



### Prediction of Tourism Demand in Iran by Using Artificial Neural Network (ANN) and Supporting Vector Machine (SVR)

Seyedehelham Sadatiseyedmahalleh<sup>1\*</sup>; Nasim Heidari Bateni; Nazanin Heidari Bateni<sup>2</sup>

<sup>1</sup> School of Housing, Building and Planning, Universiti Sains Malaysia, Penang, Malaysia

<sup>2</sup> Applied Science Anhalt, Germany

Email: [elham.sadatii@gmail.com](mailto:elham.sadatii@gmail.com)

---

#### **Abstract**

This research examines and proves this effectiveness connected with artificial neural networks (ANNs) as an alternative approach to the use of Support Vector Machine (SVR) in the tourism research. This method can be used for the tourism industry to define the tourism's demands in Iran. The outcome reveals the use of ANNs in tourism research might result in better quotations when it comes to prediction bias and accuracy. Even more applications of ANNs in the context of tourism demand evaluation is needed to establish and validate the effects.

**KeyWords:** Iran; Tourism; Artificial Neural Network (ANN); Support Vector Machine (SVR)

#### **Introduction**

Studying on tourist motivation has always been an important factor for tourism management. The ways of expressing and the changes that develop in time in tourist motivations directly influence tourism demand and supply. In order to survive in a market in constant evolution and change, characterized by fierce competition, in an economy in crisis, suppliers of tourism products and services need to know and anticipate changes in the motivations that determine tourists to buy a holiday package. Broad topic, heavily debated, tourist motivation can be defined as a psychological stimulus able to determine the desire to travel.

Accurate prediction and estimation of the tourism demand is very important in the making decision regarding investments in both the public and private sectors, also in the short term marketing decisions in the tourism industries (Crouch, 1995; Lim, 1997; Song & Li, 2008; Witt & Witt, 1995). Tourism demand analysis is important and integral part of the tourism industry.

For instance, tourism demand analysis can provide long and short term data that are valuable for anyone who has responsibility for formulating policies or promoting tourism industries (service and product). Prediction of the tourism can use to investigate the role of the tourism on the local industries and

domestic welfare. This analysis can help the government to and private investor to have a realistic overview of the tourism industries and can lead to the efficient allocation of resources.

Studying on the tourism demand research reveals the many different factors are effective in the tourism industry, such as the relative exchange rate, prices, income and safety. These parameters, as well as other factors that don't mention above have been used in many demand researches (Crouch, 1995; Johnson & Ashworth, 1990; Uysal, 1998). Many researchers have tried to formulate of these parameters and used this formulation for prediction of the tourism demand. The log – linear function is one of the functions that used widely for fitting to the tourism demand (Crouch, 1995; Johnson & Ashworth, 1990). This kind of formulation is used for study and examined the responses of different parameters on the tourism demand, although this formulation should be calibrated with local data for each country. Recently, long the conventional regression and formulation, artificial neural network are presented as a reliable method for prediction of tourism demand. ANN is used widely for a data set that have complicated nature (Kon & Turner, 2005; Law, 2001; Pai & Hong, 2005).

In this study, Iran is selected as a case study for tourism demand prediction. And also two different methods of machine learning (multi-layer perceptron ANN and supporting vector machine) are selected for analysis of this data.

## ***Material Method***

### ***Selecting the Effective Parameters on the Tourism Demand***

The tourism demand is a function of the different factors, as previously mentioned. This study has tried to localize these factors for the condition of Iran. Factors such as exchange rate, relative price rate, amount expenses for advertisements, political issues and safety are most important parameters that can be effective on the tourism demand.

Unfortunately, no sources have found for the amount of expenses that spend on advertisement in Iran, for this reason the effect of this parameter is neglected in this analysis. In addition, same problem is found for the political and safety parameters. Evaluation of the safety in small scale has been done by many researchers (Rezania & Sadatiseyedmahalleh, 2015) but in big scale, is rarely reported. To overcome with this problem these two factors are combined together that called “political safety”. A fussy description is selected for this parameter. This parameter is set rang of the 0 and 1. For most of the years, this parameter is set as one, for year of the 1979 (Islamic revolution in Iran is set as 0.5), four years between 1980 and 1988 (Iran and Iraq imposed war), it is set as the zero. And it is set at 0.75 for 2009 (Iran presidential election protest).

In this study, for both methods Iran currency exchange, political safety, price rate are selected as inputs and tourism demand is set as output of the network.

## ***Artificial Neural Networks***

Artificial neural networks have been used to create general relationships between variables in a given issue. Artificial neural networks are considered as a form of artificial intelligence and are somewhat capable to simulate the human brain and nervous system. As shown in Figure 1, artificial neural networks are made of a number of processing elements or nodes or neurons. The most widely used artificial neural networks in engineering purposes are known as feedforward multilayer perceptron network with back propagation algorithms with an input layer, an output layer, and one or more hidden layers (Fyfe, 2000; G.Dreyfus, n.d.; Gershenson, n.d.). Although there is no connection between the neurons in the same

layer, each neuron is fully connected with the rest of neurons in the next layer (Vakili et al., 2015). The output of each node could be defined by Equation 1.

$$a = f(\sum_{i=1}^n P_i w_{i,j} + b_j) \quad (1)$$

Where  $a$  is the output of node  $j$ ,  $P_i$  is the input from  $i$  th node,  $W_{i,j}$  is the connection weight between  $j$  th node of the layer and  $i$  th node of the previous layer,  $b_j$  is the bias at the  $j$  th node, and  $f$  is the transfer function.

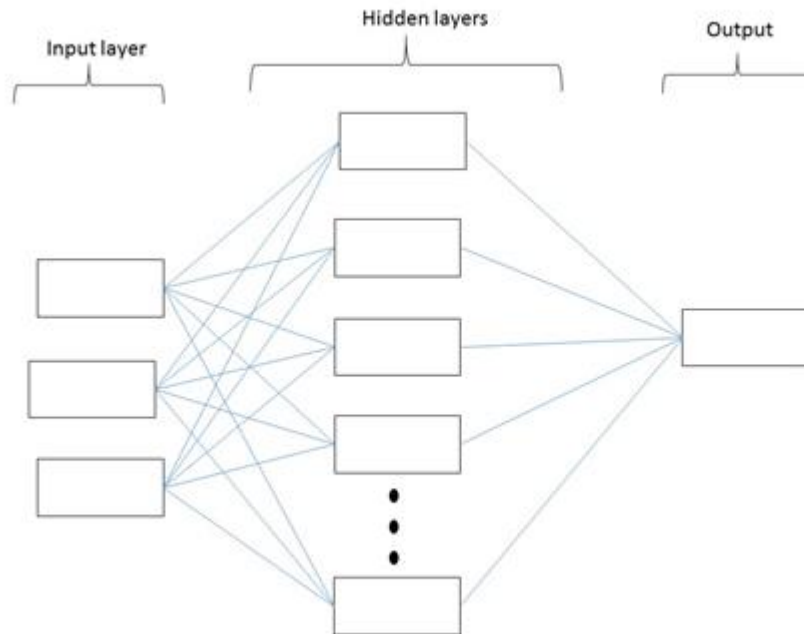


Fig.1 Schematic structure of the ANN

Sigmoid functions is usually used as transfer function of the each node, however, other functions such as hyperbolic tangent functions and linear functions could also be used, transfer function has direct effect on the performance of the ANN. Before selecting the transfer function good knowledge about the nature and behaviour of the variables is necessary.

The number of nodes in the input and the output layers, respectively represent the number of input and output variables. The neural network models are normally developed by dividing the available datasets into 3 subsets for training, validation, and test. In this paper, ANN is initially trained by using 80% of the data while 10% of the data is utilized in the validation process of ANN. This process has occurred during the training process to minimize over fitting ANN. The performance of the trained model is tested by the remaining 10% of the data. Thus, the training data is used for constructing the model and the validation and; testing data are used for controlling the accuracy of the developed model. The difference between an experimental value (target) and the network prediction value is defined as a system error.

### Support Vector Regression

The Support Vector Machine (SVM) (Boser, Guyon, & Vapnik, 1992; Schölkopf & Burges, 1999) is a method of the learning that sourced from the theoretical base of the statistical learning theory and structural risk minimization. This method is utilized in both regression and classification or analysis. The SVM method initially transferred nonlinearly the input space into the higher feature space. In order to learn nonlinear relations with a linear machine, it is required to select a set of nonlinear feature and to express the data in the new representation. This is equivalent to applying a fixed nonlinear mapping of the data to a feature space which the linear machine can be used in. Nonlinear regression problems in input space are changed to the linear regression problems in feature space. When SVM is used for regression problems, it is called as support vector regression (SVR). SVR is equivalent to maximizing the margin between training examples and the regression function. It is an alternative to other neural networks with training methods that optimize cost functions such as the mean square error.

The SVR (Vapnik, Golowich, & Smola, 1997) model is given N training data  $\{(x_i, y_i)\}_{i=1}^n \in R^m \times R$  where  $x_i$  is the input vector to the SVR and  $y_i$  is the target value, from which it learns the input–output relationship. The equation below shows this relationship (Kecman, 2001).

$$y = f(x) = \sum_{i=1}^N w_i \phi_i(x) + b = w^T \phi(x) + b \quad (2)$$

Where the function  $\Phi_i(x)$  is called the feature that is nonlinearly mapped from the input space to the output space. This equation is nonlinear regression model because the resulting hyper-surface is a nonlinear surface hanging over the m dimension input space. This equation can change into a linear regression model when the input vectors x are mapped into vectors  $\phi(x)$  of a high dimensional kernel-induced feature space. The nonlinear function is learned by a linear learning machine where the learning algorithm minimizes a convex function. The convex functional is expressed as the following regularized risk function, and support vector weight (w) and bias (b) can calculate by minimizing the risk function as follow (Na, Kim, Lim, & Kang, 2008).

$$R(w) = \frac{1}{2} w^T w + \lambda \sum_{i=1}^N |y_i - f(x)_i|_{\varepsilon} \quad (3)$$

$$|y_i - f(x)_i|_{\varepsilon} = \begin{cases} 0 & \text{if } |y_i - f(x)_i| < \varepsilon \\ |y_i - f(x)_i| - \varepsilon & \text{otherwise} \end{cases} \quad (4)$$

$\lambda$  is regularization parameter, which determines the tradeoff between the approximation error and the weight vector norm. An increasing the constant  $\lambda$  penalizes larger errors which decreased the approximation error also this can be happened when the weight vector norm is increased. Increasing the weight vector norm does not always lead to better generalization of the SVR model. The constants  $\lambda$  and  $\varepsilon$  are user-defined parameters and  $|y_i - f(x)_i|_{\varepsilon}$  is called the  $\varepsilon$ -insensitive loss function. The loss equals zero if the prediction  $f(x)$  is within an error level  $\varepsilon$ , and for all other prediction points bigger than the error level  $\varepsilon$ , the loss is equal to the value of the difference between the prediction and the error level (Fig. 4). Increasing the insensitivity zone (error level  $\varepsilon$ ) can reduce the accuracy of the prediction,

however, this increase can lead to smooth effects on modeling of the data sets that include the high level of noise (Na et al., 2008; Singh & Gupta, 2010; Vapnik et al., 1997).

For training the SVR network insensitive loss function ( $\epsilon$ ) = 0.01, kernel type= radial basis function (RBF), kernel parameter= 30 are given as the initial parameter. During the training process after adding the each sample to the function, Karush–Kuhn–Tucker (KKT) conditions was checked. If KKT condition is satisfied the sample was added, otherwise the sample was stabilized using the stabilization technique. In the stabilization technique SVR parameters (c and  $\epsilon$ ) will be changed to achieve the optimize the values.

## Result

We have analyzed the tourism demand series by considering the certain possible patterns. The purpose of analyzing the time series is to introduce the dynamic relationships among the tourism demand series and improve the accuracy of forecasts by extracting the additional information available from the associated data series at hand in the tourism demand forecasting.

The objective of the paper was the forecasting performance of different artificial neural network models, extending to tourist demand forecasting the results of previous research. With using these three different sets of models we obtained forecasts for the number of tourists from all visitor markets to Iran.

For prediction of the tourist demand two different machine learning methods are used in this paper. In the multilayer perceptron ANN different nodes number examined in the hidden layer (from 1-10) and in each try the value of the  $R^2$  was recorded and compared. The maximum value of the  $R^2$  was observed in the neural network with the 10 nodes in the hidden layer. The trend of the  $R^2$  changing is shown in figure 2. The trained network with the 10 nodes is selected for further analysis. The relation between the real tourism demand and neural network prediction is shown in the figure 3. As obvious in the figure 3, the predicted amount by ANN which is less than the 1million has great agreement with actual tourism demand.

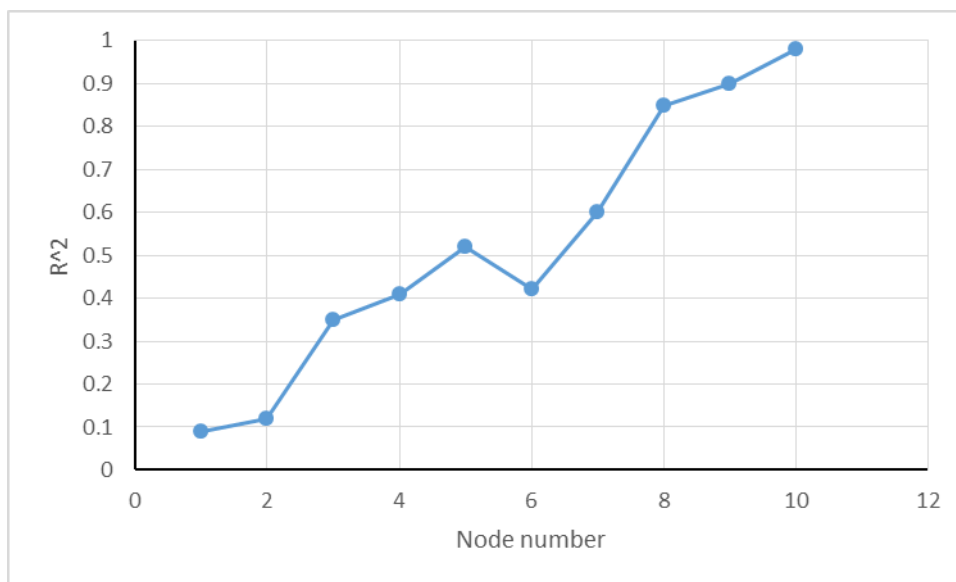


Fig. 2 Amount of the  $R^2$  for different number of the node in the hidden layer

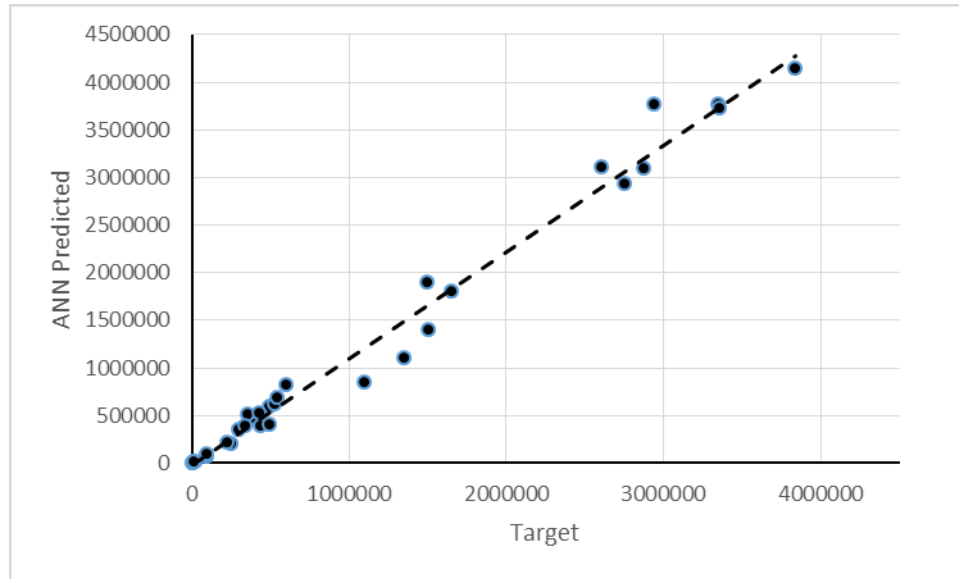


Fig. 3 Comparing the prediction of the ANN with actual tourism demand

The SVR is also trained with same data the prediction of the SVR was compared with predictions of the multilayer perceptron. This comparing is shown in figure 4. In all the years, the SVR prediction was better than the multi-layer perceptron, the value of the  $R^2$  in this method was around the 0.99 that is higher than the  $R^2$  value in the multilayer perceptron with 10 nodes in the hidden layer (0.98). The SVR was followed the same trend with actual data. However, the SVR prediction at the most of the time is less than the actual demand, otherwise the multi-layer perceptron ANN has higher prediction.

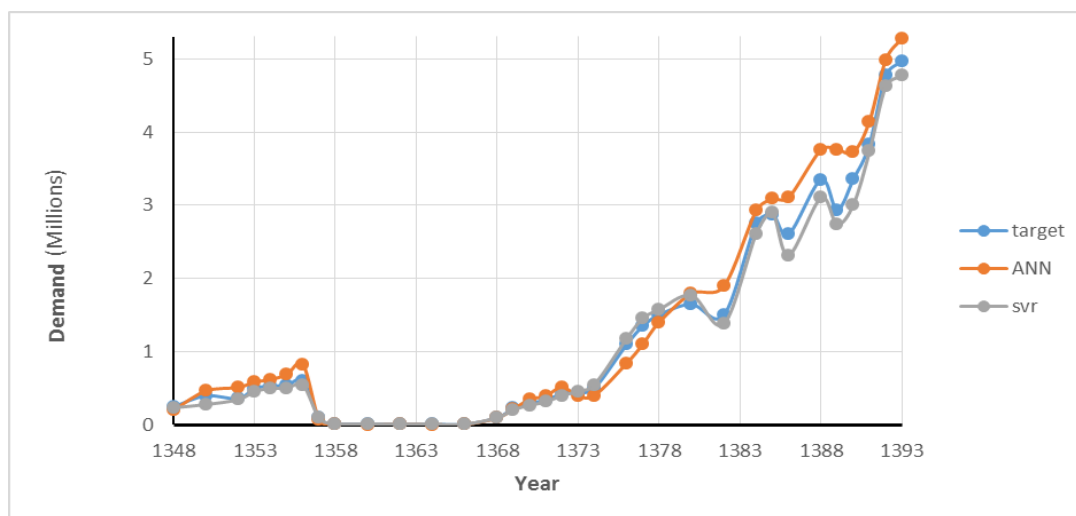


Fig. 4 Comparing the actual tourism demand with the SVR and ANN prediction

The error of prediction for both of the models were shown in figure 5. The fluctuation of the multilayer perceptron is more than the SVR error prediction.

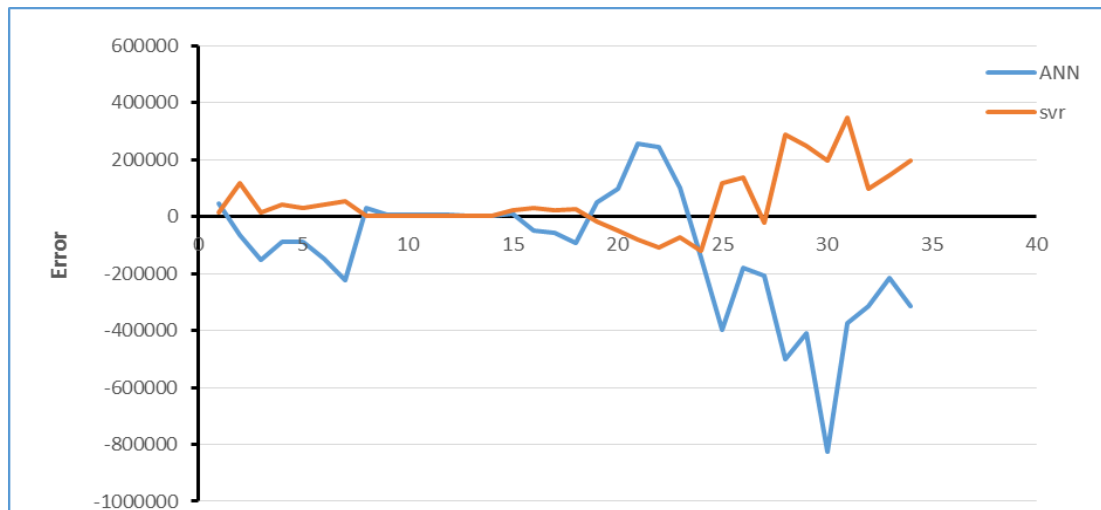


Fig. 5 The error of prediction for both of the models

## References

- Boser, B., Guyon, I., & Vapnik, V. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*.
- Crouch, G. I. (1995). A meta-analysis of tourism demand. *Annals of Tourism Research*, 22(1), 103–118. [http://doi.org/10.1016/0160-7383\(94\)00054-V](http://doi.org/10.1016/0160-7383(94)00054-V)
- Fyfe, C. (2000). *Artificial Neural Networks and Information Theory* (1.2 ed.).
- G.Dreyfus. (n.d.). *Neural Networks Methodology and Applications*. Springer-Verlag Berlin Heidelberg 2005.
- Gershenson, C. (n.d.). *Artificial Neural Networks for Beginners*, 1–8.
- Johnson, P., & Ashworth, J. (1990). *Modelling tourism demand: A summary review*. *Leisure Studies*. Retrieved from <http://www.tandfonline.com/doi/pdf/10.1080/02614369000390131>
- Kecman, V. (2001). *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models*. MIT press.
- Kon, S., & Turner, L. (2005). Neural network forecasting of tourism demand. *Tourism Economics*. Retrieved from <http://www.ingentaconnect.com/content/ip/tec/2005/00000011/00000003/art00001>
- Law, R. (2001). The impact of the Asian financial crisis on Japanese demand for travel to Hong Kong: A study of various forecasting techniques. *Journal of Travel & Tourism Marketing*. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/10548400109511558>.
- Lim, C. (1997). Review of international tourism demand models. *Annals of Tourism Research*, 24(4), 835–849. [http://doi.org/10.1016/S0160-7383\(97\)00049-2](http://doi.org/10.1016/S0160-7383(97)00049-2).

- Na, M. G., Kim, J. W., Lim, D. H., & Kang, Y. J. (2008). Residual stress prediction of dissimilar metals welding at NPPs using support vector regression. *Nuclear Engineering and Design*, 238, 1503–1510.
- Pai, P., & Hong, W. (2005). An improved neural network model in forecasting arrivals. *Annals of Tourism Research*. Retrieved from [https://scholar.google.com/scholar?q=An+improved+neural+network+model+in+forecasting+arrivals&btnG=&hl=en&as\\_sdt=0%2C5#0](https://scholar.google.com/scholar?q=An+improved+neural+network+model+in+forecasting+arrivals&btnG=&hl=en&as_sdt=0%2C5#0).
- Rezania, M., & Sadatiseyedmahalleh, S. (2015). Measuring the Safety by Using the Fuzzy Logic Methods in Universiti Sains Malaysia. *REVISTA CIENTIFICA*. Retrieved from <http://cientificaonline.com/index.php/CIENTIFICA/article/view/724>.
- Schölkopf, B., & Burges, C. (1999). *Advances in kernel methods: support vector learning*. MIT press.
- Singh, S. K., & Gupta, A. K. (2010). Application of support vector regression in predicting thickness strains in hydro-mechanical deep drawing and comparison with ANN and FEM. *CIRP Journal of Manufacturing Science and Technology*, 3(1), 66–72.
- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management*, 29(2), 203–220. <http://doi.org/10.1016/j.tourman.2007.07.016>.
- Uysal, M. (1998). The determinants of tourism demand. ... *Economic Geography of the Tourist Industry: A Supply- ...* Retrieved from <https://books.google.com/books?hl=en&lr=&id=fntlPooU2wUC&oi=fnd&pg=PA79&dq=tourism+demand+Uysal+1998&ots=U9VYN1WCGC&sig=hOC1cC1MBpAROYcvdKaMwNa0fy4>
- Vakili, A. H., Davoodi, S., Arab, A., Researcher, Y., Club, E., Branch, E., ... Tebal, N. (2015). Use of Artificial Neural Network in Predicting Permeability of Dispersive Clay Treated With Lime and Pozzolan. *Ijsres*, 3(1), 23–37.
- Vapnik, V., Golowich, S. E., & Smola, A. (1997). Support vector method for function approximation, regression estimation, and signal processing, 281–287.
- Witt, S. F., & Witt, C. A. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting*, 11(3), 447–475. [http://doi.org/10.1016/0169-2070\(95\)00591-7](http://doi.org/10.1016/0169-2070(95)00591-7).

## Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).