Artificial Intelligence Applications in Tourism

Carey Goh
The Hong Kong Polytechnic University, Hong Kong

Henry M. K. Mok
The Chinese University of Hong Kong, Hong Kong

Rob Law
The Hong Kong Polytechnic University, Hong Kong

INTRODUCTION

The tourism industry has become one of the fastest growing industries in the world, with international tourism flows in year 2006 more than doubled since 1980. In terms of direct economic benefits, United Nations World Tourism Organization (UNWTO, 2007) estimated that the industry has generated US $735 billion through tourism in the year of 2006. Through multiplier effects, World Travel and Tourism Council (WTTC, 2007) estimated that tourism will generate economic activities worth of approximately US $5,390 billion in year 2007 (10.4% of world GDP).

Owing to the important economic contribution by the tourism industry, researchers, policy makers, planners, and industrial practitioners have been trying to analyze and forecast tourism demand. The perishable nature of tourism products and services, the information-intensive nature of the tourism industry, and the long lead-time investment planning of equipment and infrastructures all render accurate forecasting of tourism demand necessary (Law, Mok, & Goh, 2007). Past studies have predominantly applied the well-developed econometric techniques to measure and predict the future market performance in terms of the number of tourist arrivals in a specific destination. In this chapter, we aim to present an overview of studies that have adopted artificial intelligence (AI) data-mining techniques in studying tourism demand forecasting. Our objective is to review and trace the evolution of such techniques employed in tourism demand studies since 1999, and based on our observations from the review, a discussion on the future direction of tourism research techniques and methods is then provided. Although the adoption of data mining techniques in tourism demand forecasting is still at its infancy stage, from the review, we identify certain research gaps, draw certain key observations, and discuss possible future research directions.

BACKGROUND

Econometric modeling and forecasting techniques are very much highly exploited to the understanding of international tourism demand. The econometric techniques have evolved and improved in accuracy and sophistication in recent years with the development of stationarity consideration, cointegration, and GARCH and time varying parameter (TVP) models. However, the development of such techniques might have reached a plateau for the time being. It is often expressed that econometric models are expanded at the expense of model comprehensiveness. That is, demand relationship is represented in a complicated system of equations in such models, which policy makers and practitioners find it like a black box that is hard to comprehend.

Besides, the conventional econometric models are founded on strict statistical assumptions and stringent economic theory (e.g., utility theory and consumption behavioral theory). These theories suggest that economic factors, such as income, price, substitute price, and advertising, are the primary influences of demand. However, tourism consumption, particularly if it involves a long-haul trip, is restricted by not only income constraint, but also on time constraint (Cesario & Knetsch, 1970; Morley, 1994), and the *intrinsic* properties of a particular tourism product or service that have not been modeled in the traditional demand framework (Eymann & Ronning, 1992; Morley, 1992; Papatheodorou, 2001; Rugg, 1973; Seddighi & Theocarous, 2002). The noneconomic factors, such as psychological, anthropological, and sociological factors, had not been analyzed in the traditional econometric travel demand studies. It is true that there exists a large amount of tourism literature that studies how these noneconomic factors could affect travel motivation, and how travel motivation affects destination choices (e.g., Edwards, 1979; Gray, 1970; Mayo & Jarvis, 1981; Um & Crompton, 1990), published articles have failed to show how noneconomic factors could affect demand for tourism, however.

Researchers have attributed the reasons for not including many relevant variables to the lack of data availability and difficulty in obtaining exact measures for the determining factors (Gonzelez & Moral, 1995; Kulendran, 1996, Song & Witt, 2000; Song *et al.*, 2000). Perhaps the true reason lies in the expense of the increasing complexity of a model in exchange for the inclusion of more determining factors, as noted in Turner and Witt (2001), Stabler (1990), as well as Faulker & Valerio (2000).

A RETROSPECTIVE VIEW ON DATA MINING IN TOURISM DEMAND ANALYSIS AND FORECASTING

In view of the growing importance of data mining in business applications, we review and analyze the relevant articles that adopt data-mining techniques to forecast tourism demand in tourism research journals. We sorted the 70 research journals in tourism, as mentioned by McKercher, Law, and Lam (2006), for the period between 1980 and early 2007, and the ScienceDirect, and the Hospitality and Tourism Complete index on EBSCOhost Web for the period between mid-2006 and early 2007. We identified 174 papers on tourism demand forecastin,g published in a 28-year span (1980 to early 2007) and among them, only 14 used data-mining techniques, with the first one published in 1999. Nine out of the 14 papers were published in the last 5 years (2003-early 2007).

A variety of neural network (NN) systems was adopted in these 14 studies, and they include supervised feed forward NN, back propagation NN, Elman's NN, multilayer perceptron NN, radial basis function NN, and Bayesian NN. Others have adopted rough sets, support vector regression, group method of data handling, fuzzy time series, and grey forecasting model.

Neural Networks

Traditionally, the term neural network had been used to refer to a network of biological neurons, but it is often refered to as artificial neural networks (ANN) nowadays. ANNs are computer software that mimics the human intelligence to deduce or learn from a dataset. What gives ANNs excellent classification and pattern recognition ability are their capabilities of representing knowledge based on massive parallel processing, and pattern recognition based on past experience. This pattern recognition ability makes ANNs a superb classification and forecasting tool for industrial applications.

Supervised Feed Forward Neural Networks (SFFNN)

The SFFNN was first adopted by Uysal and El Roubi (1999) and Law and Au (1999) in tourism demand studies. Uysal and El Roubi (1999) developed a preliminary neural network that used Canadian tourism expenditures in the United States as dependent variable; whereas per capita income of Canada, consumer price index, lagged expenditure by one period, and seasonal dummy variables were used as independent variables. Their findings showed that the neural network achieved highly accurate results with high-adjusted correlations and low-error terms. Law and Au (1999) used the model to forecast Japanese tourist arrivals to Hong Kong. Using six independent variables, including relative service price, average hotel rate in Hong Kong, foreign exchange rate, population in Japan, marketing expenses by the Hong Kong Tourist Association, and gross domestic expenditure in Japan, the neural network achieved the lowest mean average percentage error (MAPE), and highest acceptance percentage and normalized correlation coefficient compared to four other commonly used econometric models, including multiple regression, naïve I, moving average, and single exponential smoothing.

The SFFNN was adopted by Law (2001) in an attempt to capture turbulent travel behavior as a result of regional financial crisis in Asia in 1997. Law (2001) tested and compared the forecasting accuracy of this model with naïve I, naïve II, moving average, single exponential smoothing, Holt's exponential smoothing, and multiple regression in predicting Japanese demand for travel to Hong Kong. Using the same variables as in Law and Au (1999) and Law (2000), but with updated data, neural network, again, outperformed other techniques in three of the five accuracy measurements, including MAPE, trend change accuracy, and the closeness between actual and estimated values. Law (2001) concluded that no single forecasting method could outperform others in all situations when there was a sudden environmental change, but neural network appeared to perform reasonably well in terms of predicted values. Using tourism demand data of South Africa, Burger *et al.* (Burger, Dohnal, Kathrada, & Law, 2001) also investigated performance and SFFNN with a few benchmark models, such as Naïve I, moving average, decomposition, single exponential smoothing, ARIMA, autoregression, and found that neural networks with 1-month-ahead forecast generated the lowest MAPE values. Radial basis function network, a type of SFFNN, was employed by Kon and Turner (2005), but the forecasting performance falls behind that of multilayer peceptron NN model in all seasonal and deseasonalized series.

Related Content

Cost Performance Estimation in Construction Projects Using Fuzzy Time Series
Mostafa Salari, Nooshin Yousefiand Mohhmad Mahdi Asgary (2015). *International Journal of Information Technology Project Management (pp. 66-75).*
www.irma-international.org/article/cost-performance-estimation-in-construction-projects-using-fuzzy-time-series/123466

Theoretical Justification for IT Infrastructure Investments
Timothy R. Kayworth, Debabroto Chatterjeeand V. Sambamurthy (2002). *Advanced Topics in Information Resources Management, Volume 1 (pp. 73-89).*
www.irma-international.org/chapter/theoretical-justification-infrastructure-investments/4579

How Do Actuaries Use Data Containing Errors?: Models of Error Detection and Error Correction
Barbara D. Klein (1997). *Information Resources Management Journal (pp. 27-36).*
www.irma-international.org/article/actuaries-use-data-containing-errors/51041

COVID-19 Pandemic: Insights of Newspaper Trends
Jasdeep Kaur, Amit Chhabra, Munish Sainiand Nebojsa Bacanin (2022). *Journal of Information Technology Research (pp. 1-23).*
www.irma-international.org/article/covid-19-pandemic/299390

Assessing the Effectiveness of an E-Learning Framework: The Portuguese Insurance Academy Case
Nuno Penaand Pedro Isaías (2013). *Journal of Cases on Information Technology (pp. 1-18).*
www.irma-international.org/article/assessing-effectiveness-learning-framework/78354