Sosyoekonomi

2022, Vol. 30(53), 11-26

RESEARCH ARTICLE

ISSN: 1305-5577
DOI: 10.17233/sosyoekonomi.2022.03.01
Date Submitted: 14.02.2021
Date Revised: 06.04.2022
Date Accepted: 10.04.2022

Estimating the Changes in the Number of Visitors on the Websites of the Tourism Agencies in the COVID-19 Process by Machine Learning Methods

Mehmet KAYAKUŞ *(https://orcid.org/0000-0003-0394-5862), Akdeniz University, Turkey; mehmetkayakus@akdeniz.edu.tr*

COVID-19 Sürecinde Turizm Acentelerinin Web Sitelerindeki Ziyaretçi Sayısındaki Değişimin Makine Öğrenmesi Yöntemleriyle Tahmin Edilmesi

Abstract

In this study, the number of visitors of five different tourism agencies was tried to be estimated by machine learning method using the number of cases and deaths in Europe during COVID-19. Artificial neural network (ANN), support vector regression (SVR), and multiple linear regression (MLR) were used as machine learning models. A model consisting of two independent variables and one dependent variable was created. According to the analysis made according to three different techniques, the most successful results; According to R2, it was seen that ANN, DVR, and MDR, and according to other statistical methods, ANN, MDR, and DVR, respectively.

Keywords : COVID-19 Pandemic, Tourism Agency, Website, Number of Visitors, Machine Learning.

JEL Classification Codes : I10, C53, M30, S46.

Öz

Bu çalışmada COVID-19 süresince Avrupa’daki vaka ve ölüm sayıları bilgileri kullanılarak beş farklı turizm acentesinin ziyaretçi sayısı makine öğrenmesi yöntemiyle tahmin edilmeye çalışılmıştır. Yöntem olarak makine öğrenmesi modellerinden yapay sinir ağları (YSA), destek vektör regresyonu (DVR) ve çoklu doğrusal regresyon (ÇDR) kullanılmıştır. İki bağımsız değişken ve bir bağımlı değişkenden oluşan model oluşturulmuştur. Üç farklı tekniğine göre yapılan analize göre en başarılı sonuçların; R2’ye göre YSA, DVR ve ÇDR, diğer istatiksel yöntemlere göre de sırasıyla YSA, ÇDR ve DVR olduğu görülmüştür.

Anahtar Sözcükler : COVID-19 Salgını, Turizm Acentesi, Web Sitesi, Ziyaretçi Sayısı, Makine Öğrenmesi.

1. Introduction

Undoubtedly, the Internet is an effective distribution channel for consumers (Wan, 2002). The Internet has changed individuals, companies, and organisations' daily lives and information search methods. Also, the Internet's validity as a tool for advertisement and marketing has been proven (Kasavana et al., 1998).

The Internet's infinite opportunities have made it one of the crucial issues of today's tourism perspectives. The 24/7 service of the Internet available worldwide is a great advantage for tourism (Kotler et al., 2017). Moreover, other advantages brought by the Internet are its speed and relatively low cost, enabling independent operators to have access to world markets and offering the opportunity of distribution (O'Connor, 1999). The Internet provides an ever-increasing travel and tourism resource that includes a broader user base and potential users. In general, the Internet can be used in tourism in two ways: firstly, as a data source where the user has access to sources to obtain information, and secondly, to facilitate marketing and business transactions (Walle, 1996). The Internet has changed the planning, controlling, running, and integration of most business activities in the accommodation industry, including marketing activities (Kasavana et al., 1998).

A website is a form of direct contact between the travel agency and the customers, and therefore, it provides direct sales and communication channels (Cunliffe, 2000). Web-based marketing offers five advantages: Global coverage and access, convenience/fast transaction process, efficiency and flexibility in information processing, data-based management and relationship skills, and lower cost of sales and distribution (Rosenbloom et al., 1999). Therefore, tourism agencies that wish to increase sales and establish interactive communication with the customers develop their websites and mainly carry out their activities in this manner (Pulliam, 1999).

According to a report prepared by the World Tourism Organization Business Council (2001), approximately 70% of travellers used online travel agencies and called or visited tourism offices less. Besides, about 70% of the travellers who used the Internet reported that they directly called airline companies less. Approximately 60% stated that they called car rental companies and accommodation companies directly less and ordered travel brochures less by phone (So & Morrison, 2004).

The world faces a new coronavirus outbreak spread to 260 countries or regions (World Health Organization, 2020). The World Health Organization (WHO) announced an internationally crucial emergency for health epidemically on January 30, 2020. Also, it declared a pandemic on March 11, 2020. COVID-19 is not limited to losing human lives. It also has short- and long-term social, economic and political implications (Farzanegan et al., 2020).

While COVID-19 is terrorising the whole world, many countries and regions introduced travel restrictions and border closures to stop the spread of the virus. It was also claimed that the appearance of infectious diseases is a natural result of global tourism and mobility (Richter, 2003). There are many studies on crises affecting tourism (Aliperti et al., 2019; Cró & Martins, 2017; Sio-Chong & So, 2020; Wang, 2009). The COVID-19 pandemic has simultaneously affected a few countries' economic growth and welfare, along with severe consequences for international tourism (Gössling et al., 2020).

There are numerous studies in which artificial intelligence techniques were used to estimate tourism demand. Law et al. employed deep learning methods SVR and ANN models for tourism demand estimation (Law et al., 2019). Andrew et al. conducted a study on estimating hotel occupancy rates through time series models (Andrew et al., 1990).

In the study they conducted, Faranegan et al. found a significant and positive relationship between the cumulative cases and death numbers resulting from COVID-19 and the previous tourism records by using the data obtained from 90 countries and performing a multiple regression (Farzanegan et al., 2020).

Polyzos et al. designed a machine learning method to estimate the effect of the new coronavirus outbreak on the tourists from China arriving in the United States of America and Australia. They used data from the SARS outbreak in 2003 to train a model called Long Short-Term Memory (LSTM) based on artificial neural networks. The LSTM was calibrated for the details of today's pandemic (lockdowns, flight prohibition, etc.). As a result of the study, it was revealed that it would take 6 to 12 months to return to pre-pandemic levels and that this situation would have adverse and significant effects not only on the tourism industry but also on other sectors that are in interaction with it (Polyzos et al., 2020).

This study estimated the daily number of visitors visiting tourism agencies' websites using the current number of cases and deaths from the daily COVID-19 statistical data in Europe. The tourism agencies selected in the study are booking.com, tripadvisor.com, agoda.com, hotels.com and kayak.com. The websites of the five most visited tourism agencies were selected. The most important factor in his selection was the number of daily visitors. The estimation method, support vector machine, multiple linear regression and artificial neural networks method were used among the machine learning methods.

2. Data and Methodology

The study used the daily case and death numbers in 54 European countries retrieved from An Agency of the European Union (EU Agency) as the independent variable (European Union Agency, 2020). Although many tourism agencies have websites, five sites with the highest number of visitors were selected in the study. The main factor in the agencies chosen is the number of daily visitors and interactions. The website's daily visitor information was obtained from websiteiq.com (websiteiq.com, 2020). Multiple linear regression and artificial neural network methods were used among machine learning methods to estimate the number of visitor support vector machines, and the results were comparatively analysed. For the application and analyses of artificial intelligence techniques, the Knime program was chosen.

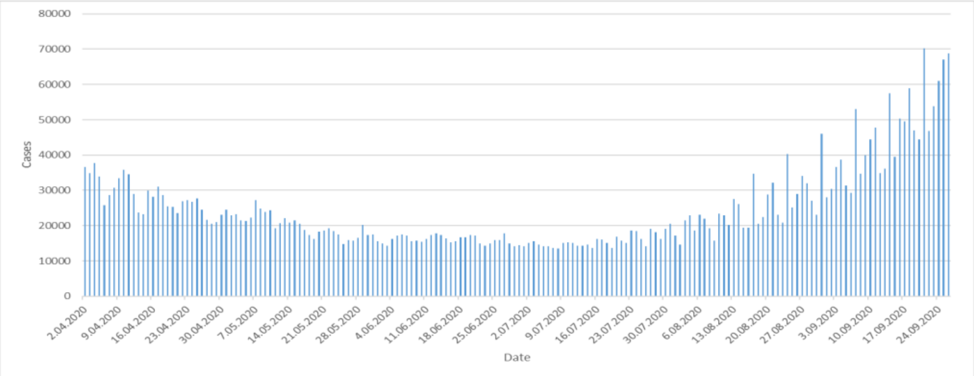
2.1. Data Set

The study investigated how the COVID-19 data affected the number of visitors visiting the websites of tourism agencies, and by using these data, the number of visitors was estimated through machine learning methods. The study estimated the number of visitors to 5 big tourism agencies' websites using 178-day COVID-19 data in Europe between 02.04.2020 and 26.09.2020.

In the study, the number of data was increased by dividing the same data samples into different groupings using the cross-validation method. In this method, initially, the data are randomly separated as testing and training data. While the training data are used in the establishment stage of the model, the testing data are not used in modelling, and the model's validity is tested over this new data set (Bishop, 1995; Temel et al., 2012).

As the study's independent variable, the numbers of cases and deaths obtained from "An Agency of the European Union" were used (EU Agency, 2020, 2020). In Figure 1, the change in the number of daily cases caused by European COVID-19.

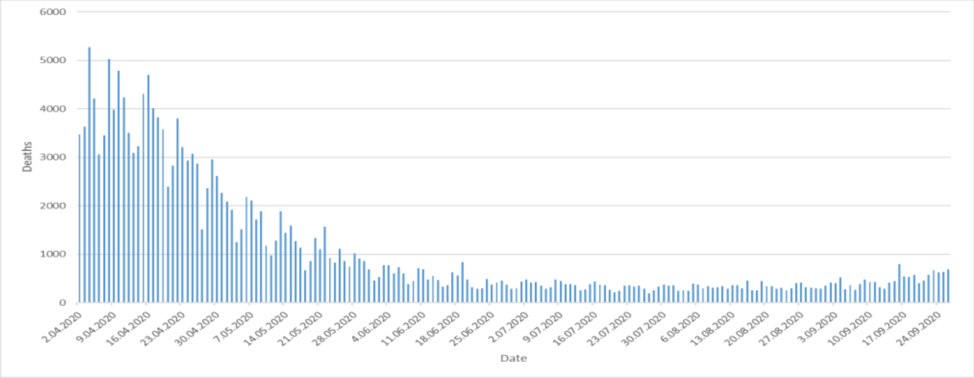
Figure: 1
Number of Daily Cases in Europe



(EU Agency, 2020)

Figure 2 shows the change in the number of daily deaths caused by COVID-19 in Europe graphically.

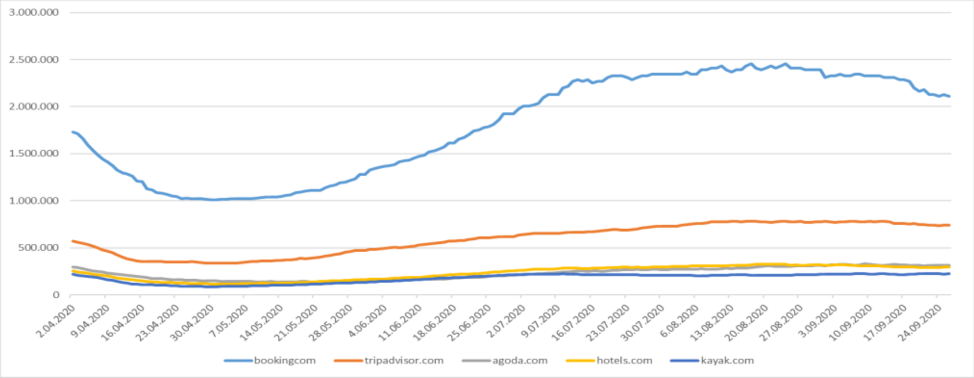
Figure: 2
Change in Daily Deaths Related to COVID-19



(EU Agency, 2020)

As the dependent variable, five big tourism agencies whose daily visitor numbers are over 170.000 on average were determined. The website’s average daily website visitors were 1.799.132 visitors for booking.com, 591.288 visitors for tripadvisor.com, 227.451 visitors for agoda.com, 233.206 visitors for hotels.com, and 176.193 visitors for kayak.com (websiteiq.com, 2020). Figure 3 shows the change in the five tourism agencies' daily visitor numbers during the study period.

Figure: 3
Five Tourism Agencies' Daily Visitor Numbers



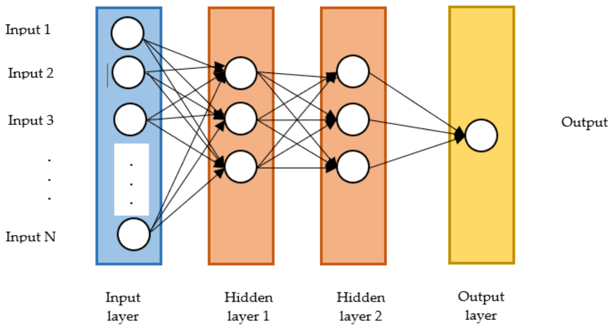
(websiteiq.com, 2020)

2.2. Artificial Neural Network

The artificial neural network (ANN) method was first used in tourism demand forecasting in the 1990s. Artificial neural networks are used in many applications in tourism management and marketing, such as estimating tourists' preference activities and behaviours, estimating the demand, and analysing guest loyalty (Bloom, 2005). It is seen that ANN mostly outpaced the multiple regression models and time series in estimating issues related to tourism (Law, 2000; Law & Au, 1999; Pattie & Snyder, 1996).

Artificial neural networks are named after the nerve cell networks in the brain. Artificial neural networks are designed to bring together the neurons' basic properties in the brain and process the data similarly to a human's brain. An artificial neural network is a mathematical function that calculates the output variable depending on the input variable (Kim et al., 2003). In another definition, the structure and functioning of artificial neural networks can be defined as data processing systems inspired by biological neural. (Palmer et al., 2006). An ANN is generally made up of numerous processing elements, known as neurons, organised into layers, as seen in Figure 4.

Figure: 4
Artificial Neural Networks



The first neuron, called the input layer, directly depends on the input vector. The purpose of the input layer is to provide data from the outside world to the network. The last layer, called the output layer, produces an output signal. Hidden layers are between the input layer and the output layer (Lippmann, 1988). Hidden layers create models for nonlinear relationships between inputs and outputs (Klimasauskas, 1992). One of the essential points to be considered in the architecture of ANN is the choice of the hidden layer structure. Although there have been efforts to develop guidelines for deciding on the most suitable hidden layer and the number of nodes, there is no standard procedure, and it is still a matter of trial and error (Hill & Remus, 1994). Each neuron is connected to other neurons by connections, each with a numerical value known as “weight” (Palmer et al., 2006).

Each neuron receives inputs (x_1, x_2, …,x_n) attached with a weight indicating the connectivity power of that neuron for each connection (Hill & Remus, 1994). Input variables are multiplied by the corresponding weight values of the neuron connection. A bias (b_i) can be expressed as a connection weight type that has a constant value different from zero that is added to the total of the inputs (Murat et al., 2014). In line with this information, the neural network model is established as:

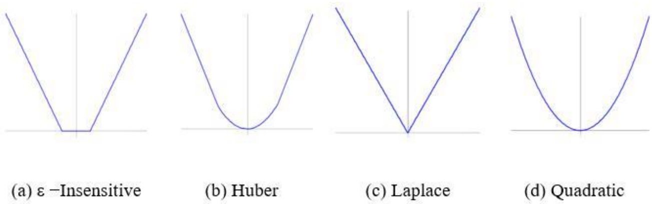
$$u_i = \sum_{j=1}^{n} w_{ij} x_j + b_i \tag{1}$$

ANNs are based on two types of architecture: feed-forward and backpropagation neural networks. The feed-forward network structure refers to a one-way flow and data processing, which allows each node to receive information only from the previous neuron. On the other hand, the backpropagation neural network is a two-way data processing method. In this method, each node receives information from the previous layer and allows the feedback to be transmitted to the next layers (Tsaur et al., 2002).

2.3. Support Vector Regression

Support Vector Machines (SVM) is a machine learning method used for classification and regression developed by Vapnik (Vapnik, 2013). Support vector machines, an algorithm based on statistical learning methods, are designed to provide a solution for limited quadratic programming and a global optimisation problem (Chen et al., 2015). It can also be applied to regression problems by adding an alternative loss function to SVMs (Smola & Schölkopf, 1998). The loss function should be changed to include a margin measurement. Figure 5 shows lost parts.

Figure: 5
Loss Functions



(Gunn, 1998)

The quadratic loss function corresponds to the traditional least-squares error criterion. The Laplacian loss function is less susceptible to outliers than the quadratic loss function. If the data's primary distribution is unknown, Huber is suggested as a proper loss function with the most suitable properties (Gunn, 1998).

Support vector regression has started to be used as a powerful approach to estimate and solve linear and non-linear systems with the appearance of Vapnik's ε-insensitive loss function (Lijuan & Guohua, 2016).

Given a data set as in Equation 2:

$$D = \{(x^1, y^1), \ldots, (x^l, y^l)\}, x \in R^n, y \in R \tag{2}$$

A linear support vector regression:

$$y = f(x) = < w, x > +b \tag{3}$$

A functional minimum provides optimal regression function:

$$\Phi(w, \xi) = \frac{1}{2}\left||w|\right|^2 + C\sum_i(\xi_i^- + \xi_i^+) \tag{4}$$

Where, the value of C is predetermined. ξ^- and ξ^+ represent the upper and lower limitations on the system's outputs.

Like classification problems, a nonlinear model is often used to model data. A non-linear mapping can be used to map the data into a high dimensional feature space where linear regression is performed. A non-linear SVR solution which uses X-Insensitive (ε-insensitive) loss function is given by,

$$\underbrace{max}_{\alpha,\alpha^*} W(\alpha, \alpha^*) = \underbrace{max}_{\alpha,\alpha^*} \sum_{i=1}^{l}(y_i - \varepsilon) - \alpha_i(y_i + \varepsilon) - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i^* - \alpha_i)\left(\alpha_j^* - \alpha_j\right)K\left(x_i, x_j\right) \tag{5}$$

$$0 \le \alpha_i, \alpha_i^* \le C, i = 1, \ldots, l \tag{6}$$

$$\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \tag{7}$$

Solving Equation 5 with Equations 6-7 determines Lagrange multipliers and α_i, α_i*, and the regression function:

$$f(x) = \sum_{SVs}(\bar{\alpha}_i - \bar{\alpha}_i^*)K(x_i, x) + \bar{b} \tag{8}$$

where,

$$\langle \bar{w}, x \rangle = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i, x_j) \tag{9}$$

$$\bar{b} = -\frac{1}{2}\sum_{i=1}^{l}(\alpha_i - \alpha_i^*)(K(x_i, x_r) + K((x_i, x_s)) \tag{10}$$

If b contains is in the Kernel function and the Kernel contained a bias term, equality constrained can be left out, and the regression function:

$$f(x) = \sum_{i=1}^{l}(\bar{\alpha}_i - \bar{\alpha}_i^*)K(x_i, x) \tag{11}$$

Optimisation principles for other loss functions can similarly be achieved by replacing the internal multiplication with a kernel function.

2.4. Multiple Linear Regression

A mathematical model is used to explain the relationship between two (simple regression) or more variables (multiple regression), and this model is called the regression model. The dependent variable is sometimes referred to as estimation or response, while independent variables are called estimators (Mata, 2011).

A simple linear regression represents the relationship between the dependent variable y and the independent variable x, as shown in Equation 12.

$$y_i = \beta_0 + \beta_1 x_i + e_i, i = 1,2,3, \ldots, n \tag{12}$$

Using the least-squares method, the best fitting line can be obtained by minimising the sum of the squares of the vertical distance from each data point over the line (Brown, 2009). According to the multiple regression model, a dependent variable correlates with two or more independent variables.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + e_i, i = 1,2,3, \ldots, n \tag{13}$$

β_0 is constant, and when all explanatory variables are 0, it will be the foreseen value of y. In a model with ρ explanatory variables, each explanatory variable will have its β coefficient (Grégoire, 2014).

The least-squares method estimates the regression coefficients in a multiple regression analysis.

2.5. Performance Analysis Methods

To measure the success of the predictions made, the Coefficient of Determination (R2) in Equation 14, Mean Squared Error (MSE) in Equation 15, Root Mean Squared Error (RMSE) in Equation 16 and Mean Absolute Error (MAE) in Equation 17 were used.

$$\mathrm{R}^2 = 1 - \frac{\sum(y_\mathrm{i} - x_i)^2}{\sum(y_\mathrm{i} - y_\mathrm{ave})^2} \tag{14}$$

$$\mathrm{MSE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} (y_\mathrm{i} - x_i)^2 \tag{15}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left| \frac{y_\mathrm{i} - x_i}{\mathrm{y}} \right|^2} \tag{16}$$

$$\mathrm{MAE} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} |y_\mathrm{i} - x_i|}{\mathrm{n}} \tag{17}$$

$$MAPE = \frac{\sum_{i=1}^{n} \left| \frac{y_i - x_i}{y_i} \right|}{n} \cdot 100 \tag{18}$$

where y_i is the observation value and x_i is the predicted value, y_{ave} average of observation values.

MAE and RMSE measure the errors that give a general idea about the difference between the observed and modelled values. MBE shows whether the observed concentrations were overestimated or underestimated. According to these criteria, high R2 and low MAE, MSE, RMSE, and MAPE values identify a model that fits well.

3. Results and Discussion

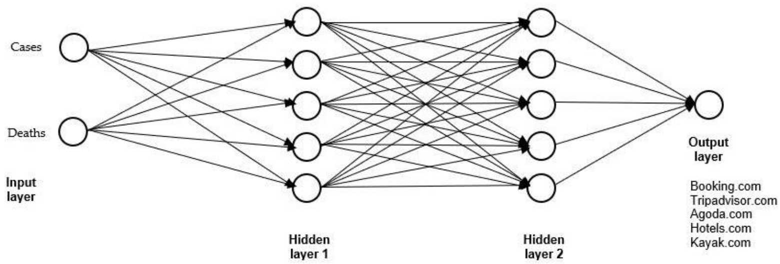
The study used 178-day COVID 19 cases and visitor numbers of data covering the period between April 2 and September 26, 2020, in Europe. The data are normalised between 0 and 1 by transferring them in Excel format into the program and using the decimal scaling method in the first place. In this method, the normalisation process is performed by moving the decimal points of the data. To normalise the data using the decimal scaling method, each data is divided by the maximum absolute value of the data. vi of the data is normalised as vi' by Equation 10:

$$v_i' = \frac{v_i}{10^j} \tag{19}$$

Here, j is the smallest integer, maxi(vi’|) being <1.

In an artificial neural network model, 70% of the data are chosen randomly for testing purposes and 30% for training purposes, and they are transferred to the designed network. An ANN is modelled to have 10.000 iterations, 2 hidden layers, and 5 neurons in each hidden layer. While creating the model, the parameters were determined to obtain the highest performance for each dependent variable. The ANN model created is seen in Figure 6.

Figure: 6
The ANN Models



Statistical results of tests performed with artificial neural networks are shown in Table 1.

Table: 1
Performances of ANN Method

Independent variables	R2	MSE	RMSE	MAE	MAPE
booking.com	0,832	0	0,022	0,015	0,089
tripadvisor.com	0,870	0,003	0,058	0,037	0,073
agoda.com	0,769	0,001	0,031	0,021	0,102
hotels.com	0,863	0,001	0,027	0,019	0,091
kayak.com	0,705	0,001	0,025	0,190	0,097

The fact that the R2 value is 1, which shows how well the data fit into a linear curve, indicates that the test data provides a linear curve. The RMSE value used to find the distance between predicted and actual values is zero, which means that the model has made no mistakes. Therefore, RMSE is desired to be close to zero. MSE shows how close a regression curve is to a few points. Estimate models close to zero can be said to be more successful. MAE is a magnitude that indicates the difference between two continuous variables. In other words, it shows the average vertical distance between the actual value and the line that best fits the data. Low MAE indicates the success of the forecast model (Kayakuş, 2021). Forecast models with MAPE values below 10% are considered models with a “high accuracy” or “very good” accuracy rating (Lewis, 1982; Witt & Witt, 1992). According to these criteria, high R2 and lowest RMSE, MSE and MAE values determine the most successful model.

According to the ANN method, the R2 value in estimating the visitor numbers of tourism agencies was determined to be 0,870 maximum and 0,705 minimum. The results can generally be considered successful. The MSE values being very close to 0 in the study shows the ideal number. The RMSE values are desired to be close to 0, and in the study, it was determined to be in the range of 0,022 - 0,058, which is very close to the desired value. For MAE and MAPE, the lowest values are accepted as ideal values. In the study, MAE took values between 0,015 and 0,190, and MAPE took values between 0,073 and 0,102, which are close to the ideal values.

In the MLR method, 2 independent variables were used in the input, and 178-day data were used to calculate each independent variable's value. Table 2 shows the performances of each independent variable according to the MLR method.

Table: 2
Performances of MLR Method

Independent variables	R2	MSE	RMSE	MAE	MAPE
booking.com	0,436	0,002	0,040	0,036	0,233
tripadvisor.com	0,55	0,012	0,109	0,092	0,180
agoda.com	0,273	0,003	0,054	0,044	0,209
hotels.com	0,426	0,002	0,049	0,044	0,219
kayak.com	0,416	0,003	0,056	0,05	0,260

According to the MLR method, the R2 value in the visitors' estimation of the websites of the tourism agencies was determined as 0,55, the highest, and 0,273, the lowest. The results were obtained even under the desired values. The MSE values are acceptable in the study as they were close to 0. The RMSE value is expected to be close to 0. Accordingly, in the study, it was within the range of 0,040 and 0,109, which is relative to the desired value. MAE took values between 0,015 and 0,190 in the survey, while MAPE was determined between 0,180 and 0,260. The lowest values are accepted as ideal values for MAE and MAPE. It was seen in the study that these values were close to the desired values.

As the SVR method, the non-linear method was used in the study. It gave the most successful result because of the tests performed; Radial Basis Function (RBF) was preferred as the Kernel function. Different combinations were tried for C and ε, which are SVR parameters; consequently, as the best performance was obtained with ε 0,01 and C 500, they were decided to be used. In Table 3, the performances of each independent variable according to the SVR method are presented.

Table: 3
Performances of the SVR Method

Independent variables	R2	MSE	RMSE	MAE	MAPE
bookingcom	0,910	0,107	0,328	0,213	0,296
tripadvisorcom	0,654	0,384	0,619	0,366	0,613
agoda.com	0,449	0,608	0,780	0,418	1,317
hotels.com	0,719	0,332	0,576	0,343	1,854
kayak.com	0,599	0,469	0,685	0,411	1,165

According to the SVR method, in estimating the number of visitors to the tourism agencies' websites, the R2 value was determined as 0,910 maximum and 0,449 minimum. The results can vary according to the tourism agencies' data, and some results yield the desired and undesired values. As MSE values are expected to be close to 0, some values obtained in the study were far from the desired values. The RMSE value is also wanted to be close to 0, and in the study, it was seen that the RMSE values ranged between 0,328 and 0,780, and some agencies took the desired value. It was determined in the study that MAE ranged between 0,213 and 0,418, and MAPE took values between 0,296 and 1,854. For MAE and MAPE, the lowest values are accepted as the ideal. The study showed that MAE was close to desired values, while MAPE was far from desired ones.

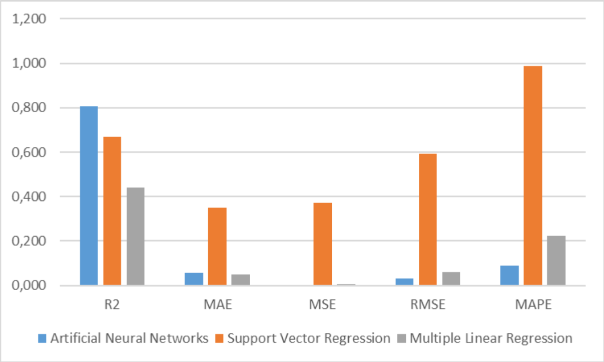
The average statistical results of the tests performed according to three different analysis techniques are presented in Table 4.

Table: 4
The Average Statistical Results of Three Different Techniques

Test techniques	R2	MSE	RMSE	MAE	MAPE
Artificial neural networks	0,808	0,001	0,033	0,056	0,090
Support vector regression	0,670	0,371	0,593	0,349	0,987
Multiple linear regression	0,442	0,004	0,058	0,051	0,222

Figure 7 is the graphic representation of the test results presented in Table 5.

Figure: 7
Average Test Results



When the results of the analyses performed in line with three different artificial intelligence techniques were examined, it was seen that the most successful results were obtained with ANN, SVR, and MLR for R2, and with ANN, MLR, and SVR for MAE, MSE, RMSE, and MAPE, respectively.

4. Conclusion

The COVID-19 pandemic affects people's health and economies and changes their social and cultural habits. In the study, using the cases and death numbers from the COVID-19 data in Europe, their effects on the number of visitors to the tourism agencies' websites and the relevant changes were estimated.

Tourism agencies are intermediaries between consumers, travel companies, and hotels. Especially recently, with the developments in the Internet, tourism agencies have directed their activities towards the worldwide web and increased their investments in this area. Today, many people get information from tourism agencies using internet sites rather than face-to-face or telephone service to make their reservations, purchase and cancel transactions.

This study researched the effect of COVID-19 on people visiting tourism agencies. Three different machine learning methods support vector regression, multiple linear regression and artificial neural networks were employed in the study, and their success rates were compared. The study results show a correlation between the changes in the number of cases and deaths and people visiting tourism agencies' websites.

The study distinguishes itself from other studies on estimating tourism data. It used the COVID-19 data and the information on the visitors of the websites of tourism companies.

With this study, it will be possible to predict how the websites of tourism agencies will be affected if the outbreak lasts for a long time or if a similar outbreak is encountered. Thus, tourism agencies and hotels can be prevented from receiving based on factual information.

References

Aliperti, G. et al. (2019), “Tourism, crisis, disaster: An interdisciplinary approach”, *Annals of Tourism Research*, 79, 102808, 1-5.

Andrew, W.P. et al. (1990), “Forecasting hotel occupancy rates with time series models: An empirical analysis”, *Hospitality Research Journal*, 14(2), 173-182.

Bishop, C.M. (1995), *Neural networks for pattern recognition*, Oxford University Press.

Bloom, J.Z. (2005), “Market segmentation: A neural network application”, *Annals of Tourism Research*, 32(1), 93-111.

Brown, S.H. (2009), “Multiple linear regression analysis: a matrix approach with MATLAB”, *Alabama Journal of Mathematics*, 34, 1-3.

Chen, R. et al. (2015), “Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm”, *Applied Soft Computing*, 26, 435-443.

Cró, S. & A.M. Martins (2017), “Structural breaks in international tourism demand: Are they caused by crises or disasters?” *Tourism Management*, 63, 3-9.

Cunliffe, D. (2000), “Developing usable Web sites-a review and model”, *Internet Research*, 10(4), 295-308.

European Union Agency (2020), <https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-COVID-19-cases-worldwide>, 23.12.2020.

Farzanegan, M.R. et al. (2020), “International tourism and outbreak of coronavirus (COVID-19): A cross-country analysis”, *Journal of Travel Research*, 60(3), 687-692.

Gössling, S. et al. (2020), “Pandemics, tourism and global change: a rapid assessment of COVID-19”, *Journal of Sustainable Tourism*, 29(1), 1-20.

Grégoire, G. (2014), “Multiple linear regression”, *EAS Publications Series*, 66, 45-72.

Gunn, S.R. (1998), “Support vector machines for classification and regression”, *ISIS technical report*, 14(1), 5-16.

Hill, T. & W. Remus (1994), “Neural network models for intelligent support of managerial decision making”, *Decision Support Systems*, 11(5), 449-459.

Kasavana, M.L. et al. (1998), “Netlurking: The future of hospitality Internet marketing”, *Journal of Hospitality & Leisure Marketing*, 5(1), 31-44.

Kayakuş, M. (2021), “Determination of Ideal Artificial Neural Networks Parameters for Software Project Effort Estimation”, *European Journal of Science and Technology*, 22, 43-48.

Kim, J. et al. (2003), “Segmenting the market of West Australian senior tourists using an artificial neural network”, *Tourism Management*, 24(1), 25-34.

Klimasauskas, C.C. (1992), “Neural networks: An engineering perspective”, *IEEE Communications Magazine*, 30(9), 50-53.

Kotler, P. et al. (2017), *Marketing for hospitality and tourism*, Pearson Education India.

Law, R. (2000), “Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting”, *Tourism Management*, 21(4), 331-340.

Law, R. & N. Au (1999), “A neural network model to forecast Japanese demand for travel to Hong Kong”, *Tourism Management*, 20(1), 89-97.

Law, R. et al. (2019), Tourism demand forecasting: A deep learning approach”, *Annals of Tourism Research*, 75, 410-423.

Lewis, C.D. (1982), *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*, Butterworth-Heinemann.

Lijuan, W. & C. Guohua (2016), “Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow”, *Knowledge-Based Systems*, 110, 157-166.

Lippmann, R.P. (1988), “An introduction to computing with neural nets”, *Acm Sigarch Computer Architecture News*, 16(1), 7-25.

Mata, J. (2011), “Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models”, *Engineering structures*, 33(3), 903-910.

Murat, Ã. et al. (2014), “Modelling and forecasting cruise tourism demand to Izmir by different artificial neural network architectures”, *International Journal of Business and Social Research*, 4(3), 12-28.

O'Connor, P. (1999), *Electronic information distribution in tourism and hospitality*, CAB International.

Palmer, A. et al. (2006), “Designing an artificial neural network for forecasting tourism time series”, *Tourism Management*, 27(5), 781-790.

Pattie, D.C. & J. Snyder (1996), “Using a neural network to forecast visitor behavior”, *Annals of Tourism Research*, 23(1), 151-164.

Polyzos, S. et al. (2020), “Tourism demand and the COVID-19 pandemic: An LSTM approach”, *Tourism Recreation Research*, 46(2), 175-187.

Pulliam, P. (1999), “To Web or Not to Web? Is not the question, but rather: when and how to Web?”, *Direct Marketing*, 62(1), 19-24.

Richter, L.K. (2003), “International tourism and its global public health consequences”, *Journal of travel research*, 41(4), 340-347.

Rosenbloom, B. et al. (1999), *Marketing Channels: A Management View*, 6th. Edition, The Dryden.

Sio-Chong, U. & Y.C. So (2020), “The impacts of financial and non-financial crises on tourism: Evidence from Macao and Hong Kong”, *Tourism Management Perspectives*, 33, 175-187.

Smola, A.J. & B. Schölkopf (1998), “On a kernel-based method for pattern recognition, regression, approximation, and operator inversion”, *Algorithmica*, 22(1-2), 211-231.

So, S.-I.A. & A.M. Morrison (2004), “Internet marketing in tourism in Asia: an evaluation of the performance of East Asian national tourism organisation websites”, *Journal of Hospitality & Leisure Marketing*, 11(4), 93-118.

Temel, G.O. et al. (2012), “Usage of Resampling Methods for Evaluating the Performance of Classification Model”, *International Journal of Informatics Technologies*, 5(3), 1-8.

Tsaur, S.-H. et al. (2002), “Determinants of guest loyalty to international tourist hotels-a neural network approach”, *Tourism Management*, 23(4), 397-405.

Vapnik, V. (2013), “*The nature of statistical learning theory*”, Springer Science & Business Media.

Walle, A. (1996), “Tourism and the Internet: opportunities for direct marketing”, *Journal of Travel Research*, 35(1), 72-77.

Wan, C.-S. (2002), “The web sites of international tourist hotels and tour wholesalers in Taiwan”, *Tourism Management*, 23(2), 155-160.

Wang, Y.-S. (2009), “The impact of crisis events and macroeconomic activity on Taiwan's international inbound tourism demand”, *Tourism Management*, 30(1), 75-82.

websiteiq.com (2020), <https://www.websiteiq.com/>, 23.12.2020.

World Health Organization (2020), “Coronavirus disease 2019 (COVID-19)”, *Situation Report*.

Witt, S.F. & C.A. Witt (1992), *Modelling and forecasting demand in tourism*, Academic Press Ltd.