) / (\text{MaxValue} - \text{MinValue}) \quad (1)$$

MaxValue、MinValue 分别为样本的最大值和最小值。

#### 2.3 预测结果检验

为了验证算法的有效性, 仅从图中难以直观得到定量结果, 因此选择全参考的客观评价测度来比较不同方法的性能。设 $x = \{x_1, x_2, \dots, x_N\}$ 表示实际的销售额数值,  $y = \{y_1, y_2, \dots, y_N\}$ 表示神经网络输出的预测的销售额,  $N$ 表示周的个数。每个数值对应的预测误差 $e_i$ 计算公式为:

$$e_i = x_i - y_i, i = 1, 2, \dots, N \quad (2)$$

采用均方误差(mean squared error, MSE)衡量算法优劣定义为:

$$MSE(x, y) = \sqrt{\frac{1}{N} \sum_{i=1}^N e_i^2} \quad (3)$$

公式中均方误差也称为标准误差，定义为各测量值误差的平方和的平均值的平方根。反映了每个神经网络预测数据对实际数据的平均能量失真。MSE 值越小,说明神经网络预测数据和对应的实际数据差异越小。由于 MSE 具有明确的物理意义，而且计算简单,被广泛应用各种统计学领域衡量算法优劣。

3 预测过程分析

3.1 隐层神经元个数确定

在实际的应用中，可以通过实验等途径确定隐含层神经元个数。在开始时设置较少的神经元，一定迭代次数后不能达到收敛误差,则再增加隐含层的神经元个数，一直达到比较合理的隐含层神经元个数为止。图 3 是不同隐含层神经元个数对应的误差收敛曲线。从图中的 3 个隐含层神经元个数收敛不到误差精度要求,4 个以上可以满足误差精度，当隐含层神经元个数在 4 个以上时，误差收敛速度很快。但是如果神经元个数选取太多，计算复杂度升高,预测的精度也会下降。综合考虑上述因素，确定 4 个作为隐含层神经元个数。

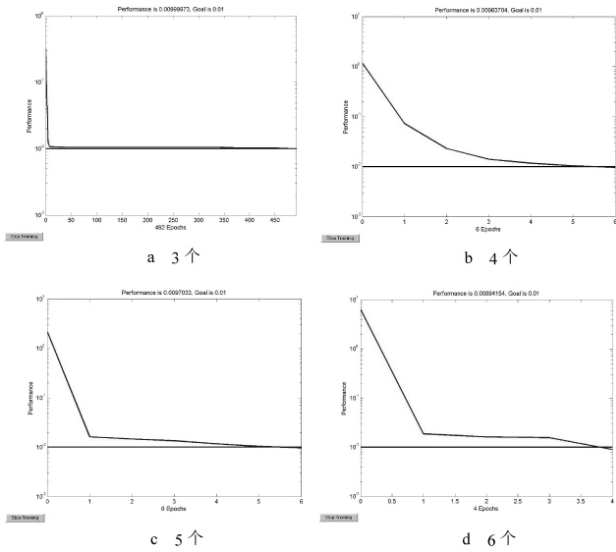


图 3 不同隐层神经元个数对应的误差收敛曲线

3.2 输入层神经元个数确定

神经网络结构设计包括输入层的输入变量的个数。其中输入层和输出层上输入变量和输出变量的个数由具体的实际情况确定。本文的预测模型是采用历史数据对未来数据进行预测，因此输入变量的个数由采用多少个历史数据进行预测来确定。输出变量的个数由需要预测多少个数据来确定。若用周的销售数据预测，输入层神经元数为 X,表示用前 X 周的数据来预测后面 X+1 周的数据,X 取值为 1。前一年的 52 周进行训练，并用于预测后面的周的销售数据。

表 1 是不同周数预测数据进行销售额预测的 MSE,从表中可以看出,由于对训练过程设置了误差精度,通过训练后，训练样本的逼近效果都较好,MSE 数据基本一致，大约在 122~124 之间,但测试样本的 MSE 差异较大，其中从测试样本的预测效果看，当使用前两周的数据预测后一周的数据时 MSE 最小,即预测效果最好,此时对应整体样本的 MSE 也最小,因此说明用

表 1 不同周数预测数据进行销售额预测 MSE

不同周数预测	训练样本	测试样本	整体
1	120.3722	140.2809	128.7521
2	121.8761	69.8904	104.463
3	122.047	165.3266	140.4426
4	119.043	232.6859	171.4598

前两周的数据预测后一周的数据的模型最为合适。

3.3 训练算法比较

神经网络有不同的训练算法，这里比较了典型的经典算法对网络进行训练，包括 Levenberg–Marquardt 反向传播算法、弹性反向传播算法、尺度化共轭梯度下降反向传播算法、BFGS 准牛顿反向传播算法。

表 2 不同训练算法预测数据进行销售额预测的 MSE

不同训练算法预测	训练样本	测试样本	整体
Levenberg–Marquardt 反向传播算法	121.8761	69.8904	104.463
弹性反向传播算法	124.1384	87.8012	111.2009
尺度化共轭梯度下降反向传播算法	123.8284	82.4872	109.3683
BFGS 准牛顿反向传播算法	124.0143	91.3849	112.2537

表 2 通过 MSE 的对比,列出了几种训练算法的精度,显然几种训练算法对训练样本的逼近效果都较好,MSE 数据基本一致,大约在 123~126 之间,但测试样本的 MSE 差异较大，其中 Levenberg–Marquardt 反向传播算法对网络进行训练获得的销售额预测效果的 MSE 最小,预测效果最为理想。Levenberg–Marquardt 反向传播算法的性能随网络规模增大而降低，但本设计中的网络结构并不复杂,因此 Levenberg–Marquardt 反向传播算法较为适合罐装饮料自动售货机销售额的预测，图 4 为 Levenberg–Marquardt 反向传播算法预测效果图。从图中可以看出自动售货机的销售情况呈现出一种季节性的变化规律,整体预测效果较好。但当前后几周数据变化跳动较大的时候,预测效果相对来说有所降低。

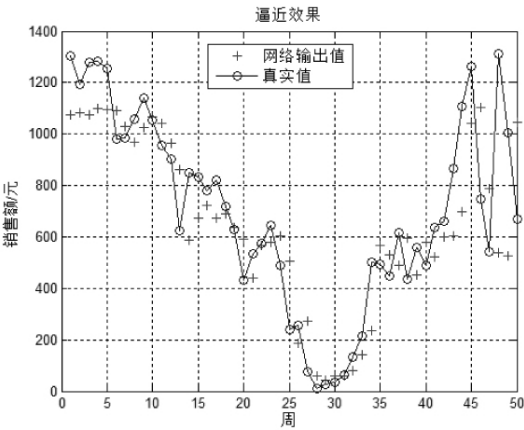


图 4 Levenberg–Marquardt 反向传播算法预测效果

4 结束语

神经网络的数据预测问题由于不同工程应用背景的数据不同,影响数据的因素也完全不同,因此需要对不同的应用情况的预测模型进行具体的分析与设计,并通过实验和优化操作,得到较优的网络结构，本文对基于神经网络的罐装饮料自动售货机销量预测模型进行了整体的设计与验证,通过实验和分析,得到了误差较小的 BP 神经网络结构模型和对应的训练算法，说明运用神经网络预测自动售货机罐装饮料销量的方法是可行的，为罐装饮料自动售货机的管理工作提供了一种新的方法。该方法不仅能对销售总额进行预测，对单一品种的销售情况也能进行相应的预测,由于本文的实验数据来源于实验室研发的样机，受实验条件的限制,在此只是对销售总额进行了预测。

参考文献

[1] 焦李成.神经网络计算[M].西安:西安电子科技大学出版社,1995  
[2] N.M.Pindoriya,S.N.Singh,S.K.Singh.Application of adaptive  
(下转第 44 页)

2 燃料的控制及煤量计算

锅炉主控:某厂的燃料控制(如图 3):锅炉燃料量指令就是锅炉负荷指令经函数运算计算出给煤量与实际主汽压力与给定主汽压力的偏差通过 PID 调节计算出算的给煤量之和(注:未经床温控制校正)。同时锅炉燃料量指令作为燃料主控的给定,而所有输入锅炉的燃煤量测量值的总量与燃油折算煤量之和作为反馈值,燃料主控进行 PID 运算输出,并经平均分配后调整各给煤机的出力,同时给煤机经指令反馈偏差进行自我调节,保证总热量输入满足锅炉负荷的要求。

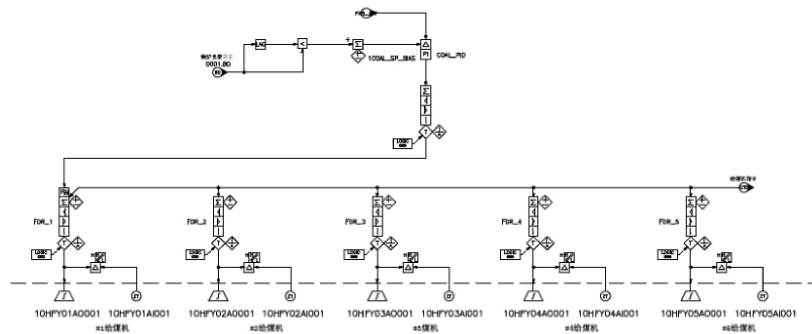


图 3 煤量控制逻辑图

锅炉主控偏置:该厂在控制逻辑中设有锅炉主控偏置,通过锅炉主控偏置设置参数直接改变锅炉燃料指令(该厂偏置设置为给煤量单位为吨),主控偏置良好的优化了因循环流化床锅炉热惯性大而带来的不足,对机组的稳定性起到了很大的作用。

煤量计算:煤量的重量信号是由给煤机上两个悬吊着重物的称重传感器产生的,在称重辊的两边是两根称重跨支承辊,该两辊之间的精确距离给出了一个进行物料称重的皮带长度。每一个称重传感器承担了在称重跨上物料重量的 25%。称重传感器输出的是一个代表物料在皮带上的 T/M 信号,这个重量数据提供给给煤率公式,给煤机根据该公式进行计算:给煤率(T/sec)=重量(T/M)×皮带速度(M/sec),经过微机处理给,给煤机给煤量显示于就地控制面板上。DCS 煤量累计计算是所有给煤机给煤量之和。采用给煤机指令反馈信号,经过换算累计计算出

表 1 给煤机校验记录

#1号炉给煤机校验记录						
名称	砝码重量	首次测量结果	误差	标定后测量结果	误差	允许误差
1号给煤机	200Kg	Kg	%	Kg	%	5‰
2号给煤机	200Kg	198.90Kg	5.55‰	199.60Kg	2‰	5‰
3号给煤机	200Kg	198.88Kg	5.6‰	199.90Kg	0.5‰	5‰
4号给煤机	200Kg	201.00Kg	5‰	199.20Kg	4‰	5‰
5号给煤机	200Kg	198.74Kg	6.3‰	198.88Kg	5.6‰	5‰
6号给煤机	200Kg	198.70Kg	6.5‰	198.80Kg	6‰	5‰
7号给煤机	200Kg	198.80Kg	6‰	199.20Kg	4‰	5‰
8号给煤机	200Kg	200.22Kg	1.1‰	Kg	%	5‰
9号给煤机	200Kg	204.20Kg	21‰	199.23Kg	3.85‰	5‰
10号给煤机	200Kg	198.83Kg	5.85‰	199.89Kg	0.55‰	5‰

总煤量。采用指令的反馈信号优点在于给煤机在启停时对系统都不会造成影响因为只要给煤机停止反馈信号输出就为零而此时指令并不一定为零。(该厂给煤机指令为 0~100,对应给煤率为 0~60 吨)。

给煤机计量精度:该厂近日进行了给煤机的实物测量确定给煤机的精度范围只有 #5、#6 两台给煤机误差超出范围测量结果(如表 1)给煤机标定前后结果,两次标定时间超过 6 个月,所以给煤机的误差不是很大。DCS 总煤量的误差还要加上反馈信号在传输过程中干扰、波动引起的误差。正确的维护和定期的标定给煤机提高煤量计量精度的

同时也能提高锅炉主控效率,减少锅炉主控指令与实际给煤量的偏差,从而优化系统。

3 结束语

通过该循环流化床机组炉内协调控制方式系统(CCS)的基本控制方式、特点及协调控制中煤量控制和煤量计算的分析,加深对控制理论的了解,解决实际问题,同时希望能够对电力爱好者学习 CFB 锅炉的基本控制理论有一定借鉴和帮助。

参考文献

[1]边立秀.热工控制系统[M].北京:中国电力出版社,2002  
[2]李彦,彭钢.循环流化床锅炉炉内控制系统调试及运行技术[M].北京:中国电力出版社,2008  
[3]陈夕松,汪木兰.过程控制系统[M].北京:科学出版社,2005

[收稿日期:2011.8.7]

(上接第 40 页)

wavelet neural network to forecast operating reserve requirements in forward ancillary services market [J].Applied Soft Computing,2011,11(2): 1811-1819  
[3]Wei Shen, Xiaopen Guo, Chao Wu, Desheng Wu.Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm[J].Knowledge-Based

(上接第 42 页)

表 2

		超调量	上升时间	稳态误差	低频振动情况	中高频振动情况
分数阶控制器	600rpm	2.83%	26s	5.67%	相对较小	相对明显
	900rpm	2.83%	25s	8.83%	相对较小	相对明显
PID 控制器	600rpm	0	43s	2.22%	非常明显	相对较小
	900rpm	1.89%	40s	5.56%	非常明显	相对较小

从表 2 可以看出,分数阶控制器比 PID 控制器在抑制电机振动方面的巨大优势。

4 结束语

从上面的分析和讨论很明显的可以看出,分数阶控制器在电机的振动抑制方面与 PID 控制器相比,有着十分明显的优势。

Systems,2011,24(3):378-385

[4]于丽娟,朱长纯,陈晓宁.基于神经网络方法预测薄膜的厚度[J].半导体学报,2000,21(4):400-403  
[5]刘星才,许有鹏,唐郑亮.改进的人工神经网络模型在水文序列预测中的应用研究[J].南京大学学报(自然科学版),2008,44(1):85-90  
[6]郭峰,王斌,刘敏.基于 BP 神经网络的时间序列预测研究[J].价值工程,2010(35):128-129

[收稿日期:2011.8.9]

参考文献

[1]曹军义,梁晋.基于分数阶微积分的模糊分数阶控制器研究[J].西安交通大学学报,2005(11)  
[2]朱呈祥,邹云.分数阶控制研究综述[J].控制与决策,2009(2)  
[3]刘进英,李文.分数阶控制器设计[J].自动控制与检测,2006(11)  
[4]李旺.分数阶系统辨识与控制器设计研究[D].合肥:中国科学技术大学,2010

[5]张弘.分数阶 PID 控制器的研究与仿真[J].西安邮电学院学报,2011(1)  
[6]白晶.分数阶模型的离散化方法研究[D].大连:大连交通大学,2009

[收稿日期:2011.6.20]