

文章编号:1672-058X(2012)09-0061-04

基于 SVM 和神经网络组合预测模型物流需求预测*

田 丽, 曹安照, 王 蒙, 周明龙, 王 静

(安徽工程大学 电气工程学院, 安徽 芜湖 241000)

摘 要:物流需求的定量数据是区域物流发展政策和规划的重要依据,影响物流需求的因素很多,传统的预测方法无法全面考虑各种因素,预测精度较低。为了提高物流需求预测的精度,采用组合预测的方法,建立一种基于支持向量机和神经网络的组合模型。首先采用支持向量机进行预测得到预测基本数据,然后通过 BP 神经网络进行残差修正,通过算例仿真分析,结果表明组合预测模型具有更高的精度,是一种有效的预测方法,为物流需求预测提供了新的思路。

关键词: 物流需求;支持向量机;神经网络

中图分类号: TP391

文献标志码: A

随着经济技术水平的不断提高,现代物流业也在快速的发展,物流需求增长迅速。物流需求的定量数据是区域物流发展政策和规划的重要依据,有利于行业管理部门制定相关的行业管理法规,对物流企业合理地配置有限的资源有着重要的作用^[1]。因此,对物流需求预测的研究就有着重要的理论意义和现实意义。

目前,很多学者和物流管理人员对物流需求预测问题进行了大量的研究,提出许多预测方法。传统的物流需求预测方法主要有:线性回归、时间序列分析、灰色等。随着人工智能的发展,支持向量机和神经网络等方法也被用于物流需求预测。不同的预测模型包含各自独立的信息,对于假设条件和预测范围有一定的局限性,有时无法全面反映物流需求的变化规律,如果将单个预测模型进行合理的组合,充分利用有用信息,可以提高预测的精度。采用支持向量机和神经网络组合的预测模型。

1 支持向量机算法

支持向量机(Support Vector Machine, SVM)是建立在统计学理论的 VC 维理论和结构风险最小原理基础上的,根据有限的样本信息在模型的复杂性(即对特定训练样本的学习精度)和学习能力(即无错误地识别任意样本的能力)之间寻求最佳折衷,以期获得最好的推广能力。

给定样本数据集 $\{(x_i, y_i)\}_{i=1}^N$, $x_i \in R$ 为输入向量, $y_i \in R$ 为输出向量, N 为训练样本个数,那么支持向量机预测的表达式为: $f(x) = \omega^T \Phi(x) + b$, 其中 ω 为权值向量, b 为偏置量^[2]。

SVM 采用结构风险最小原理,这里的风险是用 Vapnik 的 ε 不敏感损失函数来度量, ε 为:

收稿日期:2012-03-10;修回日期:2012-03-25.

* 基金项目:国家自然科学基金(71171002);安徽省自然科学基金(11040606M24).

作者简介:田丽(1962-),女,安徽芜湖人,教授,从事建模仿真研究.

$$L_\varepsilon = \begin{cases} |f(x) - y| - \varepsilon & |f(x) - y| \geq \varepsilon \\ 0 & |f(x) - y| < \varepsilon \end{cases} \quad (1)$$

为了使支持向量机在训练集上获得泛化推广能力更好的预测模型,不仅需要使经验风险的最小化,同时还要尽量使预测低模型的复杂度更低,鉴于此,支持向量机预测就是一个对下列问题进行优化求解的过程^[2]:

$$\min_{\omega, b, \xi_i, \xi_i^*} = \frac{1}{2} \omega^T \omega + c \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (2)$$

其中, ξ_i, ξ_i^* 表示松弛变量, c 表示惩罚参数,引入拉格朗日乘子, SVM 预测模型为:

$$f(x) = \sum_{i=1}^n (a_i + a_i^*) k(x, x_i) + b \quad (3)$$

其中 $k(x, x_i)$ 为支持向量机核函数,这里采用径向基函数。

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right) \quad (4)$$

2 BP 神经网络算法

BP 算法又称误差方向传播算法 (Error Back-propagation Algorithm), 是神经网络算法中应用最为广泛的一种。BP 神经网络模型是由输入层、输出层和若干隐含层组成的前向连接模型, 同层个神经元之间互不连接, 相邻层的神经元通过权重连接且为全互连结构。其拓扑结构图如图 1 所示^[3]。

当一个学习样本提供给网络后, 神经元的激活值从输入层经各中间层向输出层传播, 在输出层的各神经元获得网络的输入响应。接下来, 按照减少目标输出与实际误差的方向, 从输出层经过各中间层逐层修正各连接权值, 最后回到输入层^[4]。在神经网络的模型中, 核心的变量是各层之间的连接权重。当一个神经网络模型“学习”得到输入和输出变量之间的关系之后, 即可用于对给定输入数据的预测^[3]。

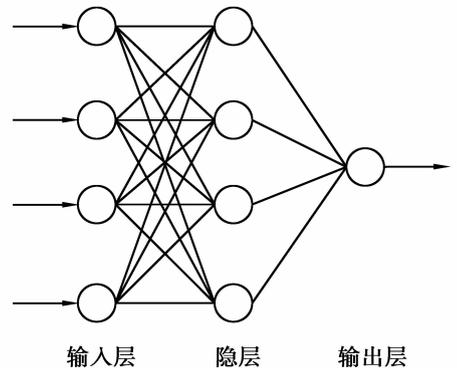


图 1 神经网络结构

3 组合预测模型

组合预测模型的结构图如图 2 所示。它实质上是一种残差修正型组合预测模型, 首先用 SVM 模型对原始样本数据进行预测分析, 得到预测结果, 记为 \hat{y}_i , 原始数据序列和预测结果序列的差为新的序列, 记为 e_i 序列, 用 BP 神经网络模型对 e_i 序列也就是残差序列进行分析, 得到修正后的残差序列, 记为 e'_i , 最后拿 SVM 模型所得到的预测结果与 BP 神经网络模型所得到的修正残差序列相加, 得到最终的预测数据。即, $\hat{y} = \hat{y} + e'_i$ 。

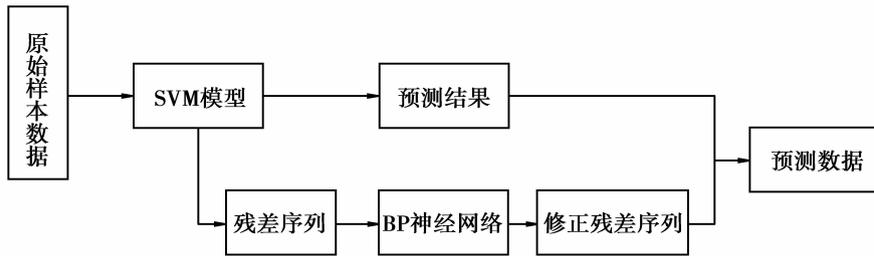


图2 组合预测模型结构

4 算例分析

选取某地区物流需求数据进行仿真分析,检验组合预测模型在物流需求预测中的效果。其中 y 为货运量(万吨); x_1 为地区国民生产总值(GDP)(亿元); x_2 为第一产业产值(亿元); x_3 为第二产业产值(亿元); x_4 为第三产业产值(亿元); x_5 为区域零售总额(亿元); x_6 为区域外贸总额(亿美元); x_7 为人均消费水平(元)。具体数据如表1所示。

表1 物流需求统计数据

指标	x_1	x_2	x_3	x_4	x_5	x_6	x_7	y
1990	623.29	187.23	214.78	121.53	227.41	12.24	713	109.34
1991	713.67	178.49	258.13	178.62	267.19	14.52	768	112.15
1992	867.24	205.01	344.98	240.86	311.23	17.31	879	115.46
1993	1 025.87	246.28	437.35	341.29	421.67	21.29	103.7	117.62
1994	1 590.21	328.46	457.12	446.78	625.43	27.65	128.9	129.07
...
2000	3 546.25	663.87	1 439.66	1 445.34	1 789.22	33.19	2 689	31 448
2001	3 879.21	698.87	1 578.33	1 643.10	1 978.28	35.79	2 967	32 890
2002	4 216.23	708.19	1 708.58	1 798.32	2 139.47	39.86	3 246	33 827
2003	4 757.74	798.21	1 967.56	2 008.43	2 367.14	51.43	3 897	35 183
2004	5 633.23	1 020.20	2 341.91	2 298.55	2 639.18	67.89	4 329	37 569
2005	4 757.74	798.21	1 967.56	2 008.43	2 367.14	51.43	3 897	35 183
2006	5 633.23	1 020.20	2 341.91	2 298.55	2 639.18	67.89	4 329	37 569

首先,采用SVM模型进行预测,为提高SVM的学习速度,对数据进行归一化处理。以1990-2001年的样本数据作为测试集,用于模型的参数优化和建模,2002-2006年的数据作为测试集,用于模型检验。

然后,采用BP神经网络模型,神经网络输入层取7个神经元,隐层取12个,输出层取1个,同样采用表1中的数据进行分析和预测。最后,采用3节提出的组合预测模型进行预测分析。所得到的预测结果见表2和图3。

表 2 预测结果及误差分析

年份	实际值	SVM 预测值	相对误差	BP 预测值	相对误差	组合预测值	相对误差
2002	33 827	33 515	0.92%	33 128	2.06%	33 689	0.41%
2003	35 183	34 827	1.01%	34 581	1.71%	35 057	0.36%
2004	37 569	37 210	0.96%	37 017	1.47%	37 418	0.40%
2005	40 175	40 462	0.71%	40 931	1.88%	40 126	0.12%
2006	43 641	43 274	0.84%	42 968	1.54%	43 573	0.16%

由上面的图表结果可以看出,组合模型的预测曲线与实际曲线的趋势基本一致,从误差分析来看,组合预测模型的精度要比单一预测模型的精度高,组合模型中采用了 BP 神经网络进行残差修正,结果较为理想。在 3 种预测模型中,组合预测模型的预测值更接近实际值,说明组合模型具有更好的性能。

5 结 论

为了提高物流需求预测的精度,建立了基于支持向量机和 BP 神经网络的组合模型,通过算例仿真分析,采用支持向量机和 BP 神经网络以及组合模型分别对样本数据进行预测分析,结果表明:组合预测模型比单一预测模型有更高的精度,是一种物流需求预测分析的有效方法,具有一定的实际意义和推广价值。

参考文献:

- [1] 宿梦思. 基于 BP 神经网络和证据理论的区域物流需求预测研究[D]. 武汉:武汉科技大学,2010
- [2] 吴洁明,李余琪,万励. 物流需求预测算法的仿真研究[J]. 计算机仿真,2011,28(9):246-249
- [3] 叶创鑫,谭满春. 基于 SVM 与人工神经网络组合模型的物流规划车辆行程时间预测[J]. 暨南大学学报,2010,31(5):451-456
- [4] 飞思科技产品研发中心. 神经网络理论与 MATLAB7 实现[M]. 北京:科技出版社,2005
- [5] 杨进伟. 基于 GM/BP 神经网络组合预测模型的区域物流量预测研究[D]. 武汉:中南大学,2011
- [6] XIAO H Y, JOHN L. Production Manufacturing and Logistics Demand forecast sharing in a dual-channel supply chain[J]. European Journal of Operational Research,2006,174:646-667
- [7] 马银军. 基于 BP 神经网络的物流量预测方法研究[J]. 物流科技,2010(5):81-83
- [8] 陈修素,陈睿. 关于供应链管理下批发价格价格依赖性的补偿策略理论[J]. 重庆工商大学学报:自然科学版,2012(1):19-22

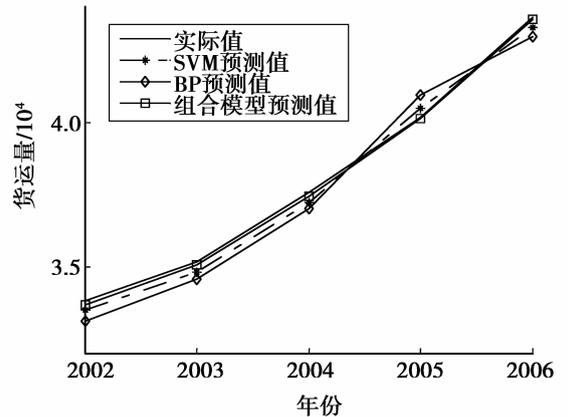


图 3 3 种模型的预测结果