TELKOMNIKA, Vol.11, No.11, November 2013, pp. 6413~6419
e-ISSN: 2087-278X

A Hybrid Neural Network Prediction Model of Air Ticket Sales

Han-Chen Huang
Department of Leisure Management, Yu Da University, Taiwan
e-mail: hchuang@ydu.edu.tw

Abstract

Air ticket sales revenue is an important source of revenue for travel agencies, and if future air ticket sales revenue can be accurately forecast, travel agencies will be able to advance procurement to achieve a sufficient amount of cost-effective tickets. Therefore, this study applied the Artificial Neural Network (ANN) and Genetic Algorithms (GA) to establish a prediction model of travel agency air ticket sales revenue. By verifying the empirical data, this study proved that the established prediction model has accurate prediction power, and MAPE (mean absolute percentage error) is only 9.11%. The established model can provide business operators with reliable and efficient prediction data as a reference for operational decisions.

Keywords: *artificial neural network, genetic algorithms, air ticket, sales revenue*

1. Introduction

Travel is not a necessity in daily livelihood, thus, the tourism industry is directly impacted during economic downturn. Travel agencies need to have good financial management and risk management capabilities in order to survive in the competitive market [1]. Financial management is critical to travel agencies. The strengthening of financial management can increase income and reduce costs. By analyzing and addressing financial problems, competitiveness can be enhanced. Among the business scope of travel agencies, air ticket sales are an important source of revenue. Travel agencies purchase a large number of discounted air tickets from airliners and sell to consumers [2]. If a travel agency can accurately predict market demand, they can purchase sufficient numbers of discounted air tickets in advance to gain higher profits, and reduce the accumulated costs caused by over purchases or order losses caused by stock depletion [3]. Therefore, this study applied the Artificial Neural Network (ANN) and Genetic Algorithms (GA) to establish a prediction model for air ticket sales revenue. The research findings can provide the industry with a practical reference of higher reliability and efficiency.

2. Literature Review

2.1. Travel Agency

Travel agencies are the intermediary between tourism product suppliers and customers, and are responsible for planning and arrangements of travel for customers to gain savings [3]. According to Taiwan's Statute for the Development of Tourism [4] Article 2: "Travel enterprise: also referred to as travel agency, a profit-taking enterprise licensed by the central administrative authority to provide tourists with arranged travel schedule, board and lodging, tour guide, and to purchase transportation tickets and apply for travel documents and visas on tourists' behalf, as well as to provide related services for remuneration". Taiwan's Regulations Governing Travel Agencies [5] Article 3: Travel agencies are divided into consolidated travel agencies, Class-A travel agencies, and Class-B travel agencies:

A. The business scope of consolidated travel agencies shall consist of the following:

1) Being commissioned to sell passenger tickets for domestic or foreign land, sea, and air transportation operators, or to buy domestic or overseas passenger tickets and to handle shipment of luggage for travelers.

Received March 29, 2013; Revised June 17, 2013; Accepted July 1, 2013

2) Being commissioned to handle inbound and outbound procedures and visa applications for travelers.
3) Soliciting of business from or receiving domestic and overseas tourists, and arranging for their tours, food, accommodation, and transportation.
4) Arranging domestic and overseas tours, food and accommodation, and transportation for travelers as well as related services, through chartered or self-organized tour packages.
5) Commissioning of Class-A travel agencies to solicit business.
6) Commissioning of Class-B travel agencies to solicit domestic tour group business.
7) Handling of communication, promotion, and price quotation on behalf of foreign travel agencies.
8) Planning of domestic and overseas tours, and arranging for tour guides or tour managers.
9) Providing of advice on domestic and overseas tours.
10) Operating other domestic and overseas tour-related businesses as approved by the central administrative authority.

B. The business scope of Class-A travel agencies shall consist of the following:

1) Being commissioned to sell passenger tickets for domestic or foreign land, sea, and air transportation operators, or to buy domestic or overseas passenger tickets and to handle shipment of luggage for the traveler.
2) Being commissioned to handle inbound and outbound procedures and visa applications for travelers.
3) Soliciting of business from or receiving domestic and overseas tourists; arranging for their tour, food, accommodation, and transportation.
4) Arranging of overseas tours, food and accommodation, and transportation for travelers as well as provision of related services, through self-organized tour packages.
5) Soliciting of businesses on behalf of consolidated travel agencies.
6) Planning of domestic and overseas tours, and arranging for tour guides or tour managers.
7) Providing of advice on domestic and overseas tours.
8) Operating other domestic and overseas tour-related businesses as approved by the central administrative authority.

C. The business scope of Class-B travel agencies shall consist of the following:

1) Being commissioned to sell passenger tickets for domestic land, sea, and air transportation operators, or to buy domestic passenger tickets and handle shipment of luggage on behalf of travelers.
2) Soliciting of business from or receiving domestic tourists; arranging for tours, food and accommodation, and transportation; and providing of related services.
3) Soliciting of businesses related to domestic group tours on behalf of consolidated travel agencies.
4) Planning of domestic tours.
5) Providing of advice on domestic tours.
6) Operation of other domestic tour-related businesses as approved by the central administrative authority.

2.2. Artificial Neural Network

ANN is a mathematical model for simulating the structure and function of a biological neural network. An ANN conducts computation by large numbers of artificial neurons [6, 7]. In most cases, ANN can change an internal structure according to external information as a type of adaptive system [6-8]. ANN is a non-linear statistical data modeling tool commonly used for the model construction of complex input-output relationships or the exploration of a data model [9].

The ANN construction principle is originated by the operation of biological (human or other animals) neural networks. ANN, like humans, has simple decision and judgment capabilities, which are more advantageous as compared with formal logical reasoning [7-10]. A common multilayer feedforward network consists of three parts (Figure 1) [11, 12]:

1) Input layer, a large number of neurons for the reception of a large amount of input information.

2) Output layer, the information is transferred, analyzed, and weighed in the neuron connections to form the output results.
3) Hidden layer, the layers formed by the numerous neurons and links between the input layer and output layer. It can be of multiple layers, but is usually only one layer. There is no recognized standard for the number of neurons in the hidden layer. However, networks with more neurons are more significantly non-linear, and thus, the neural network robustness will be more significant.

The Back Propagation Neural Network (BPNN) is the currently the most representative and most commonly used ANN [11-13]. BPNN use the steepest descent method to adjust the network parameters to determine more accurate solutions by iterative computing.

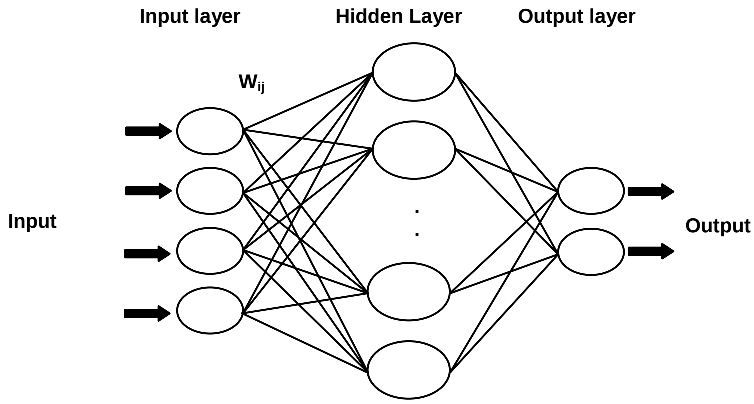

Figure 1. BPNN Architecture

2.3. Genetic Algorithms (GA)

The GA is a random search method that simulates the survival of the fittest biological evolutionary law, as proposed by John Henry Holland [14]. With the search algorithm of "survival and detection", it has been widely used in combinatorial optimization, machine learning, signal processing, self-adaptive control, and other fields. GA is one of the key technologies relating to intelligent computing of the present and has been applied by many enterprises for timetable arrangements, data analysis, future trend prediction, budgeting, and solving many other combinatorial optimization problems [15-17].

In GA, the solution to a problem is called an individual, and uses a variable sequence (known as chromosome). The chromosome is generally expressed as a simple string or numeric string. Firstly, the algorithm randomly generates a certain number of individuals. In each generation, each individual is evaluated and a fitness value is obtained by the computation of the fitness function. Individuals of the population will be sorted according to fitness values in descending order. The next step is to generate the next generation of individuals, and thus, the formation of the population. This process is implemented via selection, crossover, and mutation [16, 18]. When the fitness value is higher, the opportunity of being selected will be higher. A relatively optimized population is formed through such a selection mechanism. Afterwards, the selected individuals undergo the process of crossover. In general, GA has a range of crossover probabilities, being 0.6-1. Furthermore, it is the action of mutation, through which new "sub-individuals" are generated. In general, GA has a fixed mutation constant of 0.1 or below. After selection, crossover, and mutation, the best individual has more opportunities of being selected to generate the next generation, while individuals of lower fitness value are gradually eliminated. Such a process is repeated until the ending conditions are satisfied. The general ending conditions are as follows [15-19]:
1) Limited evolutionary times;
2) Computing resource constraints (such as the computation time and memory occupied by computation);
3) An individual has met the conditions of the optimal value;
4) Continued evolution does not produce individuals with better fitness values;

5) Human intervention;
6) Combinations of two or more of the situations described in the above.

3. Research Method
3.1. Input Variables

Based on literature review [20-26], this study used international oil price, Taiwan stock market weighted index, NTD/USD exchange rate, people traveling abroad from Taiwan each month, Taiwan's monthly unemployment rate, Taiwan's monthly monitor indicator, Taiwan's monthly composite leading index, Taiwan's monthly composite coincident index, and K travel agency's monthly air ticket sales revenue (T-1 month~T-12 month) (Table 1), as the input variables for the prediction of K travel agency's air ticket sales revenue in T month. The time period of data selection is from January 2004 to August 2012. Moreover, 68% data is randomly selected as the training data, 16% as the cross validation data, and 16% as the testing data. GA improves the performance of ANNs by selecting the optimum input features. This study used different operators for selection (Roulette, Best, Random, and Tournament) and crossover (One point, Uniform, Arithmetic, and Heuristic) operations.

Table 1. Forecast Model Variables

	Variables
	International oil price (T-1 month)
	Taiwan stock market weighted index(T-1 month)
	NTD/USD exchange rate(T-1 month)
	People traveling abroad from Taiwan each month (T-1 month)
Input	Taiwan's monthly unemployment rate(T-1 month)
	Taiwan's monthly monitor indicator(T-1 month)
	Taiwan's monthly composite leading index (T-1 month)
	Taiwan's monthly composite coincident index (T-1 month)
	K travel agency's air ticket sales revenue (T-1 month to T-12 month)
Output	Air ticket sales revenue(T month)

3.2. Architecture Design and Model Training

Regarding ANN, the input activation function uses Hyperbolic Tangent, the output error function uses Sum-Of-Squares, and the output activation function uses Logistic. The architecture design uses Test Error as the judgment standard when searching for the optimal network architecture (Figure 2).

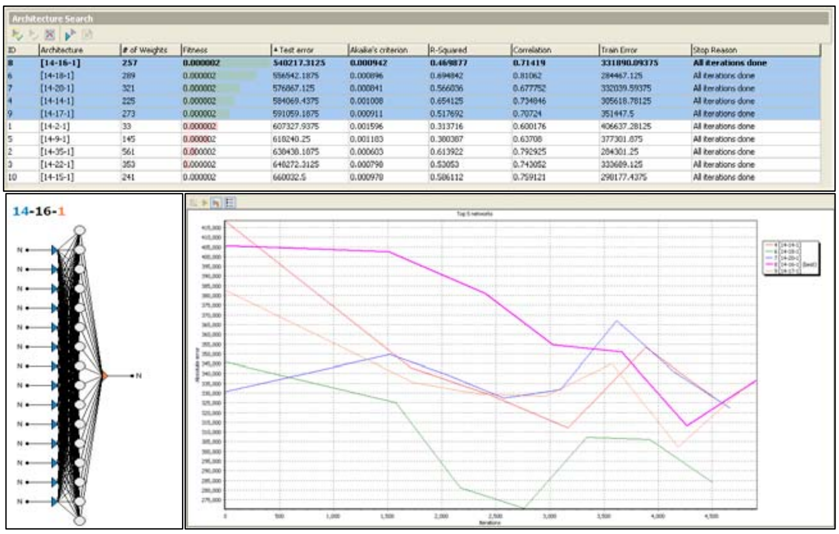

Figure 2. Best Network Architecture Search Results

ANN Training Algorithm: Quick Propagation Algorithm, Training Algorithm's Parameters include: Quick Propagation Coefficient= 1.75, Learning Rate= 0.1. Overtraining control and the weights randomization method are used to increase model accuracy (Figure 3).

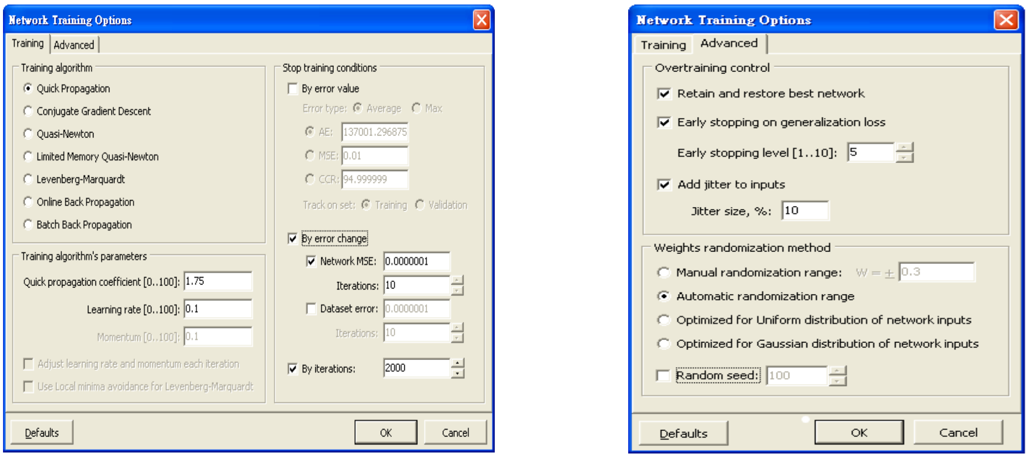

Figure 3. Network Training Options

4. Empirical Results

Mean absolute percentage error (MAPE) and correlation (r) are adopted as indicators for evaluating the model.

1) MAPE: The smaller the value, the smaller the error between the forecast value and the target value.

$$MAPE = \frac{1}{n}\sum_{1}^{n}\left|\frac{y_i - y'_i}{y_i}\right| \cdot 100\ \% \tag{1}$$

Where n is the number of the forecasting periods, y_i is the actual value for the i period, and y'_i is the forecast value for the i period.

2) Correlation (r): As r approaches 1, the ANN forecasting results improve.

The actual values and model output values distributions are as shown in Figure 4. Most of the output values are distributed along both sides of the diagonal line (Output/Target=1), indicating that the model has good predictability.

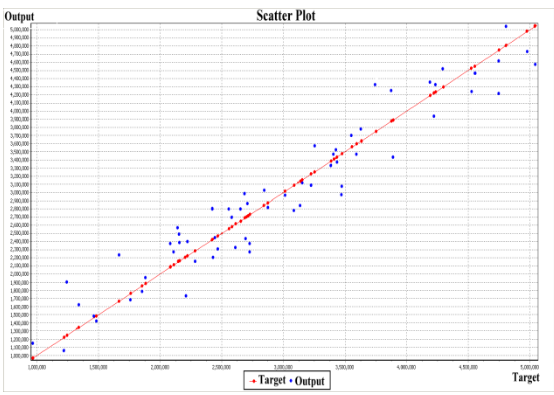

Figure 4. Scatter Plot of Actual Value and Model Output Value

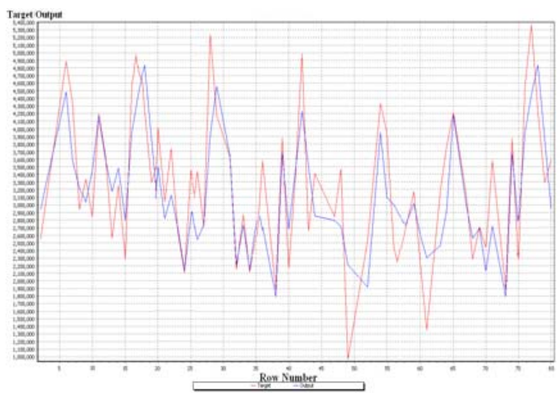

Figure 5. Actual Value and Model Output Value Graph

Table 2. Model Training Results

	Target	Output	Absolute Error	MAPE
Mean	2771725.58	2800445.86	297705.03	0.0894
SD	575737.53	541518.39	232589.28	0.1048
Correlation:0.87				

The actual value and model output value trends are as shown in Figure 5. The prediction model established for air ticket sales revenue in this study has good capability to reflect the changes in air ticket sales. The results of the model training are as shown in Table 2. Mean absolute percentage error (MAPE) is 8.94%, Correlation Coefficient is 0.87, indicating that the training model can learn the training data to a level of error below 10%. The model testing results are as shown in Table 3, MAPE is 9.11%, Correlation Coefficient is 0.83, representing that the prediction model has the ability for accurate prediction of air ticket sales revenue. The user interface of the established model is as shown in Figure 6.

Table 3. Model Forecast Results

	Target	Output	Absolute Error	MAPE
Mean	2988092.89	2835198.96	331872.74	0.0911
SD	625322.11	422310.58	277259.95	0.1169
Correlation:0.83				

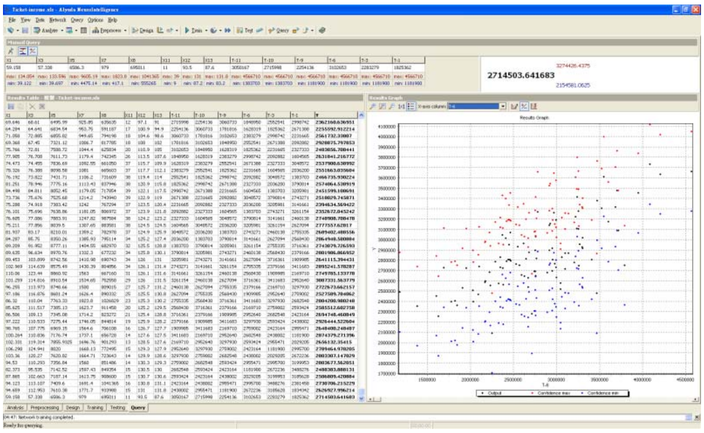

Figure 6. The User Interface of the Established Model

5. Conclusion

This study used BPNN and GA to establish a prediction model for travel agencies on air ticket sales revenue. Empirical results showed that, the proposed model has good predictability and the prediction model has the ability of accurately predicting air ticket sales revenue. The MAPE is 9.11%, and Correlation Coefficient is 0.83. If a travel agency can predict the changes in air ticket sales revenue in advance, it can purchase discounted air tickets in sufficient numbers. The proposed model can provide business with a more reliable and efficient reference basis in practice.

References

[1] Zhu Z, Zhao J. Comparative Study on the Sources of e-Business Competitive Advantages between Travel Agencies and Online Travel Service Firms. *Journal of China University of Geosciences Social Sciences Edition*. 2011; 11(2): 114-119.

[2] Chen SY, Liu HH. A Study on Business Models of On-Line Travel Agency-The Case of Eztravel. *Journal of Leisure and Tourism Industry Research*. 2010; 5(1): 90-99.

[3] Dai MH. Unfair Competition in Travel Service Industry. *Journal of Eastern Liaoning University (Social Sciences)*. 2011; 13(2): 38-42.
[4] Yang CK. Tourism Administration and Laws. Taipei: Yang Chih Ltd. 2012.
[5] Niu HY. Travel Industry Management. Taipei: Yang Chih Ltd. 2005.
[6] Guo H, Neural Network Aided Kalman Filtering for Integrated GPS/INS Navigation System. *TELKOMNIKA*. 2013; 11(3): 1221-1226.
[7] Hou Y, Mai Y. Chaotic Prediction for Traffic Flow of Improved BP Neural Network. *TELKOMNIKA*. 2013; 11(3): 1683-1690.
[8] Uncini A, Vecci L, Campolucci P. Piazza F. Complex-Valued Neural Networks with Adaptive Spline Activation Functions. *IEEE Transactions on Signal Processing*. 1999; 47(2): 505-514.
[9] Elhadi GF, Farouk RM, Issa AT. Protein Sequence for Clustering DNA based on Artificial Neural Networks. *International Journal of Computer Science Issues*. 2012; 9(1): 161-167.
[10] You C, Hong D. Nonlinear Blind Equalization Schemes using Complex-Valued Multilayer Feedforward Neural Networks. *IEEE Transactions on Neural Networks*. 1998; 9: 1442-1455.
[11] Huang HC, Chang AY, Ho CC. Using Artificial Neural Networks to Establish a Customer-Cancellation Prediction Model. *Przeglad Elektrotechniczny.* 2013; 89(1b): 178-180.
[12] Chang HT, Chen PC, Huang HC, Lin DH. A Study on the Application of Neural Network to the Prediction of Weight Control. *International Journal of Engineering Research and Development*. 2013; 5(11): 78-85.
[13] Huang HC, Ho CC. Back-Propagation Neural Network Combined with a Particle Swarm Optimization Algorithm for Travel Package Demand Forecasting. *International Journal of Digital Content Technology and its Applications*. 2012; 6(17): 194-203.
[14] Holland JH. Adaptation in Natural and Artificial Systems. Miami: The University of Michigan. 1975.
[15] Chen Y, Ma YJ, Yun WX. Application of Improved Genetic Algorithm in PID Controller Parameters Optimization. *TELKOMNIKA*. 2013; 11(3): 1524-1530.
[16] Boumediene L, Gao Z, Liu S, Leghmizi S, Yang R. Genetic Algorithm-Based Approach to Spectrum Allocation and Power Control with Constraints in Cognitive Radio Networks. *Research Journal of Applied Sciences, Engineering and Technology*. 2013; 5(1): 100-107.
[17] Guan X, Meng G, Zhao C. Manufacturing Resources Classification Based on Fuzzy Clustering Algorithm. *Journal of Theoretical and Applied Information Technology.* 2012; 46(2): 632-637.
[18] Rahmani R, Othman MF, Yusof R, Khalid M. Solving Economic Dispatch Problem using Particle Swarm Optimization by an Evolutionary Technique For Initializing Particles. *Journal of Theoretical and Applied Information Technology*. 2012; 46(2): 526-536.
[19] Peng J, Huang D, Wang Z. Profit-Driven Dynamic Spectrum Allocation in Wireless Networks. *Journal of Theoretical and Applied Information Technology*. 2012; 46(2): 615-618.
[20] Chan YM. Forecasting Tourism: A Sine Wave Time Series Regression Approach. *Journal of Travel Research*. 2011; 32(1): 58-60.
[21] Wang CH. Predicting Tourism Demand using Fuzzy Time Series and Hybrid Grey Theory. *Tourism Management*. 2004; 25(1): 367-374.
[22] Chu FL. Forecasting Tourism: A Combined Approach. *Tourism Management*. 1998; 19(6): 515-520.
[23] Chaiboonsri C, Chaitip P. The Modeling International Tourism Demand for Length of stay in India: Social Development, Economics Development. *Journal of Knowledge Management, Economics and Information Technology*. 2012; 2(2): 81-98.
[24] Lin CJ, Chen HF, Lee TS. Forecasting Tourism Demand using Time Series, Artificial Neural Networks and Multivariate Adaptive Regression Splines: Evidence from Taiwan. *International Journal of Business Administration*. 2011; 2(2): 14-24.
[25] Samsudin R, Saad P, Shabri A. Hybridizing GMDH and Least Squares SVM Support Vector Machine for Forecasting Tourism Demand. *International Journal of Research and Reviews in Applied Sciences*. 2010; 3(3): 274-279.
[26] Petrevska B. Forecasting International Tourism Demand: The Evidence of Macedonia. *UTMS Journal of Economics*. 2012; 3(1): 45-55.