

# AIZANOI FINANCE REVIEW

ISSUE: 2 / 2023

www.aizanoifr.com

info@aizanoifr.com

## COMPARISON OF ARTIFICIAL NEURAL NETWORK AND LOGISTIC REGRESSION METHODS IN CLASSIFYING THE HAPPINESS STATUS OF COUNTRIES

ÜLKELERİN MUTLULUK DURUMUNUN SINIFLANDIRILMASINDA YAPAY SİNİR AĞLARI VE LOJİSTİK REGRESYON YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Fatma Selen MADENOĞLU\*,

\*Assist. Prof. Dr., Abdullah Gül University, Faculty of Managerial Sciences, Department of Business Administration, selen.madenoglu@agu.edu.tr, <http://orcid.org/0000-0002-5577-4471>

ARTICLE INFO	ABSTRACT
<p><b>Received</b> 15.07.2023</p> <p><b>Revized</b> 13.08.2023</p> <p><b>Accepted</b> 07.09.2023</p> <p><b>Article Classification:</b> Research Article</p> <p><b>JEL Codes</b> C30 C40 M10</p>	<p>In recent years, happiness has become one of the main indicators of economic development. In this case, happiness has become a social goal rather than an individual goal. This study aims to classify the happiness status of countries. In order to achieve this goal, happiness situations are determined using the criteria. The study is applied to the countries included in the world happiness report. Logistic regression and artificial neural network methods are used to determine the happiness status of countries. As a result of the analysis, it shows that the logistic regression method is the appropriate method for classifying the happiness status of countries for this study.</p> <p><b>Keywords:</b> Happiness, Classification, Artificial Neural Network, Logistic Regression Analysis</p>

MAKALE BİLGİSİ	ÖZ
<p><b>Gönderilme Tarihi</b> 15.07.2023</p> <p><b>Revizyon Tarihi</b> 13.08.2023</p> <p><b>Kabul Tarihi</b> 07.09.2023</p> <p><b>Makale Kategorisi</b> Araştırma Makalesi</p> <p><b>JEL Kodları</b> C30 C40 M10</p>	<p>Son yıllarda mutluluk ekonomik kalkınmanın temel göstergelerinden biri haline dönüşmüştür. Bu durumda mutluluk bireysel hedef olmaktan çıkıp toplumsal hedef haline gelmiştir. Bu çalışmada ülkelerin mutluluk durumunun sınıflandırılması amaçlanmaktadır. Bu amaca ulaşmak için kriterler kullanılarak mutluluk durumları tahmin edilmektedir. Çalışma dünya mutluluk raporunda yer alan ülkelere uygulanmıştır. Lojistik regresyon ve yapay sinir ağları yöntemleri ülkelerin mutluluk durumunun belirlenmesinde kullanılmıştır. Yapılan analizler sonucunda, ele alınan çalışma için lojistik regresyon yönteminin ülkelerin mutluluk durumunu sınıflandırmada uygun yöntem olduğu belirlenmiştir.</p> <p><b>Anahtar Kelimeler:</b> Mutluluk, Sınıflandırma, Yapay Sinir Ağları, Lojistik Regresyon</p>

**Atf (Citation):** Madenoğlu, F. S. (2023). "Comparison of Artificial Neural Network and Logistic Regression Methods in Classifying the Happiness Status of Countries", *Aizanoi Finance Review*, (2): 27-48



Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

## Introduction

Happiness appears as an ultimate goal that everyone strives to achieve (Frey and Stutzer, 2022). Over the years, the concept of happiness has been presented with various aspects ranging from the subjective appreciation of individuals with their material or spiritual values to having values that will gain the appreciation of society, to their combination in various proportions, to the satisfaction with the lifestyle they lead. However, today, when considering happiness, perceived happiness, that is, whether the individual feels happy as an outcome of his/her own evaluation, has gained importance. A person's state of being happy with any subject can be learned by asking the question "how happy that person is with that thing" for his or her opinion on this subject. Such questions can refer to many topics, such as "work, health, income" or "life as a whole" (Van Praag, 2007). The happiness status of the countries is also determined by this analysis of the happiness status of the individuals.

An interdisciplinary area of computer science called machine learning automates the creation of sophisticated prediction models and algorithms by combining methods from other disciplines, including statistics, game theory, information theory, and optimization (Shalev-Shwartz and Ben-David, 2014). Since its inception in Arthur Samuel's 1952 Checkers-playing program and the early research by Hunt, et al. (1966) in inductive problem solving, Nilsson (1965) in statistical functions and data classification, Rosenblatt (1961) in neural networks, machine learning has undergone a significant evolution. It is currently a widely accepted approach that is utilized in fields including pattern recognition, the creation of new knowledge, and predictive analytics (Siegel, 2016). The successful application of machine learning as a classification techniques in areas such as: marketing and financial services, retail, travel, healthcare, sociology, and most recently social media (Finlay, 2014). This suggests that classification techniques could also be used to classify the happiness status of countries.

Saputri and Lee (2015) used a machine learning technique called Support Vector Machine for classifying the country happiness. Using data of 187 countries from the UN Development Project, the present model identify the most important factor needed to be improved by a certain country to increase the happiness of their citizens. Chaipornkaew and Prexawanprasut (2019) presented a happiness prediction model based on the survey. They used four machine learning techniques, namely KNN, Decision Tree, Naïve Bayes, and Multi-Layer Perceptron. Güneş et al. (2020) proposed an ordinal logistic regression for the variables affecting the happiness of provinces in Turkey. The data from the TUIK life satisfaction survey 2013 were used. The model resulted in the conclusion that happy people are those who are female, university graduates, upbeat, have positive relationships with their social surroundings, and are satisfied with their money, marriage, health, and social life. Ibnat et al. (2021) used machine learning, artificial intelligence, and computational strategy to predict the life satisfaction score of any specific country based on the defined parameters, emphasizing the happiest countries and regions based on the 2019 happiness report findings. You (2021) utilized the linear model and some tree models to analyze the features that are related to the happiness index and to make some predictions for Chinese adults. Jannani et al. (2021) predicted the quality of life by using different machine learning techniques on data from the World Happiness Report and made a comparative analysis to determine the most appropriate

machine learning algorithms. The best algorithms are Lasso regression, multiple linear regression, and long short-term memory (LSTM) to predict quality of life indicators for 2021. In order to investigate the impact of inequality on the degree of happiness in the globe, Kandemir (2021) employed the logistic regression by using the world happiness index data. According to the results of the logistic regression, disparities in health and education have a detrimental impact on people's levels of happiness, with disparities in education having a disproportionately greater impact. Khder et al. (2022) used the Neural Network training model and the OneR models to classify the happiness status of countries and identify the main features. The results were evaluated using different performance metrics such as accuracy and confusion matrix. Farooq and Shanmugam (2022) presented the results of the different machine learning algorithms to analyze the happiness status using the happiness report dataset which shows the data that how much citizens have freedom and the COVID-19 dataset. They identified the best algorithm based on computational studies. Lin and Horng (2023) presented the machine learning approach to explain and predict happiness scores. They used the given five prediction models: linear regression, random forest regressor, decision tree, Bayesian linear model, and Lasso Lars.

In the literature study, it was seen that there was no study used logistic regression and artificial neural networks together in the classification of the happiness status of countries using the 2021 World happiness report data. With this motivation, the presented paper aims to classify the happiness status of the countries in the World Happiness Report by using logistic regression and artificial neural networks.

The remained part of the paper has represented an order as follows. The second part explains the methodology. The third part presents the application. In the last part, results and conclusions are represented.

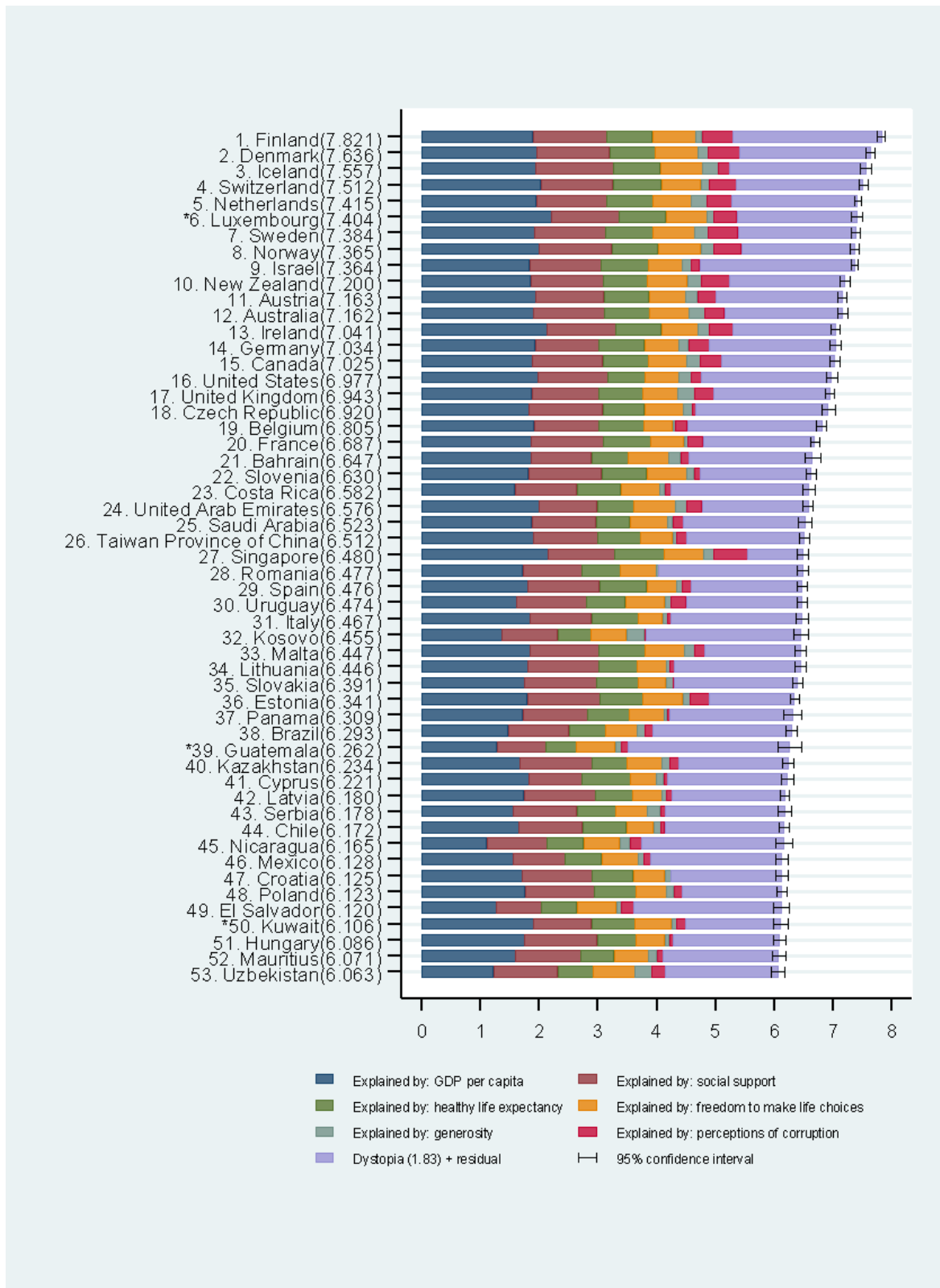
## **1. Methodology**

In the study, it's aimed to classify the happiness status of the countries. The World Happiness Index published in the World Happiness Report 2022 are used in the study. In the interest of establishing a composite index by which the countries can be ranked according to their happiness, the World Happiness Index (HI) attempts to define happiness based on a number of indicator variables, each of which is said to reflect a component of the general happiness of a country. Each of these indicators can also be connected to numerous sustainability-related topics. In support of the United Nations high-level summit on "Well-being and Happiness: Establishing a New Economic Paradigm," the World Happiness Report was first released in April 2012 and has since been released yearly. The first paper analyzed the growing science of happiness and offered current worldwide statistics on country happiness, demonstrating that a range of parameters can be used to consistently, legitimately, and reliably evaluate people's quality of life. Updated evaluations, in-depth special themes on the science of wellbeing, and a number of chapters on happiness in certain regions and countries are all included in each issue. The reports were typically prepared around a particular theme, but occasionally they had a different focus. World Happiness Report data is collected by a company called "Gallup World Poll". The company conducts this survey to approximately 1000 people over the age of 15 in more than 150 countries each year. When conducting these surveys, the company asks respondents to think of a 10-step ladder and on which rung they see themselves,

with 10 as the best possible life and 0 as the worst possible life. This ladder is called the "Cantril Ladder". The World Happiness Council combines the data from this survey with other data sources to form the World Happiness Report. The countries in the World Happiness Report are ranked according to the "Happiness Index", which is a single index, although it consists of many data and factors. The six key factors most related to the Happiness Index were determined. These factors are; economy, health, freedom, trust, generosity and corruption. These factors are used in the calculation of the happiness index. In addition to these factors, an additional factor called "Dystopia" is taken into account. Dystopia is an imaginary country with the most unhappy people in the world. The sum of these seven factors is called the "Happiness Score" and allows us to rank the countries according to their happiness with a single value. According to the latest report published in 2022, Finland is the happiest country in the world. Ranking of countries by happiness index values according to the World Happiness Report for 2022 are given in Figure 1. When the rankings in the world happiness report are examined, it is also seen how much of the total values given are caused by which factor. According to the world happiness report, Finland, Denmark, Iceland, Switzerland, Netherlands, Luxembourg, Sweden, Norway, Israel, New Zealand, which are the top 10 happiest countries, have high scores in terms of factors. According to the world happiness report, the 10 countries with the lowest scores are Zambia, Malawi, Tanzania, Sierra Leone, Lesotho, Botswana, Rwanda, Zimbabwe, Lebanon, Afghanistan. Six of the G20 countries (Australia, Germany, Canada, USA, United Kingdom, France) with the world's largest economy in dollar terms are in the top 20 of the happiness ranking. The G20 Group countries represent 85% of the world's population and 65% of the world's GDP. Only ten countries in G20 countries are in the top 50 of the happiness ranking. Turkey ranks one hundred and twelfth in this ranking.

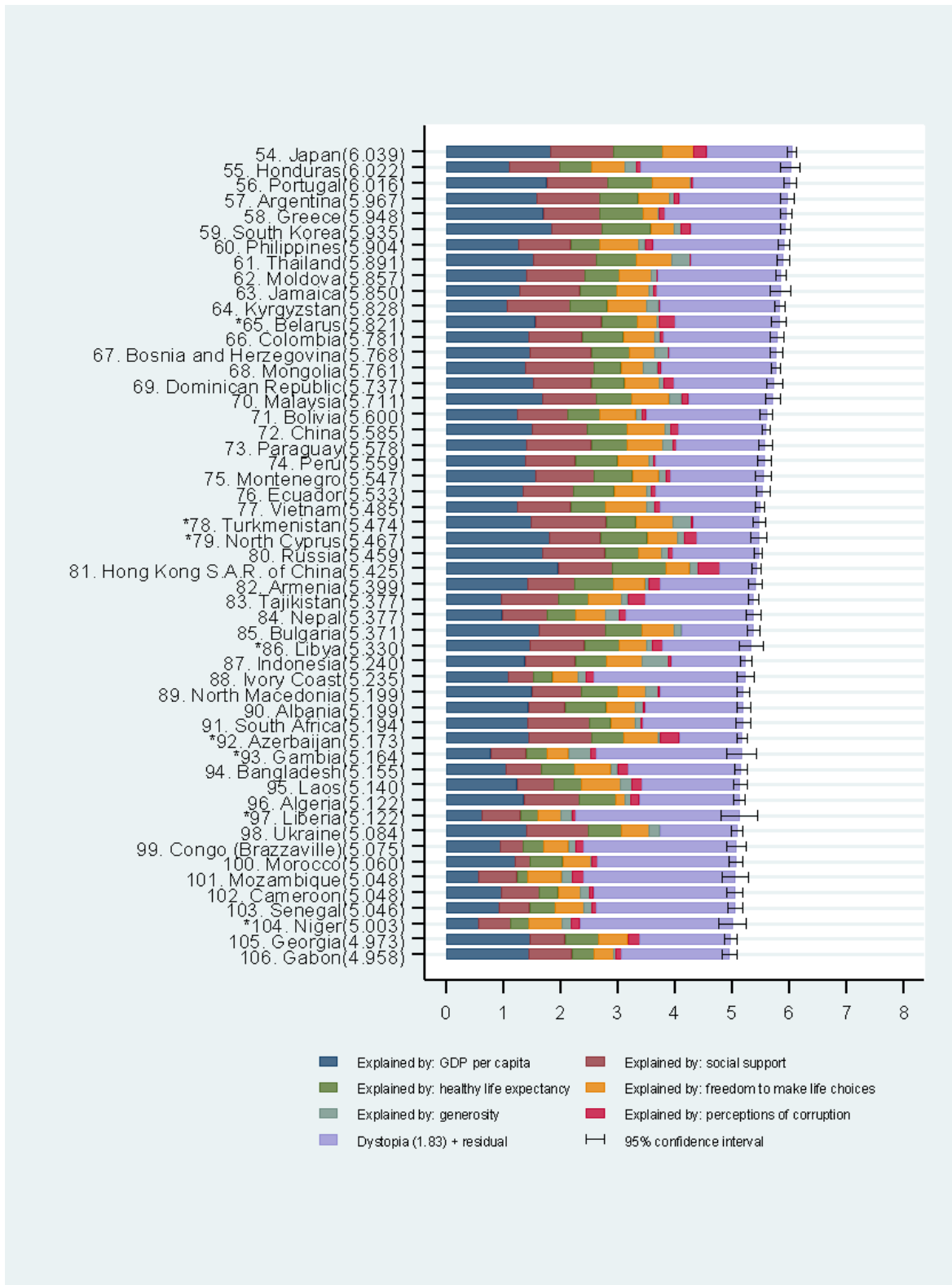
The study is carried out on behalf of 140 countries for which variable values can be obtained and are in the World Happiness Index. In this study, countries are classified according to their happiness values. First of all, the happiness index value, which is planned to be classified according to the logistic regression and artificial neural network, is rearranged into two categories. Here, while creating the categories, the average of the happiness index values given for the countries are taken and the countries below the average are coded as 0, that is, unhappy, and the countries above the average are coded as 1, that is, happy. When the data set did not contain outliers, the arithmetic mean instead of the median is taken as a measure of central position. From this point of view, in all methods to be used in the continuation of the study, the groups related to the happiness index will be expressed as output variable. Population, labor force, life expectancy at birth, school enrollment, primary (% gross), GDP per capita (current US\$), poverty headcount ratio at national poverty lines (% of population) of each country are taken into account as factors in order to examine the happiness status of countries. Population, labor force, life expectancy at birth, school enrollment, GDP per capita, poverty headcount ratio data taken from the World Bank database and happiness data is taken from the 2022 World Happiness Report.

**Figure 1.** Ranking of Countries by Happiness Index Values According to the World Happiness Report for 2022

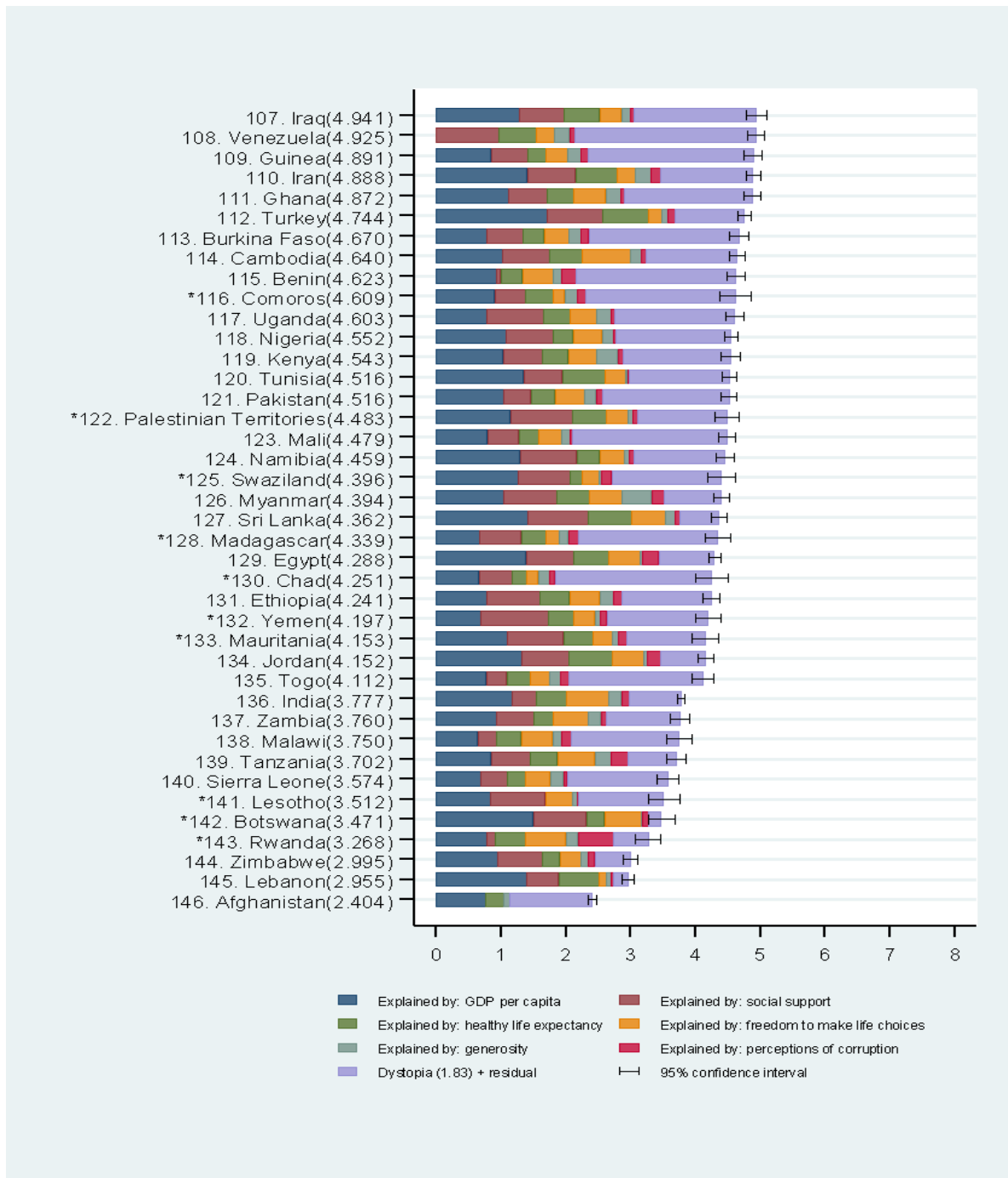


COMPARISON OF ARTIFICIAL NEURAL NETWORK AND LOGISTIC REGRESSION METHODS IN CLASSIFYING THE HAPPINESS STATUS OF COUNTRIES

**Figure 1.** Ranking of Countries by Happiness Index Values According to the World Happiness Report for 2022 (continued)



**Figure 1.** Ranking of Countries by Happiness Index Values According to the World Happiness Report for 2022 (continued)



### 1.1. Logistics Regression

In statistical applications, many regression methods have been developed to establish a relationship between the independent variable and the outcome variable. It is known that results cannot always be obtained by using only one of the developed methods. Therefore, alternative methods have been developed. The Logistic Regression (LR) has also been introduced for this purpose. LR is one of the machine learning models (Sperandei, 2014). LR is generated from a linear regression where the dependent variable is binary (Bailly et al. 2022). LR is useful in estimating a categorical variable. LR, which is similar to linear regression in many ways, is more suitable when the dependent variable is

a categorical data type with two or more options (Leech et al., 2004). Many of the variables studied in the social sciences are measured using a sensitivity scale, others are two-choice data, such as positive-negative, successful-unsuccessful, yes-no, and satisfied-not satisfied. Binary data is the most prevalent type of category data. The most popular model used in econometrics where the dependent variable has two-choice qualitative characteristics, such as yes/no, successful/unsuccessful, etc., is the logit model (Hill and Lim, 2011). For investigating the cause-and-effect relationship between the dependent variable and the independent variable, LR is utilized when the dependent variable is categorical data with two variables (Agresti, 2002).

Logit model is used in many fields such as economics, market research, politics, finance, and transportation engineering (Greene, 2012). The dependent variable is typically coded as "1" or "0" in binary logit models. The dependent variable is coded as "1" when the outcome is favorable or successful. The dependent variable is recorded as "0" since the outcome was unsuccessful or negative. Unlike conventional regression models ( $Y_i = \alpha_0 + \alpha_1 X_i + \varepsilon_i$ ), logit models conceal the error term. In logit models, there is not a  $\varepsilon_i$  as in the traditional regression model, but there is a  $\varepsilon_i$  in the background. Therefore, the estimated dependent variable can be interpreted as the probability of causing it (Kennedy, 2008).

LR allows us to construct a regression model without many assumption checks. LR is analysis where variables can be continuous or discrete. If necessary, in order to establish the LR model, a continuous result variable, in other words, the predicted variable can be transformed into a discrete variable (Tabachnick and Fidell, 1996). The method is widely used because it offers a more flexible structure and is easier to apply compared to similar analyzes.

In the LR, a model is established between one or more independent variables and the outcome variable. LR is used to model the relationships between the dependent variable measured in categories and some independent variables measured categorically or continuously. The difference from other methods is that in other regression methods, the outcome variable takes continuous values, while in LR, it is used when the outcome variable takes two or more discrete values (Hosmer et al., 2013). The major conclusion that can be reached from this investigation is that it is actually possible for the relationship between the dependent variable and the independent factors to be realized (Umar and Nashir, 2009).

The basis of the LR is based on the odds ratio. The odds ratio compares the probability that an event will occur with the probability that the same event will not occur. Thus, LR is obtained by taking the natural logarithm of the odds ratio. The maximum likelihood method is widely used when estimating the parameters of the obtained LR (Berenson and Levine, 1996).

LR is also an alternative to discriminant analysis and crosstabs. The validity criterion of the normality assumption is also discarded in this analysis because the dependent variable is one that may be classified as a binary preference with a range of 0 to 1. The results of this method are also incredibly adaptable and simple to interpret, which contributes to its high level of interest. In addition, conditions such as normality, linearity, covariance, and continuity, which are required in linear regression analysis, are not required in LR. The maximum likelihood method is used. The maximum likelihood method, on the other hand, refers to the parameter findings that provide the maximum



compatibility of the parameter values obtained from the application findings with the real data set we have (Alpar and Karabulut, 2007).

In any regression model, the expression  $E(y/x)$  represents the conditional mean value. This expression shows the mean value of the dependent variable ( $y$ ) given the independent variable ( $x$ ). The most important quantity in any regression problem is the mean value of the result for a given independent variable value. This quantity is the "conditional mean" and is denoted as " $E(y/x)$ ".  $E(y/x)$  is interpreted as "The expected value of  $y$  for a given value of  $x$ ".

$$E(y/x) = \beta_0 + \beta_1 x \quad (1)$$

Given  $x$  in LR (1), the conditional distribution mean of  $Y$  is shown as  $\pi(x) = E(y/x)$ . The special case of LR model is as follows.

$$E(y/x) = \pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (2)$$

A transformation of  $\pi(x)$  that will be at the center of our logistic regression work is the "logit\* transformation". This transformation is defined in terms of  $\pi(x)$  as follows.

$$g(x) = \ln \left[ \frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x \quad (3)$$

There is a linear relationship between the logit value and the independent and dependent variables. The logit  $g(x)$ , which is linear in its parameters, can be continuous and take values from  $-\infty$  to  $+\infty$ , depending on the definition range of  $x$ .

In the binary LR, there are two situations, 0 and 1, for the event to occur and not to occur. The ratio of the probability of an event to its probability of occurrence is defined as the odds ratio. Odds ratio (4) can take values between 0 and  $+\infty$  (Agresti, 1996).

$$\text{Odds ratio} = \frac{\pi(x)}{1 - \pi(x)} = e^{\beta_0 + \beta_1 x} \quad (4)$$

$\beta_0$  in the function represents the constant term.  $\beta_1$  is the slope coefficient.  $\beta_1$  represents the change in odds ratio versus a one-unit change in  $x$  (Gujaranti, 2010). In logistic regression, the maximum likelihood method is used instead of the least squares method in estimating the regression coefficients. In the maximum likelihood method, the probability of an event occurring is desired to be maximum (Alpar, 2013).

## 1.2. Artificial Neural Networks

Artificial neural networks (ANNs) are revealed by the design of nerve cells in the human brain and consist of interconnected processing elements with different levels of importance. ANNs are among the classification algorithms and are frequently preferred in practice today, where machine learning algorithms are increasingly being used. Although McCulloch and Pitts first used the concept of artificial neurons in their study in 1943, it started to be used quite frequently, especially after the 90s (Rojas, 1993). ANN is a mathematical model created by imitating the operation of simple biological nervous systems. The basic principles of ANN include gaining new information by learning the brain, gaining abilities such as critical thinking, problem solving and estimation. ANN is of interest to researchers because of its features such as producing solutions by learning the

relationship between input and output related to any event, whether linear or not, from the examples at hand, and correlating previously unseen events with past examples (Kaftan, 2010).

ANNs have been thoroughly utilized in many different fields for prediction in last years. Especially in non-linear time series, the superior success of ANN over traditional methods has been an important factor in the preference of this method (Zhang et al., 1998). Artificial neural networks can be applied in many different applications such as robotics, pattern recognition, medicine, power systems, signal processing, prediction and system modeling (Dogan and Atik, 2004; Ataman et al., 1998; Demir et al., 1998; Kalogirou and Bojic, 2000; Bayram et al., 2013; Sahin, 2014; Özdemir, 2011).

Haykin (1999) introduced the following definition for ANN: "A neural network is a densely parallel distributed processor composed of simple processing units, which has a natural tendency to accumulate experiential information and enable it to be used. This processor is similar to the brain in two ways: 1. Information is obtained by the network from the environment through a learning process. 2. Interneuron connection strengths, also known as synaptic weights, are used to accumulate the obtained information". As a machine learning technique, ANN is described as having the capacity to generate and find new information through behavior modification, similar to how the human brain does (Chehreh et al., 2008). Gamache et al. (2018) defined ANN a machine learning technique used to measure the relationship between input and output variables. Yakubu et al. (2019) presented artificial neural network models as an alternative to traditional regression models and defined them as a nonlinear parametric method that mimics the processing mechanism of the human brain. Kumar and Giri (2019) defined ANN as the simplification of biological neural structures and defined them as highly interconnected processing units (neurons) for understanding, acquiring information and using this information. ANN, which are commonly described as neural networks, are a prediction-based model in the common set of statistics, cognitive psychology and artificial intelligence, designed with inspiration from biological observations of human brain functions (Yang et al., 2016; Bhosale, 2016; Kasiviswanathan et al., 2016).

ANNs include traditional prediction models such as regression analysis, correlation analysis, moving averages, exponential smoothing method, box-jenkins method, autoregressive moving average, autoregressive integrated moving averages, simulation methods. The method differs from these methods in that it is a nonlinear model using the sum and transfer functions in the middle layer (Karahan, 2011; McNelis, 2005).

ANN can provide nonlinear modeling without making any assumptions, without the need for any prior knowledge between input and output variables. By giving the input information and the corresponding output information to the network, it is ensured that the network learns the relationship between input and output, and in this way, the training of the network is carried out (Hosmer, 1999).

Neural networks consist of three layers: input layer, hidden layer and output layer. Each layer consists of neurons. The basic element of the technique is the processing elements called neurons (Bağış and Konar, 2010). The input layer shows the level corresponding to the independent variable in the statistical analysis, and the output layer shows the level corresponding to the dependent variable. The hidden layer ensures that only the signals from the input layer are transmitted to the

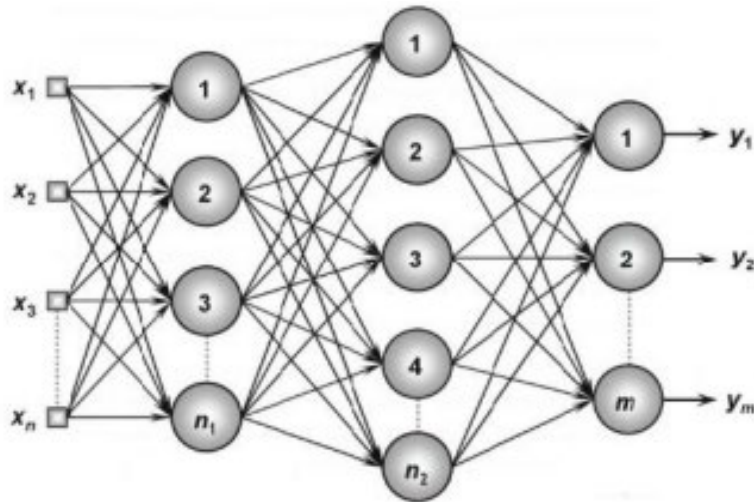
output layer (Budak and Erpolat, 2012). The connections of neurons to each other in various ways form a network. The function of the neuron is to add the values obtained by multiplying each of the input data with the weight values, with the threshold (addition function) value, and to pass this sum through the sigmoid or tanh activation functions and give it to the output. In general, neuron output is expected to be between (0-1) for *sigmoid* and (-1,1) for *tanh* (Şencan and Çiçek Bezir, 2003). ANN can be represented in many different ways. The structure of the multilayer perception model is as shown in Figure 2 (Ivan et al., 2016).

The information transmitted from the input layer is processed within the framework of certain standards and transmitted to the output layer. The main task of the network is the hidden layer. The number of hidden layers varies from network to network in line with the purpose to be realized. In the output layer, the information from the intermediate layer is processed. Output is produced in the output layer, which must be produced according to the inputs presented to the input layer. There are as many neurons in the output layer as the number of outputs of each data presented to the network. The values produced from the output layer constitute the output values created by the artificial neural networks against the problem. In modeling a problem in which the artificial neural network method is used, it is very important to determine the number of processing elements in the layers together with the number of hidden layers to be used for the best solution of the problem (Şahan and Okur, 2016).

In ANN, the process of updating the connection settings in the network to perform functions such as classification, modeling, optimization or prediction is called learning. This process, which is done in order to increase the performance of the network, is carried out in computer programs in the form of iterations. The learning process is carried out with algorithms (Çalışkan and Deniz, 2015).

Many classifications of artificial neural networks are encountered in the literature. Artificial neural networks are classified as feed forward and feedback recurrent artificial neural networks according to the direction of the information flow between the nerves in their structures, and are classified as supervised and unsupervised artificial neural networks according to the learning structure of the network. In feed forward artificial neural networks, there are multi-layer structures, one of which is the input layer, one is the output layer, and the others are hidden layers. In this artificial neural network structure, the information flow is forward-directed by being transmitted from the input layer to the hidden layer and transmitted from the hidden layer to the output layer. Examples of such networks are; multilayer perceptron and learning vector quantization. In feedback recurrent artificial neural networks, there is a multi-layered structure as in feed-forward networks. However, unlike the feed forward network structure, the information flow in this network structure can be backwards. Outputs of cells; it can also be used as an input to itself or another cell in the network (Sağiroğlu et al., 2003). Multilayer perceptrons are used in this study due to the usefulness of feed forward networks.

**Figure 2.** Structure of the Multi-Layer Sensor Model



Resource: Ivan Nunes da Silva, Danilo Hernane Spatti, Rogerio Andrade Flauzino, Luisa Helena Bartocci Liboni, Silas Franco dos Reis Alves,(2016), Artificial Neural Networks: A Practical Course, Springer, New York.

## 2. Application

### 2.1. Logistic Regression Results

LR is applied to the variables used in the study. LR is carried out with the WEKA program. The Weka program is a data analysis tool developed at the University of Waikato using the Java software language. It is also widely used in the field of data mining. There is a lot of information about machine learning, clustering and statistics available on the Weka program. The variables were determined and the significance of the model was tested. The Omnibus test was applied to evaluate the significance of the coefficients in LR. Omnibus test results obtained as a result of the analysis are presented in Table 1. Since  $p=0,000 < 0,05$ , the coefficients of the independent variables are nonzero. It is observed that there is a significant difference between the model with independent variables and the model with only constant variables. In other words, at least one of the independent variables added to the model contributes to the estimation of the dependent variable. The results show that the model is generally significant.

**Table 1.** Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	98,186	6	,000
	Block	98,186	6	,000
	Model	98,186	6	,000

The Hosmer-Lemeshow test is one of the methods used to determine whether the obtained model has fit to the data. The Hosmer-Lemeshow test is one of the methods used to determine whether the obtained model fits well with the data. This test tests whether there is a significant difference between the values predicted by the model and the observed values. The insignificance of the result of this test ( $p > 0.05$ ) indicates that the model-data fit is sufficient, in other words, the model has an

acceptable fit (Toraman and Karaca, 2016). As a result of the test, it's concluded that there is no significant difference between the estimation values and the observation values, in other words, the theoretical data represent the model well and significant ( $p=0,976>0,05$ ).

**Table 2. Hosmer and Lemeshow Test**

Step	Chi-square	df	Sig.
1	2,155	8	,976

The relationship between the estimated probabilities and the actual probabilities, showing how the model represents the data, and statistics of  $-2\log L$  are given in Table 3. In LR, the strength of the relationship between the dependent variable and the independent variables is determined by the Cox-Snell R Square and Nagelkerle R Square values. The values of the logistic regression explanatory coefficients (R Square statistics) show what percentage of the dependent variable is explained by the independent variables. As presented in Table 3, the explanatory power of the independent variables for the dependent variable was 53.6% according to the Cox&Snell R square value, while this value was 71.5% according to the Nagelkerke R square value.

**Table 3. Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	78,759	,536	,715

The results of LR are presented in Table 4. In this table, the number of countries correctly identified as happy (True positive-TP) is 60, the number of countries incorrectly as happy (False positive- FP) is 13, the number of countries correctly identified as unhappy (True negative-TN) is 56, the number of countries incorrectly identified as happy (False negative-FN) is 11. The calculated correct assignment rate must be greater than 50% to conduct the analysis. According to the results presented, countries defined as happy as a result of LR are classified as correct at a rate of 84.5%, while countries defined as unhappy are classified correctly at a rate of 81.2%. The overall accuracy rate is also found to be 84.5%. As a result, it can be said that the classification effectiveness of the model is quite good.

**Table 4. The Results of Logistic Regression Classification**

		Predicted		
		Happy	Unhappy	Percentage Correct
Happiness	Happy	60	11	84,5
	Unhappy	13	56	81,2
Overall Percentage				82,85

The results of the LR of the variables affecting the dependent variable are examined. The coefficients of the independent variables: population, labor force, life expectancy at birth, school enrollment, GDP per capita, poverty headcount ratio at national poverty lines are -0.0001, 0.0001, 0.2526, 0.0109, 0.0001, 0.0176, respectively. According to these coefficients ( $p$  values): population, labore force, school enrollment, GDP, poverty are determined as important factors, while life expectancy ( $p=0.2526>0.05$ ) are determined as unimportant factor. The odds ratios of the independent

COMPARISON OF ARTIFICIAL NEURAL NETWORK AND LOGISTIC REGRESSION METHODS IN CLASSIFYING THE HAPPINESS STATUS OF COUNTRIES

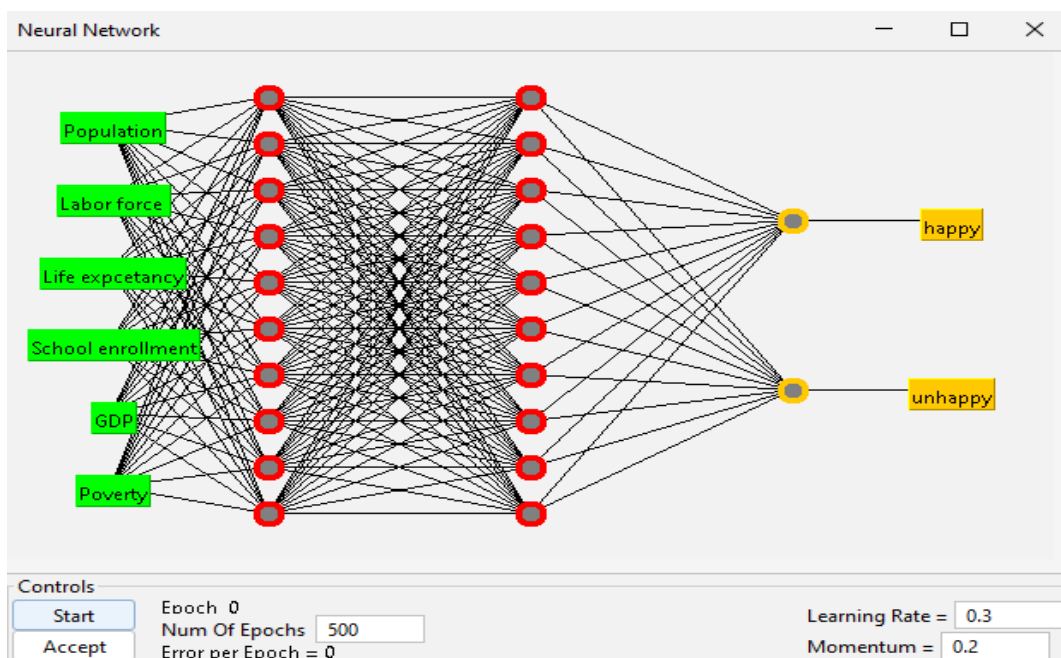
variables: population, labore force, life expectancy, school enrollment, GDP, poverty are 0.9999, 1.0001, 1.2873, 1.011, 1.0001, 1.178, respectively. The labor force, school enrollment, GDP, poverty have a positive, statistically significant effect on happiness. A one-point increase in the labor force component increases the happiness level of countries by about 1.0001 times. A one-point increase in the school enrollment component increases the happiness level of countries by about 1.011 times. A one-point increase in GDP component increases the happiness level of countries by about 1.0001 times. A one-point increase in the poverty component increases the happiness level of countries by about 1.0178 times.

### 2.2. Artificial Neural Networks Results

In the study, it is aimed to predict the happiness status of the countries. While estimating the state of happiness (happy, unhappy), artificial neural network method is used with the help of Weka software. Since the network has better learning ability in the presence of more than one layer in the artificial neural network structure to be used in the prediction analysis, the multi-layered network structure is determined as the network structure and the generated network structure is shown in Figure 3. As seen in Figure 3; The artificial neural network has an input parameter and an output parameter. In other words, the happiness status (output parameter) of the countries was estimated based on the population, labor force, life expectancy at birth, school enrollmen, GDP per capita, poverty headcount ratio at national poverty lines of the countries (input parameters).

As a result of the evaluation of the results of different parameter values, the use of the model using 2 hidden layer and 10 neurons is applied in the study. The classification models' performance for prediction was evaluated using cross-validation. It was also employed in previous studies (Hall et al., 2009; Wahbeh et al.,2011) since it evaluates how well a model works with new or test data. Cross validation is crucial, because a model is often fitted exclusively to the training dataset. Cross-validation enables the model's prediction accuracy to be seen when there is new data. The classification models in this study were evaluated using WEKA's 10-fold cross validation feature.

**Figure 3.** Artificial Neural Network Structure for the Prediction of Happiness Status of Countries



ANN classification values are presented in Table 5. In this table, the number of countries correctly identified as happy (True positive-TP) is 59, the number of countries incorrectly as happy (False positive- FP) is 14, the number of countries correctly identified as unhappy (True negative-TN) is 55, the number of countries incorrectly identified as happy (False negative-FN) is 12. The calculated correct assignment rate must be greater than 50% to conduct the analysis. According to the results presented, countries defined as happy as a result of ANN are classified as correct at a rate of 83.1%, while countries defined as unhappy are classified correctly at a rate of 79.7%. The overall accuracy rate is also found to be 81.4%. As a result, it can be said that the classification effectiveness of the model is quite good.

**Table 5.** *The Results ANN Classification*

		Predicted		
		Happy	Unhappy	Percentage Correct
Happiness	Happy	59	12	83,1
	Unhappy	14	55	79,7
Overall Percentage				81,4

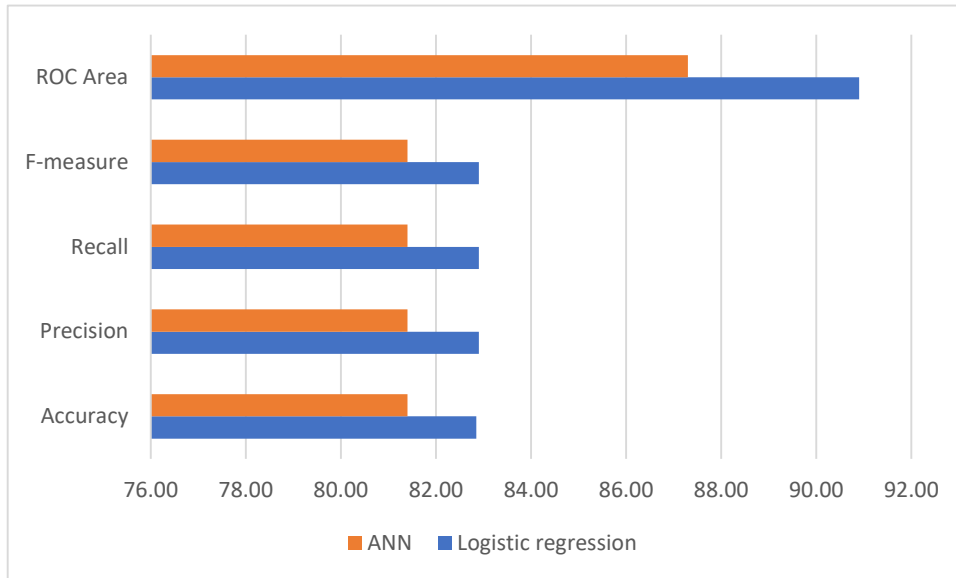
### 2.3. Performance Evaluation

To evaluate the performance of LR and ANN, different evaluation metrics are taken into account. An indicator of the model's performance across all classes is accuracy. It is measured as the proportion of correctly predicted events to all predicted events. The accuracy is calculated as the ratio of Positive samples that are correctly classified to all samples that are classified as Positive (either correctly or incorrectly). The precision assesses how well the model categorizes a sample as positive. The recall is determined as the proportion of Positive samples that are correctly identified as Positive to all Positive samples. The recall evaluates how well the model can identify Positive samples. The weighted harmonic mean of recall and precision is the F-measure. ROC area is the area under the curve of a plot that presents the True Positive Rate against the False Positive rate and measures the overall performance of the model. Table 6 and Figure 4 present the result of the evaluation metrics of LR and ANN. The given chart shows that LR gives better results for all evaluation metrics and has better accuracy for this dataset.

**Table 6.** *The Performance Evaluation of LR and ANN*

Method/Measurement	LR	ANN
Accuracy	82.85	81,4
Precision	82.9	81.4
Recall	82.9	81.4
F-measure	82.9	81,4
ROC Area	90.9	87.3

**Figure 4. Accuracy Chart**



Additional evaluation metrics are considered to compare the performance of the presented approaches. The consolidation between two sets of categorized data is measured by the kappa statistic (Melville et al., 2005). Kappa values range from zero to one. The stronger the agglomeration, the higher the value of the Kappa statistic. There is perfect agreement if Kappa = 1. There is no agreement when Kappa is equal to 0. If Kappa statics values vary between 0.40 and 0.59, they are regarded as moderate, 0.60 to 0.79, as substantial, and above 0.80, as extraordinary (Landis and Koch, 1977). The mean absolute error (MAE) is calculated by dividing the total absolute error by the number of estimates. It is a measurement of how closely a predicted model matches a given collection of actual values. The square root of the sum of squared errors divided by the number of estimates is the root mean square error (RMSE). The disparities between values assumed by a model and the values actually observed are measured. Better accuracy of the model is shown by a low RMSE value. So, the more accurate and precise the estimate, the lower the RMSE and MAE. The variation in errors in a group of forecasts can be identified by combining the MAE and the RMSE. The variation in the individual errors in the sample will always be bigger than the MAE or greater than the RMSE, depending on how much of a difference there is between the two. All errors have the same magnitude if the value of RMSE is equal to the value of MAE. When comparing a mean error (also known as a residual) to errors generated by a simple model, the relative absolute error (RAE) is expressed as a ratio. A ratio of less than one is the outcome of an acceptable model, or one that generates results that are better to those of a simple model. A ratio that is nearly zero will be generated by an accurate forecasting model; A poor model (one that performs worse than the simplistic model) will result in a ratio higher than 1. The square root of the squared errors of a predictive model normalized by the squared errors of a simple model is known as the Root Relative Squared Error (RRSE). The model outperforms the simple model when the RRSE is less than one. Therefore, the model is better the lower the RRSE.

Figure 3 show the obtained results by using LR. Figure 4 presents the results of ANN. These results state that the performance of LR better than the performance of ANN. There are a number of evidence that utilize LR is superior to ANN. Firstly, the kappa statistic offers a substantial boost in



predictions, and LR is of greater value than ANN. Secondly, a prediction is more accurate if the RMSE and MAE values are as low as possible. When compared to the ANN approach, the RMSE and MAE values for LR are the lowest. Third reason is the the value of RAE and RRSE. LR has higher RAE. LR outperforms ANN has lower RRSE.

**Figure 3.** *The Obtained Results by Using Logistic Regression*

```

=== Summary ===
Correctly Classified Instances      116      82.8571 %
Kappa statistic                    0.6569
Mean absolute error                0.2292
Root mean squared error            0.345
Relative absolute error            45.8431 %
Root relative squared error        68.9899 %
Total Number of Instances         140

```

**Figure 4.** *The Obtained Results by Using Artificial Neural Networks*

```

=== Summary ===
Correctly Classified Instances      114      81.4286 %
Kappa statistic                    0.6283
Mean absolute error                0.2272
Root mean squared error            0.3606
Relative absolute error            45.4449 %
Root relative squared error        72.1061 %
Total Number of Instances         140

```

### 3. Results

In recent years, there has been a general tendency to study the concept of happiness as an indicator of economic development. There has been an increase in the studies carried out on this subject, that happiness is the goal that is desired to be achieved socially rather than an individual goal. The status of happiness of countries is examined by survey studies and by considering different development factors as well as economic status indicators. Within the scope of this study, the classification of the happiness status of the countries is discussed by using LR and ANN. The happiness status of countries is determined by using the variables: population, labor force, life expectancy at birth, school enrollment, GDP per capita, poverty headcount ratio at national poverty lines. The world happiness index values is used to determine the happiness status of countries. Since the dependent variable in the model has 2 categories (happy, unhappy), the results of the multiple classification method were examined to determine the relationship between the independent variables. Within the scope of this study, current happiness values were first adapted to the analysis. The missing data in the data set was cleaned and made suitable for the .arff file format. The results obtained were divided into two categories, happy and unhappy countries. The cleaned data set was estimated using LR, and ANNs, which is the estimation tool of WEKA software. Statistical analyzes was performed with SPSS statistic software package.

The results show that population, labor force, school enrollment, GDP per capita, poverty headcount ratio at national poverty lines have a positive, statistically significant effect on happiness status. Different evaluation metrics are used to compare the performances of the classification methods.

COMPARISON OF ARTIFICIAL NEURAL NETWORK AND LOGISTIC REGRESSION METHODS IN CLASSIFYING THE  
HAPPINESS STATUS OF COUNTRIES

These are as follows: Accuracy, precision, recall, F-measure, ROC Area, Kappa statistic, mean absolute error, root mean squared error, relative absolute error, root relative square error. The comparative analysis revealed that LR offers better results. It is concluded that the LR is more appropriate than ANN for determining the happiness situations of countries using given variables. The study is expected to assist decision-makers by offering data assistance during the process of making strategic decisions and by identifying the indications that countries should take into account while attempting to improve their position.

**Author Contributions (Yazar Katkı Oranı):** Fatma Selen MADENOĞLU (%100)

**Ethical Responsibilities of Authors (Yazarın Etik Sorumlulukları):** This study was prepared in accordance with the rules of the required ethical approval

**Conflicts of Interest (Çıkar Çatışması):** There is no conflict of interest with any institution related to the study.

**Plagiarism Checking (İntihal Denetimi):** This study has been checked for plagiarism using a plagiarism scanning programme.

**REFERENCES**

- Agresti, A. (1996). *An Introduction To Categorical Data Analysis* (Vol. 1350), New York: Wiley.
- Agresti, A. (2002), *Categorical Data Analysis*, Second Edition, John Wiley and Sons Inc., New Jersey.
- Alpar , R. (2013). *Çok Değişkenli İstatistiksel Yöntemler* (4. Baskı), Detay Yayıncılık, Ankara.
- Alpar, R. & Karabulut, E. (2017). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*. 5.Baskı, Detay Yayıncılık, Ankara.
- Ataman, F., Kaynak, T., & Yüncü, S. *Analysing of Solutions Containing Artificial Intelligence Through System Modeling on Computer*. In *Electrical, Electronic and Computer Engineering 8th National Congress* (pp. 677).
- Bailly, A., Blanc, C., Francis, É., Guillotin, T., Jamal, F., Wakim, B., & Roy, P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine*, 213, 106504.
- Bağış, A. ve Konar, M. (2010) “Uçuş Kontrol Sistemi Yakıt Parametresinin Yapay Sinir Ağları Kullanılarak Belirlenmesi”, *Akıllı Sistemlerde Yenilikler Ve Uygulamaları Sempozyumu*, 21-24 Haziran 2010, Kayseri&Kapadokya, Türkiye.
- Bayram, S., Kaplan, K., Kuncan, M., & Ertunç, H. M. (2013, September). Ball bearings space of time statistical feature extraction and neural networks with error estimation method size. In *Automatic Control National Meeting, TOK2013, Malatya* (Vol. 26, pp. 28).
- Berenson, M. L. ve Levine, D. M. (1996), *Basic Business Statistics: Concepts and Applications*, Sixth Edition, Prentice-Hall International, New Jersey.
- Bhosale, B. (2016). Curvelet Interaction with Artificial Neural Networks. *Artificial Neural Network Modelling*, 109-125.
- Budak, H. & Erpolat, S., (2012). “Kredi Riski Tahmininde Yapay Sinir Ağları ve Lojistik Regresyon Analizi Karşılaştırması”. *Online Academic Journal of Information Technology*, 3(9), 23-30.
- Chaipornkaew, P., & Prexawanprasut, T. (2019, October). A prediction model for human happiness using machine learning techniques. In *2019 5th International Conference on Science in Information Technology (ICSITech)* (pp. 33-37). IEEE.
- Chehreh, C S, James, C H, Jorjani, E, Mesroghli, S & Bagherieh, A H, (2008). “Prediction of Coal Grindability Based on Petrography, Proximate and Ultimate Analysis Using Multiple Regression And Artificial Neural Network Models”. *Fuel Processing Technology*, 89,13- 20.
- Çalışkan, M. M. T. & Deniz, D. (2015) “Yapay Sinir Ağlarıyla Hisse Senedi Fiyatları ve Yönlerinin Tahmini”, *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 31, 177-194.
- Demir, Y., Tuntaş, R., & Köksal, M. (1998). Analysis of Switched Circuit with Neural Networks. In *Electrical and Computer Engineering 8th National Congress* (pp. 673).
- Dogan, H., & Atik, K. (2004). Providing Persistence in Comfort Conditions by the Application of Artificial Neural Networks to Air-Conditioning Systems. *Technology*, 7(2).
- Farooq, S. A., & Shanmugam, S. K. (2022). A Performance Analysis of Supervised Machine Learning Techniques for COVID-19 and Happiness Report Dataset. In *Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021* (pp. 591-601). Springer Singapore.
- Finlay, S. (2014). *Predictive analytics, data mining and big data: myths, misconceptions and methods*. Basingstoke: Palgrave Macmillan., pp. 6 – 9.
- Frey, B. S. & Stutzer, A. (2002). “What Can Economists Learn From Happiness Research?”. *Journal of Economic Literature*, 40(2), 402-435.

COMPARISON OF ARTIFICIAL NEURAL NETWORK AND LOGISTIC REGRESSION METHODS IN CLASSIFYING THE HAPPINESS STATUS OF COUNTRIES

- Gamache, K., Giardino, J. R., Zhao, P. & Owens, R. H., (2018). Bivouacs of the Anthropocene: Urbanization, Landforms, and Hazards in Mountainous Regions. *Urban Geomorphology: Landforms and Processes in Cities*. London: Elsevier, 205-230.
- Greene, W. H., (2012). *Econometric Analysis*. USA: Pearson Education Inc. ISBN: 0-13-066189-9.
- Gujarati, D. N. (2010). *Temel ekonometri*. (Çev. Ü. Şenesen ve G.G. Şenesen), Literatür Yayıncılık, İstanbul.
- Güneş, B., İsmail, T. A. Ş., & Sedat, A. C. A. R. (2020). "Türkiye'de İllerin Mutluluğu". *Ankara Üniversitesi SBF Dergisi*, 75(2), 747-774.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). "The WEKA Data Mining Software: An Update". *ACM SIGKDD Explorations Newsletter*, 11(1), 10-18.
- Haykin, S., *Neural Networks: A Comprehensive Foundation*, Perenctice Hall, New Jersey, 1999.
- Hill, R. C., Griffiths, W. E. & Lim, G. C., (2011). *Principles Of Econometrics*. Usa: Wiley. ISBN: 978-0-470-62673-3
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression (Vol. 398)*. John Wiley & Sons.
- Ivan Nunes da Silva, Danilo Hernane Spatti, Rogerio Andrade Flauzino, Luisa Helena Bartocci Liboni, Silas Franco dos Reis Alves,(2016), *Artificial Neural Networks: A Practical Course*, Springer, New York.
- Jannani, A., Sael, N., & Benabbou, F. (2021, December). Predicting Quality of Life using Machine Learning: case of World Happiness Index. In *2021 4th International Symposium on Advanced Electrical and Communication Technologies (ISAECT)* (pp. 1-6). IEEE.
- Kaftan, İ., 2010. Batı Türkiye Gravite ve Deprem Katalog Verilerinin Yapay Sinir Ağları ile Değerlendirilmesi. Doktora Tezi, Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, İzmir.
- Kalogirou, S. A., & Bojic, M. (2000). "Artificial Neural Networks for the Prediction of the Energy Consumption Of A Passive Solar Building". *Energy*, 25(5), 479-491.
- Kandemir, O., & Kürkcü, M . (2016). "Bir Refah Göstergesi Olan "Yaşam Endeksi" Bağlamında TR82 (Kastamonu, Çankırı, Sinop) Bölgesinin Analizi". *Kastamonu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 14(4) , 22-35.
- Kandemir, O. (2021). "Dünya Ülkelerinde Sosyal Eşitsizliklerin Mutluluk Düzeyine Etkisi: Lojistik Regresyon Analizi". *Uluslararası Avrasya Sosyal Bilimler Dergisi*, 12(45), 612-625.
- Karahan, M. (2011). İstatistiksel tahmin yöntemleri: yapay sinir ağları metodu ile ürün talep tahmini uygulaması. Yayınlanmamış Doktora Tezi. Konya: Selçuk Ü. Sosyal Bilimler Enstitüsü.
- Kasiviswanathan, K. S., Sudheer, K. P., & He, J. (2016). "Quantification Of Prediction Uncertainty İn Artificial Neural Network Models". *Artificial Neural Network Modelling*, 145-159.
- Kennedy, P., 2008. *A Guide to Econometrics*. (2008): Blackwell Publishing. ISBN: 978-1405182577.
- Khder, M. A., Sayfi, M. A., & Fujo, S. W. (2022). Analysis of World Happiness Report Dataset Using Machine Learning Approaches. *International Journal of Advances in Soft Computing & Its Applications*, 14(1), 14-34.
- Kumar, H. & Giri, S., (2019). "A Flow Shop Scheduling Algorithm Based On Artificial Neural Network". *Bulletin of Pure and Applied Sciences*, 38(E)(1), 62-71.
- Landis, J. R., & Koch, G. G. (1977). "The Measurement of Observer Agreement for Categorical Data". *biometrics*, 159-174.
- Leech, N. L., Barrett, K. C. & Morgan, G. A. (2004), *Spss For Intermediate Statistics: Use and Interpretation*, Second Edition, Lawrance Erlbaum Associates Publishers, Manwah New Jersey, London.
- Lin, Y. L., & Horng, S. C. (2023, May). Data analysis and prediction on world happiness score 2020. In *AIP Conference Proceedings* (Vol. 2685, No. 1). AIP Publishing.

- McNelis, P. D. (2005). *Neural Networks in Finance: Gaining Predictive Edge in the Market*. Elsevier, Amsterdam.
- Melville, P., Yang, S. M., Saar-Tsechansky, M., & Mooney, R. (2005). Active learning for probability estimation using Jensen-Shannon divergence. In *Machine Learning: ECML 2005: 16th European Conference on Machine Learning, Porto, Portugal, October 3-7, 2005. Proceedings 16* (pp. 268-279). Springer Berlin Heidelberg.
- Özdemir, V. E. Y. S. E. L. (2011). Determination of Turkey's Carbonization Index Based On Basic Energy Indicators By Artificial Neural Networks. *Journal of The Faculty of Engineering and Architecture of Gazi University*, 26(1), 9-15.
- Rojas, R. (1993). *Theorie der Neuronalen Netze*, Springer, Berlin.
- Sahin, İ. (2014). "Prediction Of Surface Roughness Of Al/Sic Composite Material With Artificial Neural Networks Yapay Sinir Ağları İle Al/Sic Kompozit Malzemenin Yüzey Pürüzlülüğünün Tahmini". *Journal of the Faculty of Engineering and Architecture of Gazi University*, 29(1), 209-216.
- Sağiroğlu, Ş.; Beşdok E. & Erler, M., *Mühendislikte Yapay Zeka Uygulamaları I: Yapay Sinir Ağları*. Kayseri: Ufuk Kitap KirtasiyeYayıncılık, 2003.
- Saputri, T. R. D., & Lee, S. W. (2015). "A Study of Cross-National Differences in Happiness Factors Using Machine Learning Approach". *International Journal of Software Engineering and Knowledge Engineering*, 25(09n10), 1699-1702.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory To Algorithms*. Cambridge University Press.
- Siegel, E. (2016). *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*. Hoboken, NJ: Wiley., pp.1-16.
- Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia medica*, 24(1), 12-18.
- Şahan, M. & Okur, Y., (2016). Akdeniz Bölgesine Ait Meteorolojik Veriler Kullanılarak Yapay Sinir Ağları Yardımıyla Güneş Enerjisinin Tahmini. *SDU Journal of Science (E-Journal)*, 11(1), 61- 71.
- Şencan, A. & Çiçek Bezir, N., 2003. Ölçülebilen Meteorolojik Verilerle Güneş Radyasyonunun Yapay Sinir Ağları Metoduyla Tayini. *TMMOB Makine Mühendisleri Odası Güneş Enerjisi Sistemleri Sempozyumu ve Sergisi*, 20-21 Haziran 2003, Mersin, pp. 235-239.
- Tabachnick, B. G. & Fidell, L. S. (1996). *Using multivariate statistics*. New York, USA: Harper Collins College Publishers.
- Toraman, C. & C. Karaca (2016), "Kimya Endüstrisinde Faaliyet Gösteren Firmalar Üzerinde Mali Başarısızlık Tahmini: Borsa İstanbul'da Bir Uygulama". *Muhasebe ve Finansman Dergisi*, 111-127.
- Umar İ. K. & Bashir S. (2019). "Investigation of the Factors Contributing to Truck Driver's Involvement in an Injuty Accident". *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*. 26(3), 402 – 408.
- Van Praag, B. M. (2007). "Perspectives from the Happiness Literature and the Role of New Instruments for Policy Analysis". *CESifo Economic Studies*, 53(1), 42-68.
- Wahbeh, A. H., Al-Radaideh, Q. A., Al-Kabi, M. N., & Al-Shawakfa, E. M. (2011). "A Comparison Study between Data Mining Tools Over Some Classification Methods". *International Journal of Advanced Computer Science and Applications*, 8(2), 18-26.
- Yakubu, A., Dahloun, L., Shoyomb, A. & Yahaya, U., (2019). "Modelling Hatchability and Mortality in Muscovy Ducks Using Automatic Linear Modelling and Artificial Neural Network". *Journal of the Indonesian Tropical Animal Agriculture*, 44(1), 65-76.
- Yang, J., Ma, J., & Howard, S. K. (2016). "A Structure Optimization Algorithm of Neural Networks for Pattern Learning from Educational Data". *Artificial Neural Network Modelling*, 67-82.

COMPARISON OF ARTIFICIAL NEURAL NETWORK AND LOGISTIC REGRESSION METHODS IN CLASSIFYING THE HAPPINESS STATUS OF COUNTRIES

- You, L. (2021, March). Utilizing machine learning to predict happiness index. In *2021 2nd International Conference on E-Commerce and Internet Technology (ECIT)* (pp. 233-238). IEEE.
- Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). "Forecasting with Artificial Neural Networks: The State of the Art". *International Journal of Forecasting*, 14(1), 35-62.
- Saputri, T. R. D., & Lee, S. W. (2015). "A Study of Cross-National Differences in Happiness Factors Using Machine Learning Approach". *International Journal of Software Engineering and Knowledge Engineering*, 25(09n10), 1699-1702.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory To Algorithms*. Cambridge University Press.
- Siegel, E. (2016). *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*. Hoboken, NJ: Wiley., pp.1-16.
- Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia medica*, 24(1), 12-18.
- Şahan, M. & Okur, Y., (2016). Akdeniz Bölgesine Ait Meteorolojik Veriler Kullanılarak Yapay Sinir Ağları Yardımıyla Güneş Enerjisinin Tahmini. *SDU Journal of Science (E-Journal)*, 11(1), 61- 71.
- Şencan, A. & Çiçek Bezir, N., 2003. Ölçülebilir Meteorolojik Verilerle Güneş Radyasyonunun Yapay Sinir Ağları Metoduyla Tayini. *TMMOB Makine Mühendisleri Odası Güneş Enerjisi Sistemleri Sempozyumu ve Sergisi*, 20-21 Haziran 2003, Mersin, s. 235-239.
- Tabachnick, B. G. & Fidell, L. S. (1996). "Using multivariate statistics. New York, USA: Harper Collins College Publishers.
- Toraman, C. & C. Karaca (2016), "Kimya Endüstrisinde Faaliyet Gösteren Firmalar Üzerinde Mali Başarısızlık Tahmini: Borsa İstanbul'da Bir Uygulama". *Muhasebe ve Finansman Dergisi*, 111-127.
- Umar İ. K. & Bashir S. (2019). "Investigation of the Factors Contributing to Truck Driver's Involvement in an Injuty Accident". *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*. 26(3), 402 – 408.
- Van Praag, B. M. (2007). "Perspectives from the Happiness Literature and the Role of New Instruments for Policy Analysis". *CESifo Economic Studies*, 53(1), 42-68.
- Wahbeh, A. H., Al-Radaideh, Q. A., Al-Kabi, M. N., & Al-Shawakfa, E. M. (2011). "A Comparison Study between Data Mining Tools Over Some Classification Methods". *International Journal of Advanced Computer Science and Applications*, 8(2), 18-26.
- Yakubu, A., Dahloun, L., Shoyomb, A. & Yahaya, U., (2019). "Modelling Hatchability and Mortality in Muscovy Ducks Using Automatic Linear Modelling and Artificial Neural Network". *Journal of the Indonesian Tropical Animal Agriculture*, 44(1), 65-76.
- Yang, J., Ma, J., & Howard, S. K. (2016). "A Structure Optimization Algorithm of Neural Networks for Pattern Learning from Educational Data". *Artificial Neural Network Modelling*, 67-82.
- You, L. (2021, March). Utilizing machine learning to predict happiness index. In *2021 2nd International Conference on E-Commerce and Internet Technology (ECIT)* (pp. 233-238). IEEE.
- Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). "Forecasting with Artificial Neural Networks: The State of the Art". *International Journal of Forecasting*, 14(1), 35-62.