

基于智能算法的价格预测模型探究

孙吉红^{1,2}, 张丽莲^{2,3}, 武尔维^{2,3}, 郜鲁涛^{2,3}, 彭琳^{2,3}, 钱晔^{2,3}

(1. 云南农业大学 园林园艺学院, 云南 昆明 650201;

2. 云南农业大学 云南省高校农业信息技术重点实验室, 云南 昆明 650201;

3. 云南农业大学 基础与信息工程学院, 云南 昆明 650201)

摘要:文中主要是建立基于智能算法的价格预测模型,解决商品价格预测的难题。首先通过德尔菲技术进行专家意见征询,确定影响商品价格的因子;然后数据化处理影响因子后,利用拉依达准则剔除异常数据;再以影响因子为输入量,采用BP神经网络算法建立商品预测模型;采用聚类分析算法建立预测模型与之对比。以玫瑰鲜切花为例建立价格预测模型,实验结果表明:该商品价格预测模型规避了单纯BP神经网络算法的缺陷,具有预测商品价格的普遍性、实用性。

关键词:BP神经网络;聚类分析;预测;价格

中图分类号:TP31

文献标识码:A

文章编号:1673-629X(2014)11-0107-03

doi:10.3969/j.issn.1673-629X.2014.11.027

Research on Price Prediction Model Based on Intelligent Algorithm

SUN Ji-hong^{1,2}, ZHANG Li-lian^{2,3}, WU Er-wei^{2,3}, GAO Lu-tao^{2,3},

PENG Lin^{2,3}, QIAN Ye^{2,3}

(1. School of Landscape Gardening, Yunnan Agricultural University, Kunming 650201, China;

2. University Key Laboratory of Agricultural Information Technology in Yunnan, Yunnan
Agricultural University, Kunming 650201, China;

3. School of Basis and Information Engineering, Yunnan Agricultural University,
Kunming 650201, China)

Abstract: In order to estimate commodity prices, the purpose in this paper is to establish the forecasting price model. Firstly, through expert consulting by the Delphi technique, the factors influencing commodity prices are determined. Secondly, after processing the factors, data are got, and abnormal data of them are eliminated according to Pauta criterion. Thirdly, with the influencing factors as input data, establish the forecasting commodity prices model by means of the BP neural network algorithm. Finally, this algorithm is compared with the clustering analysis algorithm applied in the forecasting model. This forecasting model has been applied to predict the rose cut flowers. The experimental results show that this forecasting model can avoid defects of traditional BP algorithm, and has certain generality and practicality.

Key words: BP neural network; clustering analysis; forecasting; price

0 引言

预测是对于未来事情的一种考量,在生产和技术上都具有重要的意义^[1]。而价格预测是预测在实际生活中的重要部分,是人们生活不可或缺的部分。价格预测方法发展至今已经形成了150余种,主要包括定性预测方法和定量预测方法。定量预测方法相比定性预测方法更具有科学性、客观性。定量预测方法包括

了传统预测方法和智能预测方法。传统预测方法主要是根据专家的丰富经验进行预测,具有一定的主观性和局限性。智能预测方法能够突破传统方法的局限性,处理较为复杂的问题,处理不完全信息的预测问题,是预测方法的一种新的趋势。人工神经网络作为智能算法中的一种,是模拟生物神经网络进行信息处理的一种数学模型。它是一种人工智能范畴内的方

收稿日期:2013-12-11

修回日期:2014-03-14

网络出版时间:2014-09-11

基金项目:“十二五”国家科技支撑计划课题(2012BAD35B04);国家自然科学基金资助项目(31260292);云南省教育科学研究基金项目(2012y367,2013J103);云南省自然基金项目(2008ZC050M)

作者简介:孙吉红(1983-),男,云南建水人,硕士研究生,CCF会员,研究方向为计算机应用、人工智能。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140911.0942.009.html>

法,通过对大脑的生理过程进行模拟以实现一些特定的功能,具有很强的非线性映射能力^[2]。用于非线性预测的效果很好^[3]。与传统的定性预测方法相比较,神经网络预测方法在建模等方面更简便,效果更好。同时,对于商品价格预测问题中包含的多个影响因子之间,所产生一系列复杂问题的拟合能力更强、预测精确度更高^[4-5]。Leal I 等运用时间序列、神经网络等方法进行商品价格的预测^[6],Hu 等提出小波神经网络预测方法,对水产品的价格进行预测^[7],都取得了良好的预测效果。而 BP 神经网络是一种多层前馈神经网络,该网路的主要特点是信号前向传递,误差反向传播^[8]。它可以实现从输入到输出的任意非线性映射^[9]。文中尝试用改进后的 BP 人工神经网络算法构建价格预测模型。

1 价格预测模型的构建

文中采用拉依达准则、BP 人工神经网络,构建商品价格预测模型,采用聚类分析算法进行实验对比,具体如下。

1.1 利用拉依达准则对数据进行检验,剔除异常数据

(1) 补充数据。由于收集到的价格、产量、销售量数据存在着少量的数据空缺,结合空缺数据前后几日的数据、不同年份但日期相同的数据进行数据填补。

(2) 使用 premnmx 函数,将数据进行归一化处理,对预测商品的价格、产量、销售量等数据进行归一化处理,使数据一致化,并且将各指标的值得控制在 $(-1, 1)$ 的范围内。避免输入、输出数据数量级差别较大造成网络预测误差较大的现象。

$$P = \text{premnmx}(p)$$

$$2 * (m - \min) / (\max - \min) - 1$$

(3) 采用拉依达准则剔除异常数据。公式如下^[10]

$$v_b = |x_b - y| > 3\delta, 1 \leq b \leq n$$

将 x_b 剔除。在实验过程中,保证实验数据 $n > 10$ 。

1.2 确定价格影响因子

(1) 针对预测的某种商品,查阅文献之后,初步确定影响该商品价格的因子。

(2) 把查阅文献之后的资料,整理成专家咨询意见表,采用德尔菲技术进行专家意见征询,最终确定影响该商品价格的因子。

(3) 查询文献,找到 2 篇以上的核心期刊与专家意见相对应,最终确定影响该商品价格的因子。

1.3 确定权重

权重是综合评价中的一个重要的指标体系,合理地分配权重是量化评估的关键^[11]。因此,权重的构成是否合理,直接影响到评估的科学性^[11]。在此,通过回收调查问卷的打分来确定权重,每个权重反映了它

们对商品价格的影响程度,确定权重计算公式^[12],即准则层的权重计算公式如下:

$$\text{准则层权重} = \frac{\text{准则 } q \text{ 分数} + \text{准则 } q \text{ 分数总分}}{\text{总和}}$$

其中,选择/非常小计 1 分;/挺小计 2 分;/一般计 3 分;/挺大计 4 分;/非常大计 5 分。

1.4 构建商品价格预测的智能模型

1.4.1 构建神经网络模型,对商品价格进行预测

BP 神经网络是典型的多层前馈型网络,由输入层、隐含层和输出层组成,层与层之间一般采用全部连接方式,同一层单元之间不存在相互连接^[13]。当采用 3 层 BP 神经网络构建价格预测模型时,将取得较好的预测效果。模型包括了输入层、竞争层、输出层 3 个部分。其中,综合考虑专家意见和多篇论文的观点之后,将 m 个数据化后影响因子作为输入层,输入层包括了 m 个神经元;输出层包括 1 个神经元,为实际价格。隐含神经元数为 7 个。本模型已经进行了数据补充、归一化处理,并通过拉依达准则对数据进行检验,剔除异常数据。得到一组个数相对较少的、彼此不相关的新输入变量^[14],再用确定后的影响因子作为 BP 神经网络的输入,提高了神经网络的预测性能。

1.4.2 构建聚类分析模型,对商品价格进行预测

聚类分析法也可以对商品的价格进行预测,用分类的方法把商品的价格分为 5 个(昂贵、较贵、正常、较便宜、非常便宜)不同的类别。训练样本时,打破时间的界限,把相同类别的因子分配到相同的类别中,即生成第 1~5 类目标距离像,把测试数据生成未知类目标距离像。计算目标距离像与这 5 类目标距离像中的哪类距离更近,从而实现识别。

1.5 商品预测价格的评价准则

价格预测效果评价方面的问题,关键还是看其预测的准确性和有效性。因此,为了更好地评价算法预测的性能^[15],文中设计了准确率评价指标。准确率就是预测结果和实际结果相差的比率。假定 A 为模型预测出的结果, B 为实际价格,则准确率等于 $|B - A| / B \times 100\%$ 。

2 价格预测模型的应用及分析

2.1 实验数据和平台

文中以玫瑰鲜切花为例,以影响鲜切花价格因素为输入层数据,建立通用价格预测模型,预测商品价格。由于存在一些数据噪声,这些噪声会影响预测的正确率,所以在预测前,必须把这些噪声数据剔除。

2.2 确定影响玫瑰鲜切花的价格因子

根据秦开大博士毕业论文中提到:昆明国际花卉拍卖市场中,玫瑰的供货量和成交量呈现一致性的变

化趋势,而平均成交价格与供货量、成交量的变化呈相反的趋势,平均成交价格与流拍率的趋势变化关系不明显^[11]。说明供货量、成交量是影响鲜切花价格的重要因素。同时,邓冰在硕士论文中也明确提到:影响鲜切花价格最重要的因素包括:供给量、需求量、生产成本、质量等级、供货商和批发商计划获利的多少、拍卖方面的现场购买人数、现场购买者参与拍卖的策略^[16]。丽都花卉专家认为:供货商和批发商计划获利的多少、生产成本并不能影响花卉价格。吴红芝教授、赵燕副教授等专家认为:拍卖方面的现场购买人数并不能影响鲜切花价格。供给量、需求量、质量等级、生产成本、节假日前后,会影响鲜切花的价格。

在查阅以上文献之后,进行德尔菲技术进行专家意见征询,确定影响鲜切花价格因子:供给量、需求量、质量等级。

2.3 实验结果与分析

文中以 1.4.1 节的模型输入和 1.1 节的算法剔除异常数据建立预测模型,将 2011–2012 年昆明玫瑰鲜切花的拍卖价格、影响因子进行训练,设定最大迭代次数为 1 000 次,学习率为 0.01。利用 2013 年 4–5 月玫瑰鲜切花的价格进行预测,建立一个包括输入层、隐含层、输出层在内的三层网络结构的 BP 神经网络。对玫瑰鲜切花价格进行预测时,经过训练之后,迭代次数达到了 1 000 次,运算时间较短,提高了算法的准确率,得到了较为准确的玫瑰鲜切花的预测数据。

2.4 实验对比与分析

2.4.1 确定权重

由于邓冰已经证明了鲜切花供给量的权重为 0.5,需求量的权重为 0.5,因此,根据权重的比值将玫瑰鲜切花的产量与需求量进行合并,即玫瑰鲜切花的产量 * 0.5 + 需求量 * 0.5,作为聚类分析中的一维数据。然后根据 1.3 节的方法计算玫瑰鲜切花权重。

2.4.2 实验对比

文中把 2011–2012 年玫瑰鲜切花的价格分为 5 个类别,分别是 x_1 :0.8~0.99 元, x_2 :1.0~1.2, x_3 :0.4~0.6 元, x_4 :随机抽取, x_5 :2.0 元以上。抽取 2013 年玫瑰鲜切花数据 20 组进行测试,价格在 1.05~1.25 元范围内。利用聚类分析算法,计算测试数据和 $x_1 \sim x_5$ 数据之间的距离。

通过表 1 的计算结果可以发现, x_2 与预测样本的距离最接近,说明预测样本的金额在 1.0~1.2 元之间,属于鲜切花较贵的时期。

因此,文中提出的聚类分析算法,预测效果较好,精确度较高。但是,聚类分析只能把预测对象进行分类,并不能具体地预测出玫瑰鲜切花的具体价格,在实际应用中,缺乏说服力。

表 1 2011–2012 年建模样本与预测样本之间的距离

x_i	鲜切花	价格范围	测试数据与 x_i 距离
x_1	一般	0.8~0.99 元	6.744 7
x_2	较贵	1.0~1.2 元	5.215 3
x_3	便宜	0.4~0.6 元	8.788 3
x_4	随机抽取		11.674 0
x_5	昂贵	2.0 元以上	8.900 0

表 2 所示为部分月季鲜切花经过 BP 神经网络算法预测的数据展示。其中, R 值越接近 1 表示预测值越准确,充分显示了基于智能算法的鲜切花价格预测模型在短期,最小维度预测中的预测结果。从表中可以清楚地看到月季鲜切花的预测准确率保持在 90% 以上,精确地展现了该智能模型的优势。

表 2 部分月季鲜切花预测数据

月季鲜切花 实际价格	最终误差	鲜切花预 测率/%	R 值
0.72	0.014 9	97.97	0.915 58
0.78	0.014 7	98.12	0.915 58
0.78	0.037 4	95.21	0.915 58
0.76	0.014 8	98.05	0.915 88
0.85	0.006 2	99.27	0.915 88
0.74	0.030 9	95.82	0.915 88
0.82	0.024 4	97.02	0.915 88
0.82	0.011 1	98.65	0.915 88
0.83	0.015 5	98.13	0.915 88

3 结束语

(1)利用拉依达准则、BP 人工神经网络相结合的方法,构建商品价格预测模型。该模型利用了德尔菲技术进行专家意见征询,确定了影响商品的价格因素。

(2)利用拉依达准则剔除异常数据,有效减少了输入因子的个数,提高了输入因子的准确度,提高了 BP 神经网络算法的运算速度,增强了 BP 神经网络算法预测的准确率。

(3)利用聚类分析算法对商品价格进行预测,并与该商品价格预测模型进行对比,更加突出智能预测模型的优势。

(4)文中以玫瑰鲜切花为例,构建的基于智能算法的商品预测模型具有广泛的适用性,适用于所有商品价格预测,可进行推广运用。

参考文献:

[1] 陈振伟,郭拯危.小波神经网络预测模型的仿真实现[J].计算机仿真,2008,25(6):147–150.

5 结束语

文中通过结合二维 Gabor 小波与稀疏保持投影算法,提出了一种新的特征提取方法。该方法的优点在于:用二维 Gabor 小波提取人脸图像特征,克服了人脸图像对光照、姿态、表情变化的影响;由于 Gabor 特征的维数很高,文中又利用了稀疏保持投影算法对 Gabor 特征进行降维,使 Gabor 特征图像在高维空间中的重构关系在低维空间中也能够保留,利于后续的分类识别。ORL 人脸数据库上的实验表明文中方法与 PCA,LDA,SPP 算法相比,有效地提高了人脸的识别率。

参考文献:

[1] Turk M, Pentland A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.

[2] Belhumeur P N, Hespanha J P, Kriegman D. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.

[3] Lades M, Vorbruggen J C, Buhmann J, et al. Distortion invariant object recognition in the dynamic link architecture [J]. IEEE Transactions on Computers, 1993, 42(3): 300-311.

[4] Liu Chengjun, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition [J]. IEEE Transactions on Image Processing, 2002, 11(4): 467-476.

[5] Qiao Lishan, Chen Songcan, Tan Xiaoyang. Sparsity preserving

projections with applications to face recognition [J]. Pattern Recognition, 2010, 43(1): 331-341.

[6] Cai D, He X, Han J. Spectral regression for dimensionality reduction[R]. [s. l.]: UIUC, 2007.

[7] 邵诗强, 施立欣, 周龙沙. 基于环形 Gabor 小波与 CS-LBP 算法在人脸识别中的应用[J]. 光电子技术, 2012, 32(3): 180-184.

[8] The Yale face database B and the extended Yale face database B[EB/OL]. [2011]. <http://cvc.yale.edu/projects/yalefaces>.

[9] Wright J, Ma Yi, Mairal J, et al. Sparse representation for computer vision and pattern recognition [J]. Proceedings of the IEEE, 2010, 98(6): 1031-1044.

[10] Cai Deng, He Xiaofei, Han Jiawei. Spectral regression: a unified approach for sparse subspace learning[C]//Proceedings of international conference on data mining. [s. l.]: [s. n.], 2007.

[11] Murray J, Kreutz-Delgado K. Visual recognition and inference using dynamic overcomplete sparse learning[J]. Neural Computation, 2007, 19(9): 2301-2352.

[12] Baraniuk R. A lecture on compressive sensing[J]. IEEE Signal Process Mag, 2007, 24(4): 118-121.

[13] Donoho D. Compressed sensing[J]. IEEE Trans on Information Theory, 2006, 52(4): 1289-1306.

[14] Donoho D, Tsaig Y. Fast solution of l_1 -norm minimization problems when the solution may be sparse[R]. USA: Stanford University, 2006.

+++++

(上接第 109 页)

[2] 任伟宏. 农产品市场价格预测方法探析[J]. 中国农学通报, 2011, 27(26): 209-212.

[3] 刘 芹, 余一娇, 谭连生. 一种利用 BP 神经网络的 Internet 流量预测算法[C]//中国计算机大会. 出版地不详: 出版者不详, 2003.

[4] Conejo A J, Plazas M A, Espinola R, et al. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models[J]. IEEE Trans on Power System, 2005, 20(2): 1035-1042.

[5] Park D C, El-Sharkawi M A, Marks R J, et al. Electric load forecasting using an artificial neural network[J]. IEEE Trans on Power Systems, 1991, 6(2): 442-449.

[6] Leal I, Melin P. Time series forecasting of tomato prices in Mexico using modular neural networks and processing in parallel[J]. Hybrid Intelligent Systems, 2007, 208: 385-402.

[7] Hu Tao, Zhang Xiaoshuan, Hou Yunxian, et al. A hybrid model for forecasting aquatic products short-term price integrated wavelet neural network with genetic algorithm [J]. LNCS,

2005, 3611: 352-360.

[8] 史 峰, 王小川, 郁 磊, 等. MATLAB 神经网络 30 个案例分析[M]. 北京: 北京航空航天大学出版社, 2010.

[9] 张德丰. MATLAB 神经网络应用设计[M]. 北京: 机械工业出版社, 2009.

[10] 王 岩, 隋思涟. 试验设计与 MATLAB 数据分析[M]. 北京: 清华大学出版社, 2012.

[11] 秦开大. 鲜活农产品拍卖市场价格波动研究[D]. 昆明: 昆明理工大学, 2012.

[12] 刘枚莲, 朱美华, 黄 键. 港口吞吐量预测影响因素筛选方法研究[J]. 水运工程, 2011(3): 76-80.

[13] 彭 琳, 杨林楠. 基于 GRA/ BPNN 的农作物害虫发生量预测模型[J]. 农机化研究, 2013, 35(6): 19-24.

[14] 桂 斌, 黄立冬, 周 杰, 等. 金融时间序列模糊边界预测研究[J]. 小型微型计算机系统, 2012, 33(10): 2283-2286.

[15] 李 婧. 浅谈影响商品价格的具体因素[J]. 新课程, 2012(11): 175-175.

[16] 邓 冰. 鲜活农产品拍卖价格的影响因素研究[D]. 昆明: 昆明理工大学, 2010.

基于智能算法的价格预测模型探究


作者：

孙吉红, 张丽莲, 武尔维, 郜鲁涛, 彭琳, 钱晔, SUN Ji-hong, ZHANG Li-lian, WU Er-wei, GAO Lu-tao, PENG Lin, QIAN Ye

作者单位：

孙吉红, SUN Ji-hong(云南农业大学 园林园艺学院, 云南 昆明650201; 云南农业大学 云南省高校农业信息技术重点实验室, 云南 昆明 650201), 张丽莲, 武尔维, 郜鲁涛, 彭琳, 钱晔, ZHANG Li-lian, WU Er-wei, GAO Lu-tao, PENG Lin, QIAN Ye(云南农业大学 云南省高校农业信息技术重点实验室, 云南 昆明 650201; 云南农业大学 基础与信息工程学院, 云南 昆明 650201)

刊名：

计算机技术与发展 

英文刊名：

Computer Technology and Development

年, 卷(期):

2014(11)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wj fz201411027.aspx