

Support vector regression with genetic algorithms in forecasting tourism demand

Kuan-Yu Chen^{a,*}, Cheng-Hua Wang^b

^aDepartment of Business Administration, Far East College, 49 Chung Hua Road, Hsin-Shih, Tainan County 744, Taiwan, ROC

^bDepartment of Business Administration, Chang-Jung Christian University

Received 31 May 2005; accepted 5 December 2005

Abstract

This study applies a novel neural network technique, support vector regression (SVR), to tourism demand forecasting. The aim of this study is to examine the feasibility of SVR in tourism demand forecasting by comparing it with back-propagation neural networks (BPNN) and the autoregressive integrated moving average (ARIMA) model. To build an effective SVR model, SVR's parameters must be set carefully. This study proposes a novel approach, known as genetic algorithm (GA)-SVR, which searches for SVR's optimal parameters using real value GAs, and then adopts the optimal parameters to construct the SVR models. The tourist arrivals to China during 1985–2001 were employed as the data set. The experimental results demonstrate that SVR outperforms the BPNN and ARIMA models based on the normalized mean square error (NMSE) and mean absolute percentage error (MAPE).

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Support vector machine; Genetic algorithms; Support vector regression; Back-propagation neural networks; ARIMA

1. Introduction

In the past few decades, the tourism industry has emerged as the fastest growing sector, and has spread widely around the world. Tourism expenditure has become an important source of economic activity, employment, tax revenue, income and foreign exchange. Therefore, every country needs to understand its international visitors and tourism receipts, to help formulate responsive policies on tourism quickly. Furthermore, the tourism industry is characterized by unstockpiled economics (such as empty hotel rooms and air flight seats), resulting in a requirement for accurate forecasts of tourism demand, in both the short and long term.

In tourism demand forecasting, the relevant research models firstly emphasize econometrics (Hiemstra & Wong, 2002; Smeral, Witt, & Witt, 1992; Song & Witt, 2000). These investigations concluded that econometric models can help the policymakers formulate appropriate economic

strategies in order to influence tourism demand and generate accurate demand forecasts. However, such methods are time-consuming, costly and the influent factors are difficult to research. Although time series methods cannot fully explain the relations between variables, they require less empirical evidence than econometrics and, hence, are appropriate for tourism research issues. For instance, Lim and McAleer (2002) employed Box–Jenkins' autoregressive integrated moving average (ARIMA) model to forecast tourist arrivals to Australia from Hong Kong, Malaysia and Singapore. Goh and Law (2002) applied the time series SARIMA and MARIMA with interventions in forecasting tourism demand using ten arrival series for Hong Kong.

Recently, neural networks have been successfully employed for modelling time series. Unlike conventional statistical models, neural networks are data-driven, non-parametric, weak models which let “the data speak for themselves”. Law and Au (1999) first applied feed-forward neural networks to forecast Japanese demand for travel to Honk Kong. Law (2000) extended the applicability of neural networks in tourism demand forecasting by incorporating the back-propagation learning process into

*Corresponding author. Tel.: +886 6 5977611; fax: +886 6 5977610.
E-mail address: t530903@ms7.hinet.net (K.-Y. Chen).

nonlinearly separable tourism demand data. Experimental results demonstrate that the neural networks model forecasts outperformed multiple regression, moving average and exponent smoothing. However, the neural networks suffer from several weaknesses, such as the need for a large number of controlling parameters, difficulty in obtaining a stable solution and the danger of over-fitting.

In 1995, Vapnik developed a neural network algorithm called support vector machine (SVM), which is a novel learning machine based on statistical learning theory, and which adheres to the principle of structural risk minimization seeking to minimize an upper bound of the generalization error, rather than minimize the training error (the principle followed by neural networks). This induction principle is based on the bounding of the generalization error by the sum of the training error and a confidence interval term depending on the Vapnik–Chervonenkis (VC) dimension. Based on this principle, SVM achieves an optimum network structure by striking a right balance between the empirical error and the VC-confidence interval. This balance eventually leads to better generalization performance than other neural network models (Tay & Cao, 2001). Additionally, the SVM training process is equivalent to solving linearly constrained quadratic programming problems, and the SVM-embedded solution meaning is unique, optimal and unlikely to generate local minima. Originally, SVM has been developed to solve pattern recognition problems. However, with the introduction of Vapnik's ε -insensitive loss function, SVM has been extended to solve nonlinear regression estimation problems, such as new techniques known as support vector regression (SVR), which have been shown to exhibit excellent performance (Vapnik, Golowich, & Smola, 1997). Recently, SVR has emerged as an alternative and powerful technique to solve the nonlinear regression problem. It has achieved great success in both academic and industrial platforms due to its many attractive features and promising generalization performance.

Despite its superior features, SVR is limited in academic research and industrial applications because the user must define various parameters (so-called hyper-parameters) appropriately. To construct the SVR model efficiently, SVR's parameters must be set carefully (Duan, Keerthi, & Poo, 2001; Keerthi, 2002; Lin, 2001). Inappropriate parameters in SVR lead to over-fitting or under-fitting (Lin, 2001). Different parameter settings can cause significant differences in performance. Therefore, selecting optimal hyper-parameter is an important step in SVR design. However, no general guidelines are available to select these parameters (Cristianini & Shawe-Taylor, 2000; Gunn, 1997; Schölkopf & Smola, 2002; Vapnik, 1995, 1998). Therefore, in this study, we propose real-value genetic algorithms (RGA) to determine free parameters of SVR, known as GA-SVR, which optimizes all SVR's parameters simultaneously from the training data. Then,

the tourism demand, as represented by the number of world-wide visitors to China, was predicted. The proposed approach was compared with the back-propagation neural networks (BPNN) and traditional time series models, such as ARIMA, so as to show that the SVR model is substantially featured with an excellent forecasting capacity.

This study includes eight sections. Section 2 introduces theories related to SVR. Section 3 summarizes the existing practical approaches to the choice of hyper-parameters. Section 4 elaborates on the GA-SVR model proposed in this study. Section 5 describes the data source and experimental settings. Section 6 analyzes the results of RGA and optimizes SVR's parameters, and also explains the determining parameters process of the BPNN and ARIMA models. Section 7 discusses and analyzes the experimental results. Section 8 concludes the study and suggests directions for future investigations.

2. Support vector regression

Recently, a regression version of SVM has emerged as an alternative and powerful technique to solve regression problems by introducing an alternative loss function. In the sequel, this version is referred to as SVR. Here, a brief description of SVR is given. Detailed descriptions of SVR can be found in Vapnik (1995, 1997, 1998), Schölkopf and Smola (2002) and Cristianini and Shawe-Taylor (2000).

The SVR formulation follows the principle of structural risk minimization, seeking to minimize an upper bound of the generalization error rather than minimize the prediction error on the training set (the principle of empirical risk minimization). This equips the SVR with a greater potential to generalize the input–output relationship learnt during its training phase for making good predictions, for new input data. The SVR maps the input data x into a high-dimensional feature space F by nonlinear mapping, to yield and solve a linear regression problem in this feature space (see Fig. 1). The regression approximation estimates

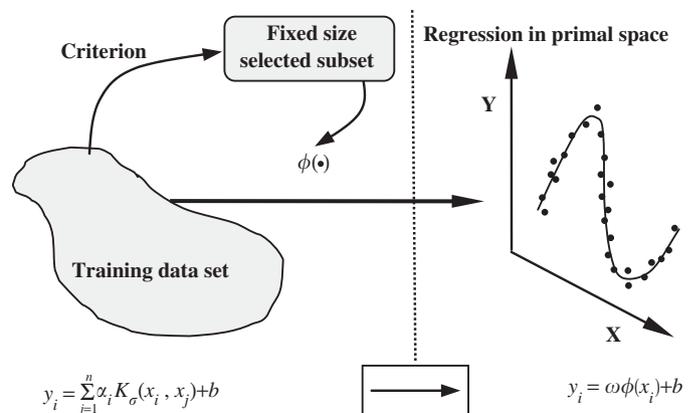


Fig. 1. Mapping input space x into high-dimensional feature space (from Peng et al., 2004).

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات