Expert Systems with Applications 37 (2010) 1185–1191

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Forecasting tourist arrivals by using the adaptive network-based fuzzy inference system

Miao-Sheng Chen [a], Li-Chih Ying [b,*], Mei-Chiu Pan [a,1]

[a] Department of Business Administration, Nanhua University, 32, Chung Keng Li, Dalin, Chiayi 622, Taiwan, ROC
[b] Department of Marketing Management, Central Taiwan University of Science and Technology, 666, Po-tzu Road, Pei-tun District, Taichung City 40601, Taiwan, ROC

ARTICLE INFO

Keywords:
Adaptive network-based fuzzy inference system
Tourist arrivals

ABSTRACT

Since accurate forecasting of tourist arrivals is very important for planning for potential tourism demand and improving the tourism infrastructure, various tourist arrivals forecasting methods have been developed. The purpose of this study is to apply the adaptive network-based fuzzy inference system (ANFIS) model to forecast the tourist arrivals to Taiwan and demonstrate the forecasting performance of this model. Based on the mean absolute percentage errors and statistical results, we can see that the ANFIS model has better forecasting performance than the fuzzy time series model, grey forecasting model and Markov residual modified model. Thus, the ANFIS model is a promising alternative for forecasting the tourist arrivals. We also use the ANFIS model to forecast the monthly tourist arrivals to Taiwan from Japan, Hong Kong and Macao, and the United States.

1. Introduction

The tourism industry has grown rapidly over the past few decades. Due to the perishable nature of the tourism industry, the need for accurate forecasts is crucial. Both the public non-profit government sectors and the private profit companies are interested in finding an accurate forecasting technique to make operational, tactical and strategic decisions. Companies can, using the results from this forecasting technique, plan for potential tourism demand successfully and invest in tourism related facilities and equipments sufficiently, and government sectors can play a significant role in maintaining and improving tourism infrastructure. Therefore, various forecasting methods have been developed. They include exponential smoothing (Cho, 2003), ARIMA (Cho, 2003; Chu, 1998; Goh & Law, 2002; Lim & McAleer, 2002), vector autoregressive (Song & Witt, 2006; Wong, Song, & Chon, 2006), neural networks (Chen & Wang, 2007; Cho, 2003; Law, 2000; Law & Au, 1999), fuzzy time series (Wang, 2004; Wang & Hsu, 2008), grey model (Hsu & Wen, 1998; Wang, 2004), econometric (Hiemstra & Wong, 2002; Smeral, Witt, & Witt, 1992; Song & Witt, 2000; Witt & Martin, 1987), regression-based model (Chan, 1993; Crouch, Schultz, & Valerio, 1992; Kulendran & Witt, 2001) and genetic algorithm (Chen & Wang, 2007; Hernández-López & Cáceres-Hernández, 2007; Hernández-López, 2004; Hurley, Moutinho, & Witt, 1998). But we can not find any paper adopting an adaptive network-based fuzzy inference system, referred to as ANFIS (Jang, 1993), to forecast tourist arrivals. So, the purpose of this paper is to fill this gap, and we also try to compare the results with those of other models and use the ANFIS model to forecast the monthly tourist arrivals to Taiwan from the top three markets.

In this paper, the data used were from the Tourism Bureau of Republic of China (ROC), and, for comparison, the annual tourist arrivals to Taiwan from the three markets: Hong Kong, the United States and Germany from 1989 to 2003 were considered. But, according to the numbers of tourist arrivals to Taiwan, we apply the ANFIS to forecast the monthly tourist arrivals to Taiwan from the top three markets: Japan, Hong Kong and Macao, and the United States.

The rest of this paper is organized as follows: Section 2 introduces the architecture and the hybrid learning algorithm of an ANFIS with a simple illustration. Section 3 compares the forecasting accuracy for the different models. Section 4 presents the application of the ANFIS to forecast the monthly tourist arrivals to Taiwan from Japan, Hong Kong and Macao, and the United States. The last section contains some concluding remarks.

2. Adaptive network-based fuzzy inference system

An ANFIS (Jang, 1993) can help us find the mapping relation between the input and output data through hybrid learning to determine the optimal distribution of membership functions. Five layers are used to construct this inference system. Each layer contains several nodes described by the node function. Adaptive nodes, denoted by squares, represent the parameter sets that are adjustable

* Corresponding author. Tel.: +886 4 2239 0071; fax: +886 4 2239 4465.
E-mail address: yingdear@gmail.com (L.-C. Ying).

[1] Ph.D. student.

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.06.032

in these nodes, whereas fixed nodes, denoted by circles, represent the parameter sets that are fixed in the system. The output data from the nodes in the previous layers will be the input in the present layer.

To illustrate the procedures of an ANFIS, for simplicity, we consider only two inputs x, y and one output f_{out} in this system. The framework of ANFIS is shown in Fig. 1, and the node function in each layer is described below.

Layer 1: Every node in this layer is an adaptive node with node function as:

$$O_{1,i} = \mu_{A_i}(x), \quad \text{for} \quad i = 1,2 \tag{1}$$

$$O_{1,i} = \mu_{B_{i-2}}(y), \quad \text{for} \quad i = 3,4 \tag{2}$$

where x(or y) is the input of the node, A_i(or B_j) is the linguistic label, $\mu(x)$(or $\mu(y)$) is the membership function, usually adopting bell shape with maximum and minimum equal to 1 and 0, respectively, as follows:

$$\mu(x) = \frac{1}{1 + \left(\frac{x-c_i}{a_i}\right)^{2b_i}} \tag{3}$$

or

$$\mu(x) = \exp\left\{-\left(\frac{x-c_i}{a_i}\right)^2\right\} \tag{4}$$

where $\{a_i, b_i, c_i\}$ is the parameter set. As the values of these parameters change, the bell shaped functions vary accordingly. The parameters in this layer are named premise parameters.

Layer 2: Every node in this layer is a fixed node, marked by a circle and labeled $\prod$, with the node function to be multiplied by input signals to serve as output signal

$$O_{2,i} = \mu_{A_i}(x) \cdot \mu_{B_i}(y) = \omega_i \quad \text{for } i = 1,2 \tag{5}$$

The output signal ω_i represents the firing strength of a rule.

Layer 3: Every node in this layer is a fixed node, marked by a circle and labeled N, with the node function to normalize the firing strength by calculating the ratio of the ith node firing strength to the sum of all rules' firing strength.

$$O_{3,i} = \frac{\omega_i}{\sum \omega_i} = \frac{\omega_i}{\omega_1 + \omega_2} = \bar{\omega}_i \quad \text{for } i = 1,2 \tag{6}$$

Layer 4: Every node in this layer is an adaptive node, marked by a square, with node function

$$O_{4,i} = \bar{\omega}_i \cdot f_i \quad \text{for } i = 1,2 \tag{7}$$

where f_1 and f_2 are the fuzzy if-then rules as follows:

Rule 1: if x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$
Rule 2: if x is A_2 and y is B_2 then $f_2 = p_2x + q_2y + r_2$
and where $\{p_i, q_i, r_i\}$ is the parameters set, referred to as the consequent parameters.

Layer 5: Every node in this layer is a fixed node, marked by a circle and labeled Σ, with node function to compute the overall output by

$$O_5 = \sum_i \bar{\omega}_i \cdot f_i = f_{out} \tag{8}$$

As mentioned above, an ANFIS is a multilayer feedforward network in which each node performs a node function on incoming signals as well as a set of parameters belonging to this node. Suppose that the given training data set has n entries. We define the overall error measure by

$$E = \sum_{i=1}^{n} E_i = \sum_{i=1}^{n} (T_i - f_{outi})^2 \tag{9}$$

where E_i is the error measure for the ith entry of the given training data set, T_i is the desired output of the ith entry and f_{outi} is the output of the ANFIS using the ith entry.

From the architecture of the ANFIS, we know that if the premise parameters $\{a_i, b_i, c_i\}$ are fixed, the output f_{outi} of the whole system will be a linear combination of the consequent parameters $\{p_i, q_i, r_i\}$ as follows:

$$\begin{aligned} f_{out} &= \sum \bar{\omega}_i \cdot f_i = \bar{\omega}_1 \cdot f_1 + \bar{\omega}_2 \cdot f_2 = \bar{\omega}_1(p_1x + q_1y + r_1) \\ &\quad + \bar{\omega}_2(p_2x + q_2y + r_2) = (\bar{\omega}_1 x)p_1 + (\bar{\omega}_1 y)q_1 + \bar{\omega}_1 r_1 \\ &\quad + (\bar{\omega}_2 x)p_2 + (\bar{\omega}_2 y)q_2 + \bar{\omega}_2 r_2 \end{aligned} \tag{10}$$

Let matrices

$$f = \begin{bmatrix} f_{out1} \\ f_{out2} \\ \cdot \\ \cdot \\ \cdot \\ f_{outn} \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} p_1 \\ q_1 \\ r_1 \\ p_2 \\ q_2 \\ r_2 \end{bmatrix}, \text{ and}$$

$$\mathbf{B} = \begin{bmatrix} \bar{\omega}_1 x_1 & \bar{\omega}_1 y_1 & \bar{\omega}_1 & \bar{\omega}_2 x_1 & \bar{\omega}_2 y_1 & \bar{\omega}_2 \\ \bar{\omega}_1 x_2 & \bar{\omega}_1 y_2 & \bar{\omega}_1 & \bar{\omega}_2 x_2 & \bar{\omega}_2 y_2 & \bar{\omega}_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{\omega}_1 x_n & \bar{\omega}_1 y_n & \bar{\omega}_1 & \bar{\omega}_2 x_n & \bar{\omega}_2 y_n & \bar{\omega}_2 \end{bmatrix} \tag{11}$$

Then, Eq. (10) can be expressed in matrix form as

$$\boldsymbol{f} = \mathbf{B}\theta \tag{12}$$

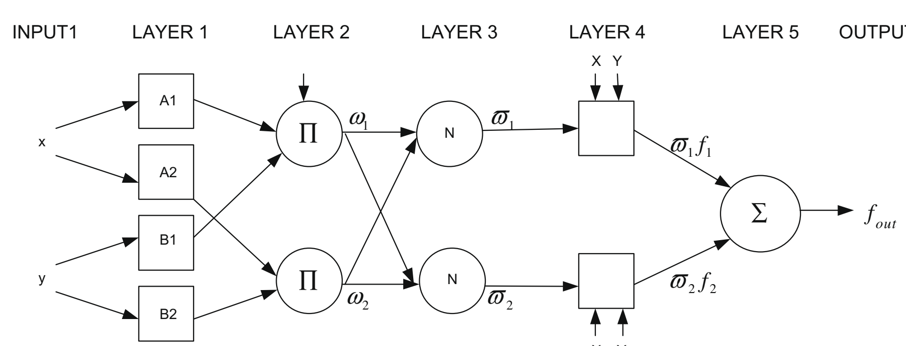

Fig. 1. The framework of an ANFIS.