

基于 BP 神经网络的药品销售预测模型设计

马新强, 黄 羿

(重庆文理学院 数学与计算机科学系, 重庆 永川 402160)

[摘 要] 提出一种根据药品的历史销售数据来建立数据仓库的方法, 并建立了药品销售预测模型. 该模型利用 BP 神经网络的学习理论建立算法来预测下一年药品的销售量, 为药品生产提出优化方法, 从而形成一个有效的药品生产决策系统模型.

[关键词] 数据仓库; OLAP; 数据挖掘; BP 神经网络; 决策系统

[中图分类号] TP311 [文献标识码] A [文章编号] 1673-8012(2008)02-0064-03

随着市场竞争的加剧, 信息系统的用户已不满足于计算机仅用于日常事务处理, 迫切需要它支持决策制定过程的信息. 在此背景下, 数据仓库应运而生. 数据仓库^[1] (Data Warehouse, DW) 使人们从一个全新的角度认识到信息系统的重要性. 数据不仅可用于检索, 还可以用来分析未来的发展趋势, 并为决策和管理提供依据, 这被称为数据仓库应用的联机分析处理阶段^[2] (On-Line Analytical Processing). 数据挖掘^[3] (Data Mining) 是数据仓库知识发现的一个关键步骤, 其目标是在大量的数据中发现人们有用的信息. 药品决策系统是进、销、存 3 部分的有机组成, 而药品生产系统受销售情况、库存状况、药品需求市场、药品原材料供应等的影响. 采取从药品的历史销售数据角度出发, 运用这些数据构建数据仓库, 在此基础上利用 BP 神经网络的学习理论进行预测, 形成药品生产决

策系统模型.

1 根据历史销售数据构建数据仓库

被称为“数据仓库之父”的 W. H. Inmon 在“Building the Data Warehouse”中对数据仓库的定义是: “数据仓库是一个用以更好地支持企业或组织决策分析处理的、面向主题的、集成的、不可更新的、随时间不断变化的数据集合, 用于支持管理层的决策过程.” 根据这一定义, 构建数据仓库系统不是一个一次性完成的项目, 而是一个随着业务发展的需要平滑地不断丰富、充实的维护过程. 整个数据仓库的构建过程关键在于数据仓库体系结构的构建.

数据仓库的结构框架^[4] 通常可分为数据源、多数据源的集成、数据仓库 (核心)、工具及接口等部分. 本文结合药品销售管理信息系统, 对数据仓库的体系结构进行描述. 如图 1 所示.

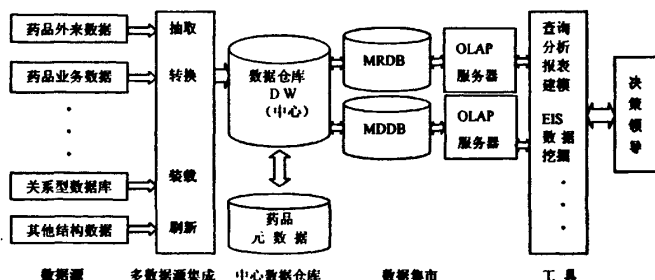


图 1 数据仓库体系结构应用实例

* [收稿日期] 2007-12-11

[基金项目] 重庆文理学院项目资助金资助项目 (Y2007SJ43)

[作者简介] 马新强 (1979 -), 男, 山东鱼台人, 讲师, 硕士, 主要从事数据库技术与人工智能研究. E-mail: mxq345@sohu.com

2 药品销售预测的模型和主要算法

2.1 利用 BP 神经网络建立销售预测模型

BP 神经网络是一个十分有效的智能预测方式.例如:基于遗传算法的 BP 神经网络在企业资信评估的应用^[5]中,提出了一种新的企业资信评估方法等.在本文中,由于各药品的下一年销售数量未知,所以在计划生产之前应该建立药品的历史销售数据仓库,作为预测药品的下一年销售数量的依据.

设某种药品的历史销售数据(共 8 年)向量 $C = (c_1, c_2, c_3, \dots, c_8)$, 其中 c_1 为本年销售量, c_2 为前一年销售量, c_3 为前二年销售量, 以此类推. 把销售数据向量 C 规格化成向量 $SC = (sc_1, sc_2, sc_3, \dots, sc_8)$, 从中构造出 4 个样本向量 X_i 及对应目标输出 Y_j , 具体如表 1 所示.

表 1 样本向量

样本序列	样本向量 X_i	目标输出 Y_j
1	(sc_2, sc_3, sc_4, sc_5)	sc_1
2	(sc_3, sc_4, sc_5, sc_6)	sc_2
3	(sc_4, sc_5, sc_6, sc_7)	sc_3
4	(sc_5, sc_6, sc_7, sc_8)	sc_4

这样就使某年的销售量与其前 4 年的销售量建立了关联,然后将这 4 个样本及对应目标输出不断地输入 BP 神经网络网内进行学习,调整各神经元之间的权值,直至所有样本的输出误差达到下限 ε , 则学习阶段完成.此时,网络的输出 Y_j 与输入 X_i 之间建立了映射函数 $Y = W(X)$, 而且以上 4 个样本与其目标输出满足此种映射,即该网络能够根据历史数据准确预测前 4 年的销售数量,因此它能够预测下一年的销售数量.只要把前 4 年的销售量(sc_1, sc_2, sc_3, sc_4) 作为样本输入,网络的输出中与前 4 年误差最小者即为下一年的预测数量.反向传播网络如图 2 所示.

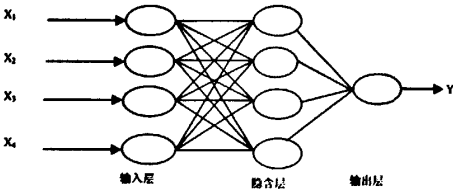


图 2 反向传播网络(BP)

2.2 利用 BP 神经网络建立销售预测模型算法

神经网络主要通过两种学习算法进行训练,即指导式学习算法和非指导式学习算法^[6]. 此

外,还有第 3 种学习算法,即强化学习算法,可把它看作是有师学习的一种特例.有师学习算法需要提供期望或目标输出信号.无师学习算法不需要知道期望输出,在训练过程中,只要向神经网络提供输入模式,神经网络就能够自动地适应连接权,以便相似特征把输入模式分组聚集.

本实例中的反向传播网络^[7]是一种有师学习算法,它采用 4 个输入节点,4 个隐结点和 1 个输出节点,节点的作用函数选用 $f(x) = \frac{1}{1 + e^{-x}}$

型函数,学习算法如下:

第 1 步:选用权值系数初值 W (若权值的选择为本年的最优销售量,这样可以减少权值的修改步骤).重复以下步骤直至每个样本的目标输出 Y_j 和实际输出 Y_j' 满足 $|\frac{y_j - y_j'}{y_j}| < \varepsilon$.

第 2 步:对 j 从 1 到 4,计算每个隐含层节点的输入 I_{jk} 和输出 O_{jk} 以及输出节点的输出 Y_j' (正向计算过程).

第 3 步:根据误差计算各个权值的误差梯度 $\frac{\partial E_j}{\partial W_{ij}}$.修正权值 $W: W_{ij} = W'_{ij} \Delta W_{ij}$, 其中 $\varepsilon = 0.05, W \in (0.2, 0.3)$. 权值的修正采用动量修正法: $\Delta W_{ij}(m) = \eta \sum \delta_{ij} O_{ij} + \alpha \Delta W_{ij}(m-1)$, 即不是每给一个样本后都立即修改权值系数,而是所有样本都施加后,对每个样本效果计算误差.权值系数调节量与误差总和 $\sum \delta_{ij} O_{ij}$ 及上次调节量 $\Delta W_{ij}(m-1)$ 有关(m 为迭代次数; $\eta, \alpha \in (0.1, 0.2)$; 实际取值 $\eta = 0.3, \alpha = 0.4$).

3 根据药品的销售预测来优化生产

生产厂家在从数据仓库中提取有效的历史数据,并结合本年度库存量,运用 BP 神经网络学习理论建立算法模型预测下一年药品销售量的过程中,还应考虑到噪声数量问题.

这里的噪声数量 σ 主要考虑受销售增长率均值影响的数量.增长率均值的求法:首先求下年预测销售量的增长率 $\theta_0 = \frac{c - c_1}{c_1} \times 100\%$, 再求今年销售量的增长率 $\theta_1 = \frac{c_1 - c_2}{c_2} \times 100\%$, 依次求出 $\theta_2, \dots, \theta_8$; 最后的均值增长率 $\theta = \frac{\theta_0 + \theta_1 + \dots + \theta_8}{9} \times 100\%$ ($-1 < \theta < 1$). 下年预测销售量的增长率 θ_0 与均值增长率 θ 之间构成

二维图,且这两个增长率构成一条直线,由其斜率 k 控制噪声数量 σ 的正负. 如图 3 所示.

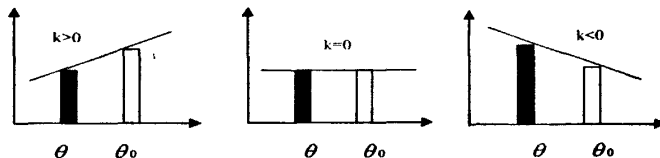


图 3 θ 与 θ_0 构成直线斜率的正负比较情况

根据斜率 k 的正负可判断下年增长率的_{增长趋势}:当 $k > 0$ 时,噪声数量 σ 取正;当 $k = 0$ 时,噪声数量 σ 取零;当 $k < 0$ 时,噪声数量 σ 取负. 因此,有了确定的噪声数量 σ ,下一年的生产量可以得到进一步的优化.

优化的生产量: $Q = C - S \pm \sigma$ (Q :生产量; C :预测的销售量; S :库存量; σ :噪声数量).

在本生产决策方法中,整个生产活动可看作一个多阶段决策过程,对一种药品的生产可看作一个阶段的决策,只要每次(种)药品的生产是最优的,则整个生产活动是最优的. 药品生产决策系统的结构如图 4 所示.

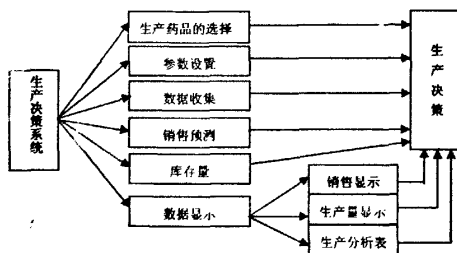


图 4 生产决策系统结构

4 结语

药品生产决策系统的建立是一项复杂的系统工程. 本文采取数据仓库技术收集大量历史数据,并利用数据挖掘技术来分析提取其中的有用

信息,在此基础上结合人工智能的 BP 神经网络对药品的销售量进行预测,并在有了较为精确的销售量和库存量后,再考虑到噪声参数因素,提出了一种优化的生产决策系统方法. 在本销售预测模型中,考虑到的是“年销量”的预测,同样还可以进一步做到“月销量”的预测. 但在生产决策模型与算法的细化工作方面以及更多的噪声变量因素上,还有待进一步研究和分析.

[参考文献]

- [1] W H Inmon. Building the Data Warehouse[M]. 王志海,林友芳,译. 北京:机械工业出版社,2003.
- [2] 谷岩,冯华. 利用数据仓库技术解决异构数据库的集成问题[J]. 计算机应用与软件,2005,22(6):24-26.
- [3] Jiawei han, Micheline Kamber. Data Mining Concepts and Techniques[M]. 范明,孟小峰,译. 北京:机械工业出版社,2007.
- [4] 施伯乐,朱扬勇. 数据库与智能数据分析[M]. 上海:复旦大学出版社,2005.
- [5] 孟凡超,张洪伟,徐剑. 基于遗传算法的 BP 神经网络在企业资信评估中的应用[J]. 计算机应用研究,2007,24(8):301-305.
- [6] 蔡自兴,徐光佑. 人工智能及其应用[M]. 北京:清华大学出版社,2004.

The Design of Medicine Sale Forecast Model Based on BP Neural Network

MA Xin-qiang, HUANG Yi

(Dept. of Math & Computer Science, Chongqing University of Arts and Sciences, Yongchuan Chongqing 402160, China)

Abstract: This artical proposed an approach to build the data warehouse according to the historical data of the drugs sales, created a model to predict the medicine sales volume of next year on the basis of the theory of BP neural network and offered one optimized production method. So as to form an efficient medicine production decision system model.

Key words: data warehouse; OLAP; data mining; BP neural network; decision system

作者：[马新强](#), [黄羿](#), [MA Xin-qiang](#), [HUANG Yi](#)
作者单位：[重庆文理学院, 数学与计算机科学系, 重庆, 永川, 402160](#)
刊名：[重庆文理学院学报（自然科学版）](#)
英文刊名：[JOURNAL OF CHONGQING UNIVERSITY OF ARTS AND SCIENCES](#)
年, 卷(期)：2008, 27 (2)
被引用次数：4次

参考文献(6条)

1. [W H Inmon](#); [王志海](#); [林友芳](#) [Building the Data Warehouse](#) 2003
2. [谷岩](#); [冯华](#) [利用数据仓库技术解决异构数据库的集成问题](#) [期刊论文]-[计算机应用与软件](#) 2005 (06)
3. [Jiawei han](#); [Micheline Kamber](#); [范明](#); [孟小峰](#) [Data Mining Concepts and Techniques](#) 2007
4. [施伯乐](#); [朱扬勇](#) [数据库与智能数据分析](#) 2005
5. [孟凡超](#); [张洪伟](#); [徐剑](#) [基于遗传算法的BP神经网络在企业资信评估中的应用](#) [期刊论文]-[计算机应用研究](#) 2007 (08)
6. [蔡自兴](#); [徐光佑](#) [人工智能及其应用](#) 2004

本文读者也读过(10条)

1. [凤德伟](#), [张忠能](#), [凌君逸](#) [基于神经网络的短期销售预测](#) [期刊论文]-[计算机工程](#) 2004, 30 (z1)
2. [李雅莉](#), [LI Ya Li](#) [基于数据挖掘的销售预测研究](#) [期刊论文]-[微型机与应用](#) 2009, 28 (8)
3. [查振中](#), [王晖](#) [灰色模型GM\(1, 1\) 结合Excel实现药品销售预测](#) [期刊论文]-[中国医院管理](#) 2004, 24 (5)
4. [王宪庆](#), [涂冰](#), [文诗琪](#), [WANG Xianqing](#), [TU Bing](#), [WEN Shiqi](#) [BP神经网络在药品销售预测中的应用](#) [期刊论文]-[中国医药技术经济与管理](#) 2009, 3 (5)
5. [李雪娟](#), [蒋世忠](#), [黄展鹏](#) [基于LM算法的BP网络在药品销售预测中的应用](#) [期刊论文]-[福建电脑](#) 2007 (10)
6. [娄青](#) [销售预测审计量表开发和应用研究](#) [学位论文] 2005
7. [张志宇](#) [数据挖掘技术在食品经销管理决策中的应用研究](#) [学位论文] 2004
8. [耿东升](#), [张文斌](#) [某药材站药品销售行为分析](#) [会议论文]-2001
9. [吴正佳](#), [王文](#), [周进](#), [Wu Zheng-jia](#), [Wang Wen](#), [Zhou Jin](#) [BP神经网络在备货型企业销售预测中的应用](#) [期刊论文]-[工业工程](#) 2010, 13 (1)
10. [李儒勋](#), [张洪伟](#) [基于BP的CRM系统销售预测的数据挖掘](#) [期刊论文]-[计算机应用](#) 2004, 24 (11)

引证文献(5条)

1. [王宪庆](#), [涂冰](#), [文诗琪](#) [BP神经网络在药品销售预测中的应用](#) [期刊论文]-[中国医药技术经济与管理](#) 2009 (5)
2. [马健](#), [盛魁](#) [基于遗传BP神经网络组合模型的中药销售预测研究](#) [期刊论文]-[河北北方学院学报（自然科学版）](#) 2013 (4)
3. [甘永祥](#), [张淑兰](#) [数据挖掘在自动化药房中的应用](#) [期刊论文]-[中国医院药学杂志](#) 2013 (19)
4. [舒展](#) [基于BP神经网络算法的三区电加热炉温度微机控制系统](#) [期刊论文]-[九江职业技术学院学报](#) 2011 (2)
5. [赵晓娇](#), [董绍华](#), [赵宁](#) [基于人工神经网络的钢管销售预测](#) [期刊论文]-[物流技术](#) 2010 (10)

引用本文格式：[马新强](#), [黄羿](#), [MA Xin-qiang](#), [HUANG Yi](#) [基于BP神经网络的药品销售预测模型设计](#) [期刊论文]-[重庆文理学院学报（自然科学版）](#) 2008 (2)