Int. J. Production Economics 128 (2010) 603–613

Contents lists available at ScienceDirect

Int. J. Production Economics

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

Combining independent component analysis and growing hierarchical self-organizing maps with support vector regression in product demand forecasting

Chi-Jie Lu, Yen-Wen Wang *

Department of Industrial Engineering and Management, Ching Yun University, 229 Chien-Hsin Road, Chung-Li 32026, Taiwan

ARTICLE INFO

Article history:
Received 31 December 2009
Accepted 5 July 2010
Available online 14 July 2010

Keywords:
Demand forecasting
Support vector regression
Independent component analysis
Growing hierarchical self-organizing maps

ABSTRACT

In the evaluation of supply chain process improvements, the question of how to predict product demand quantity and prepare material flows in order to reduce cycle time has emerged as an important issue, especially in the 3C (computer, communication, and consumer electronic) market. This paper constructs a predicting model to deal with the product demand forecast problem with the aid of a growing hierarchical self-organizing maps and independent component analysis. Independent component analysis method is used to detect and remove the noise of data and further improve the performance of predicting model, then growing hierarchical self-organizing maps is used to classify the data, and after the classification, support vector regression is applied to construct the product demand forecasting model. In the experimental results, the model proposed in this paper can be successfully applied in the forecasting problem.

1. Introduction

Companies excelling in customer's demand know the importance of effective supply chain planning. Successful businesses need an accurate picture of demand to drive production, inventory, distribution, and buying plans across their operations. These challenges are intensified by the effects of seasonality, promotions, and product proliferation, not to mention growth through mergers and acquisitions.

Customer demand forecasting can be implemented to manage a global demand plan with multiple channels to market. An accurate forecast, which, in turn, drives more responsive customer service with lower inventories and reduced obsolescence, especially in the 3C(computer, communication, and consumer electronic) market. According to the statistics provided by Taiwan's Institute for Information Industry, the value of 3C related products was 1.3 trillion NT dollars in 2008 and 1.1 trillion NT dollars in 2009. For this reason, the 3C market plays an important role in Taiwan's semiconductor manufacturing and, as a result, accurate forecast of product demand has become increasingly important. Accordingly, many researchers have begun to pay more attention to the integration of product demand into supply chain management models. Reiner and Fichtinger (2009) has developed an extended dynamic demand forecast and inventory model for a two stage supply chain process, assuming purchase decisions are made by rational actors. In addition, Dolguia and Pashkevich (2008) has proposed a new demand model for multiple slow-moving inventory items with short request histories and unequal demand variance. They both note the importance of demand forecast in the supply chain management model.

With the recent development of artificial intelligence models, several methods have been found to work more effectively than traditional methods when applied to forecast models. However, many researchers (Chang et al., 2006; Wang et al., 2009) have mentioned that no matter what kind of data, some noise may influence the forecast result a lot. It seems data preprocessing to be more and more important. An integrated demand forecast model will be developed in this research, in which noise detecting and removing task will be considered first and then all data will be clustered to increase the accuracy and the practicability of the model. The detailed methodologies introduction and literatures review applied in this research can be found in Section 2. The framework of proposed forecast model will be illustrated in Section 3. Section 4 will present the sales data set and results of our experiment, and Section 5 will discuss the contributions of our model.

2. Methodology and its review

This paper constructs a support vector regression (SVR) predicting model to mitigate the problem of product demand

* Corresponding author. Fax: +886 3 4591933.
E-mail addresses: jerrylu@cyu.edu.tw (C.-J. Lu), ywwang@cyu.edu.tw (Y.-W. Wang).

0925-5273/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpe.2010.07.004

forecast. It is aided by the utilization of growing hierarchical self-organizing maps (GHSOM) and independent component analysis (ICA): hereafter referred to as the ICA–GHSOM–SVR model. The ICA method is used to detect and remove the noise of data and further improve the performance of SVR; the GHSOM is then used to classify the data. After the classification, SVR is applied to construct the product demand forecasting model. The framework of ICA–GHSOM–SVR model is shown in Fig. 1, and the detailed introduction and literature review of each method can be seemed in the following sections.

2.1. Independent component analysis

ICA is a novel statistical signal processing technique designed to find independent sources given only observed data composed of mixtures of unknown sources, without any prior knowledge of the mixing mechanism. In the operation of the basic ICA model, the observed mixture signals $\mathbf{X}$ can be expressed as $\mathbf{X}=\mathbf{AS}$, where $\mathbf{A}$ is an unknown mixing matrix and $\mathbf{S}$ represents the latent source signals that cannot be directly observed from the mixture signals $\mathbf{X}$. The ICA model describes how the observed mixture signals are generated by a process utilizing the mixing matrix $\mathbf{A}$ to linearly mix the latent source signals $\mathbf{S}$. The source signals are assumed to be mutually and statistically independent. Based on this assumption, the ICA solution is obtained in an unsupervised learning process that identifies a de-mixing matrix $\mathbf{W}$. The de-mixing matrix $\mathbf{W}$ is used to transform the observed mixture signals $\mathbf{X}$ to yield the independent signals $\mathbf{Y}$, i.e., $\mathbf{WX}=\mathbf{Y}$. The independent signals $\mathbf{Y}$ are then used as the estimates of the latent source signals $\mathbf{S}$. The rows of $\mathbf{Y}$, called independent components (ICs), are required to be as mutually independent as possible. Even though the basic ICA model has been widely applied in medical signal processing, audio signal processing, face recognition, feature extraction, and quality control, there have still been few applications using ICA in financial time series forecasting.

Back and Weigend (1997) used ICA to ascertain the features of the daily returns of 28 of Japan's largest stocks. The results show that dominant ICs can reveal more underlying structure and information of the stock prices than principal component analysis. Kiviluoto and Oja(1998) employed ICA to find the fundamental factors affecting the cash flow of 40 stores in the same retail chain. They found that the cash flow of the retail stores was mainly affected by holidays, season, and competitors' strategies. Oja et al.(2000) applied ICA in foreign exchange rate time forecasting. They first used ICA to estimate independent components and the mixing matrix from the observed time series data and then filtered the independent components to reduce the effects of noise through linear and nonlinear smoothing techniques. Thereupon the autoregression (AR) model was employed to predict the smoothed independent components. Finally, they combined the predictions of each smoothed IC by using a mixing matrix and thus obtained predictions for the original observed time series. In our study the ICA method is used to detect and remove the noise of data and further improve the performance of SVR.

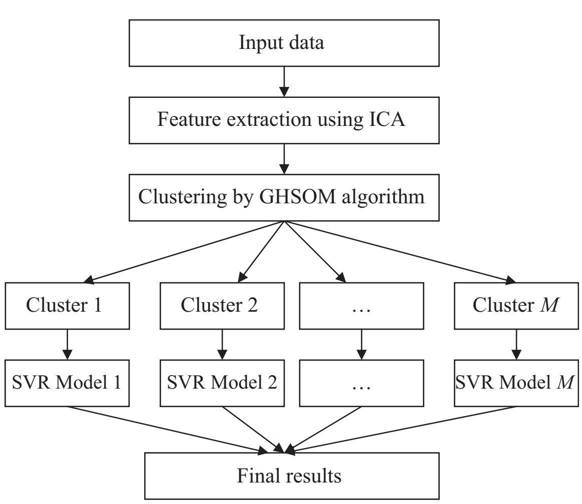

Fig. 1. The proposed hybrid ICA–GHSOM–SVR forecasting model.

2.2. Growing hierarchical self-organizing maps

Current research indicates that the self-organizing map (SOM) approach is able to detect structures from high-dimensional space to low-dimensional space (Hung and Tsai, 2008). Facing forecasting problems, most of this research was able to use SOM to cluster the original data set into the different groups. Forecast procedures could then be applied to each data set to make an accurate forecasting result. Tay and Cao (2001) integrated SOM and SVM to construct a financial time series forecasting model. Cao (2003) also combined SOM and SVM to deal with the time series forecasting problem. Lai et al. (2009) used a K-means procedure to cluster their data set, as well as applied a fuzzy decision tree to forecast the stock price. Huang and Tsai (2009) hybridized SVR with the self-organizing feature map (SOFM) technique and a filter-based feature selection to reduce the cost of training time and improve prediction accuracies.

Hsu et al. (2009) first used SOM to decompose the whole input space into regions where data points with similar statistical distributions were grouped together, so as to contain and capture the non-stationary property of financial series. After decomposing heterogeneous data points into several homogenous regions, SVR was applied to forecast financial indices. However, some shortcomings have to be mentioned. One possible solution for an extensive data set is using the SOM in a hierarchical manner. Rauber et al. (2002) proposed the GHSOM, which is based on a recursive SOM algorithm, to provide a SOM hierarchy automatically. Nevertheless, most research has applied GHSOM to deal with image recognition and web mining problems. Thus, this study will first apply GHSOM to the task of product demand forecasting.

2.3. Support vector regression forecasting model

Neural network is the most commonly used tool applied in forecasting. The most popular neural network training algorithm for forecasting is the back-propagation neural network (BPN), which has a simple architecture but powerful problem-solving ability. The BPN, however, suffers from a number of shortcomings, such as the need for a large number of controlling parameters, difficulty in obtaining a stable solution, and the risk of model over-fitting (Tay and Cao, 2001a, 2003; Chang et al., 2005). As a result of these weaknesses, many models have been developed to improve on the BPN model.

Support vector machines (SVMs) have a novel neural network algorithm based on statistical learning theory (Vapnik, 1999, 2000). They have great potential and provide superior performance in practical applications. This is largely due to the SVMs structure risk minimization principles, which have greater generalization ability and are superior to the empirical risk minimization principle utilized by traditional neural networks. SVMs guarantee global optima, whereas BPN face the risk of getting stuck in local optima and are not guaranteed to achieve global optima (Tay and Cao, 2001a, 2003). As a consequence of