Decision Support Systems 54 (2012) 584–596

Contents lists available at SciVerse ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Sales forecasting for computer wholesalers: A comparison of multivariate adaptive regression splines and artificial neural networks

Chi-Jie Lu [a], Tian-Shyug Lee [b,*], Chia-Mei Lian [c]

[a] Department of Industrial Management, Chien Hsin University of Science and Technology, Taiwan
[b] Department of Business and Administration, Fu Jen Catholic University, Taiwan
[c] Graduate School of Business Administration, Fu Jen Catholic University, Taiwan

ARTICLE INFO

Article history:
Received 27 October 2010
Received in revised form 27 May 2012
Accepted 11 August 2012
Available online 21 August 2012

Keywords:
Sales forecasting
Computer wholesaler
Multivariate adaptive regression splines
Artificial neural networks
IT industry

ABSTRACT

Artificial neural networks (ANNs) have been found to be useful for sales/demand forecasting. However, one of the main shortcomings of ANNs is their inability to identify important forecasting variables. This study uses multivariate adaptive regression splines (MARS), a nonlinear and non-parametric regression methodology, to construct sales forecasting models for computer wholesalers. Through the outstanding variable screening ability of MARS, important sales forecasting variables for computer wholesalers can be obtained to enable them to make better sales management decisions. Two sets of real sales data collected from Taiwanese computer wholesalers are used to evaluate the performance of MARS. The experimental results show that the MARS model outperforms backpropagation neural networks, a support vector machine, a cerebellar model articulation controller neural network, an extreme learning machine, an ARIMA model, a multivariate linear regression model, and four two-stage forecasting schemes across various performance criteria. Moreover, the MARS forecasting results provide useful information about the relationships between the forecasting variables selected and sales amounts through the basis functions, important predictor variables, and the MARS prediction function obtained, and hence they have important implications for the implementation of appropriate sales decisions or strategies.

1. Introduction

In the consumer-centric environment of today's business world, enterprises seeking good sales performance often need to maintain a balance between meeting customer demand and controlling inventory costs. Carrying a larger inventory allows customer demand to be satisfied at all times, but can result in over-stocking, leading to problems such as tied up capital, inventory writedowns, and reduced profit margins. Lower inventory levels, in contrast, may reduce inventory costs, but can result in opportunity costs arising from missed sale opportunities, reduced customer satisfaction, and other problems. Sales forecasting can be used to determine the required inventory level and avoid the problem of under/over-stocking. In addition, sales forecasting can have implications for corporate financial planning, marketing, client management, and other areas of business. Improving the accuracy of sales forecasts has therefore become an important aspect of operating a business.

There is an extensive body of literature on sales forecasting in such industries as textiles and clothing [52–54], fashion [41,47,58], books [7,10,48], and electronics [9,11–13]. However, very few studies center on sales forecasting in the information technology (IT) industry, especially for computer wholesalers. Lu and Wang [34] employed a combination of independent component analysis, growing hierarchical self-organizing maps, and support vector regression analysis to develop a hybrid sales forecasting model for a computer dealer. In the wake of technological advances and rapid changes in consumer demand, IT products have come to be characterized by their variety, constant changes in specifications, and rapid price declines. These factors have made sales forecasting in the IT industry an important but difficult task. This paper focuses on sales forecasting for computer wholesalers in light of the important role they play in the IT industry by distributing IT products to retailers and customers.

Artificial neural network (ANN) algorithms such as backpropagation neural networks (BPN) and support vector regression (SVR) have been found to be useful techniques for sales/demand forecasting due to their ability to capture subtle functional relationships among empirical data, even where the underlying relationships are unknown or hard to describe [9,22,34,63]. Unlike traditional time series forecasting models such as the Box–Jenkins ARIMA model and multivariate regression analysis, ANNs are data-driven and non-parametric. They require no strong model assumptions, and can map any nonlinear function without a priori assumptions about the properties of the data [24,56,65]. Coupled with their superior performance in constructing non-linear models,

* Corresponding author at: Department of Business and Administration, Fu-Jen Catholic University, Hsin-Chuang, Taipei 24205, Taiwan. Tel.: +886 2 2905 2653; fax: +886 2 2905 2186.
E-mail addresses: 036665@mail.fju.edu.tw, mleetw@yahoo.com (T.-S. Lee).

0167-9236/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dss.2012.08.006

ANNs have been successfully applied in sales/demand forecasting [9,22,53].

Kuo and Xue [28] used ANNs to forecast sales for a beverage company. Their results showed that the forecasting ability of ANNs is indeed better than that of ARIMA specifications. Chang and Wang [9] applied a fuzzy BPN to forecast sales for the Taiwanese printed circuit board industry. Hyunchul et al. [27] first used independent component analysis to screen variables before employing an ANN algorithm to predict sales for a Korean shopping mall. They also showed that the proposed forecasting scheme is superior to a forecasting method in which principal component analysis is first used to screen variables before an ANN algorithm is applied. Yang et al. [63] reported that SVR is a promising method for predicting Chinese tobacco sales. Luis and Richard [37] combined ARIMA and ANN models to forecast sales for a Chilean supermarket. Their results showed that this combined forecasting technique can help firms make correct decisions. Sun et al. [47] successfully used an extreme learning machine (ELM) to forecast sales for a fashion retailer. Wu [61] utilized the combination of a wavelet support vector machine and particle swarm optimization to develop a hybrid model for auto sales forecasting. The results indicated that the forecasting ability of the hybrid model is indeed better than that of ARIMA models. Wong and Guo [58] integrated an extreme learning machine with a harmony search algorithm to develop a hybrid sales forecasting model for fashion retail supply chains.

Despite the existence of a significant body of literature on sales forecasting using ANNs, the difficulty of identifying important forecasting variables makes ANNs less attractive for sales predictions, as the selection of important forecasting variables is crucial to the construction of sales forecasting models, given that the variables selected will usually affect the accuracy of the model. Having too many forecasting variables will add complexity to the forecasting model, while having too few may result in an ineffective model. Important forecasting variables that have an impact on sales forecasting results are often the key focus areas or indicators requiring managerial attention. Discussing and understanding these important forecasting variables will lead to improved management and sales efficiency. This study therefore utilizes a methodology that enables both faster processing and the selection of variables – multivariate adaptive regression splines (MARS) – to construct sales forecasting models for computer wholesalers. Through the outstanding variable screening ability of MARS, variables important to sales forecasting for these wholesalers can be obtained to make better sales management decisions.

MARS is a nonlinear and non-parametric regression methodology [23]. It is a flexible procedure that requires no advance specification of a functional form. Rather, it attempts to adapt to the unknown functional form using a series of piecewise regression splines. This makes it very suitable for modeling complex non-linear relationships among variables. Moreover, unlike ANNs, MARS can identify 'important' independent (forecasting) variables and investigate the relationship between the selected independent and dependent (target) variables through the built basis functions when many potential forecasting variables are considered. Finally, MARS does not require a long training process and hence can save lots of time in the model building process. The power of MARS in building prediction models has been demonstrated in many applications such as network intrusion detection [40], electricity price forecasting [2], cancer diagnosis [18,33], software engineering [5,45,66], and credit scoring [16,31]. However, to the best of the authors' knowledge, in no reported study has MARS been used to forecast sales.

This study applies five different methodologies – MARS, BPN, SVR, ELM, and a cerebellar model articulation controller neural network (CMACNN), the latter being an effective neural network model [59] – in sales forecasting for computer wholesalers, and compares the merits of MARS against those of the three other ANN algorithms. To achieve this objective, forecasting models are developed using backpropagation neural networks, a support vector machine, an extreme learning machine, a CMACNN, and four two-stage forecasting schemes. First, MARS, BPN, SVR, and CMACNN are applied to construct sales forecasting models using all forecasting variables. In the case of the two-stage forecasting schemes, MARS is first used as a screening tool for the forecasting variables, after which the important forecasting variables obtained are used as input variables for forecasting models developed using BPN, SVR, ELM, and CMACNN. This results in four two-stage forecasting models respectively called the MARS–BPN, MARS–SVR, MARS–ELM, and MARS–CMACNN models. Finally, to compare the results of the sales forecasting models proposed in this study, experiments are carried out using monthly sales data from two Taiwanese computer wholesalers.

The rest of this paper is organized as follows. Section 2 gives a brief introduction to the MARS, SVR, BPN, and CMACNN methodologies. The experimental results and related discussions are presented in Section 3. Section 4 concludes the paper.

2. Methodology

2.1. Multivariate adaptive regression splines

MARS is a nonlinear and non-parametric regression methodology first proposed by Friedman [23] as a flexible procedure in which model relationships are nearly additive or involve interactions with fewer variables. The MARS modeling procedure is based on a divide-and-conquer strategy in which training data sets are partitioned into separate regions, each of which is assigned its own regression equation. MARS excels at finding not only optimal variable transformations and interactions, but also in identifying the complex data structures often concealed in high-dimensional data. This makes MARS particularly suitable for problems with high input dimensions.

MARS essentially builds flexible models by fitting piecewise linear regressions; that is, the non-linearity of a model is approximated through the use of separate linear regression slopes in distinct intervals of the independent variable space. Therefore, the slope of the regression line is allowed to change from one interval to the other as the two 'knot' points are crossed. The variables to be used and the end points of the intervals for each variable are found through a fast but intensive search procedure. In addition to searching for variables one by one, MARS also searches for interactions between variables, allowing any degree of interaction to be considered as long as the built model provides a better fit with the data.

Fig. 1 depicts a simple example of how MARS would attempt to fit data in a two-dimensional space with piecewise linear regression. Note that y and x in Fig. 1 are the dependent and independent variables, respectively. It can be observed that k_1 and k_2 in Fig. 1 are two knot points delimiting three intervals where different linear relationships are identified.

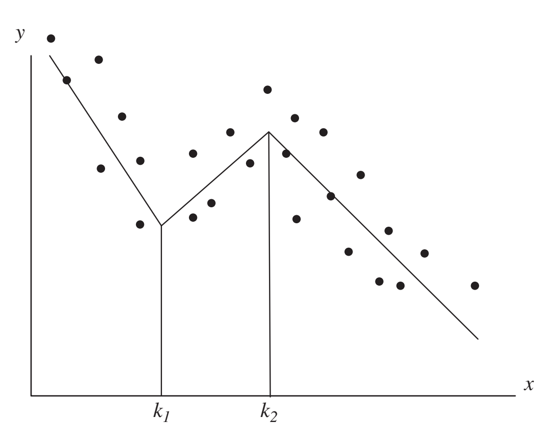

Fig. 1. A simple example of MARS.