International Journal of Contemporary Educational Research (IJCER)

www.ijcer.net

Comparison of Artificial Neural Networks and Logistic Regression Analysis in PISA Science Literacy Success Prediction

Ali Bozak[1], Eren Can Aybek[2]
[1]Republic of Turkey Ministry of National Education
[2]Pamukkale University

To cite this article:

Bozak, A. & Aybek, E.C. (2020). Comparison of artificial neural networks and logistic regression analysis in PISA science literacy success prediction. *International Journal of Contemporary Educational Research,* *7*(2), 99-111. DOI: https://doi.org/10.33200/ijcer.693081

This article may be used for research, teaching, and private study purposes.

Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles.

The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material.

International Journal of Contemporary Educational Research
Volume 7, Number 2, December 2020, 99-111 ISSN: 2148-3868

Comparison of Artificial Neural Networks and Logistic Regression Analysis in PISA Science Literacy Success Prediction

Ali Bozak[1*], Eren Can Aybek[2]
[1]Republic of Turkey Ministry of National Education
[2]Pamukkale University

Abstract

The present study aims to determine which analysis technique-Artificial Neural Networks (ANNs) or Logistic Regression (LR) Analysis-is better at predicting the science literacy success of the 15-year Turkish students who participated in PISA research carried out in 2015 by using learning time spent on science, test anxiety, environmental awareness, environmental optimism, epistemological beliefs, inquiry-based science teaching and learning practices, instrumental motivation, and disciplinary climate in science classes as the predictor variables. For this purpose, the data from 5895 students who participated in the PISA 2015 test were analyzed. Models were developed using LR and ANNs, and the results were compared. As a result, although the classification performance of artificial neural network is significantly better compared to LR, it is understood that practical significance is low due to the intersection of AUC confidence intervals.
Keywords: Artificial Neural Networks, Logistic Regression, PISA, Science Literacy

Introduction

In today's world, education and training activities are considered not only as transferring information to students but also the acquisition of high-level skills such as applying the learned knowledge to real life problems, conducting team studies, and learning to learn. As the perspectives on education have recently changed, it is also valuable for students to gain these skills and to measure how much these skills are gained.

The research such as Trends in International Mathematics and Science Study (TIMMS), Progress in International Reading Literacy Study (PIRLS) are conducted worldwide to observe the results of investments made in the field of education and make comparisons with other countries. The Programme for International Student Assessment (PISA), which is implemented by the Organisation for Economic Co-operation and Development (OECD) countries, is one of the researches carried out for this purpose. PISA started to be implemented in 2000 and is held every three years with the participation of countries that are OECD members and non-member countries, as well. In PISA, in which the students who complete the compulsory education of the countries attend, the aim is to assess the level of students' basic knowledge and skills required to be successful in real life. 15-year-old students from seventh and above grade levels participate in PISA research (Republic of Turkey Ministry of National Education [MEB], 2016). In PISA, survey studies are also conducted on variables that may affect student performance. Through the items related to socio-cultural, economic, and educational fields, information is gathered about the students, parents, and schools.

The competencies of students in mathematics, science and reading are measured within the scope of PISA research. At this stage, the concept of *literacy* becomes prominent. Literacy is defined as the ability to use the knowledge and learning outcomes that they gained from the problems they face, to be able to make inferences about the subject by analyzing and to effectively communicate with others. In 2015 PISA research, mainly science literacy was measured. Science literacy is defined as the ability to deal with science-related issues (MEB, 2016).

It is crucial for policy makers to identify the characteristics that affect student success to make better decisions. In this context, PISA results are a very useful data source to discover student performance and variables that describe the performance. In this context, besides regression or structural equation models for predicting student performance, analyses such as ANN, support vector machines, decision trees, and data envelopment analysis

* Corresponding Author: *Ali Bozak, alibozak@gmail.com*

can also be used. LR and ANN analysis, which are used in this research, are discussed in detail in the next sections.

Logistic Regression

In LR analysis, prediction of the dependent variable is provided by using independent / predictor variables. The independent/predictor variables in this analysis can be continuous or categorical. However, the dependent variable must be categorical. Predicting of group membership is an example of a LR problem (Tabachnick & Fidell, 2013). In cases which the dependent variable is categorical, LR analysis is preferred rather than multiple regression analysis (Kline, 2011). LR is more flexible compared to other techniques. Contrary to most analyzes, the predictor variables do not have to be normally distributed. These variables can also be continuous, categorical, or mixed. However, while the negative predictor value can be produced in multiple regression analysis, LR does not produce negative values (Tabachnick & Fidell, 2013). There are two methods used in LR analysis. These are called standard and stepwise methods. Stepwise methods are two parts, forward and backward (Çokluk, 2010).

In cases where the dependent variable has two categories, binary LR is used. In LR, independent variables are determined through logarithmic operations. The LR equation with one independent variable is as follows (Rençber, 2018):

$$p(x) = \ln\left[\frac{E(y/x)}{1 + E(y/x)}\right] = \ln\left[\frac{\exp(\beta_0 + \beta_1 x)}{1 + exp(\beta_0 + \beta_1 x)}\right] \quad (1)$$

The value to be obtained from the formula varies between 0 and 1. β_0 indicates the constant in regression. β_1 regression coefficient specifies the effect of the independent variable on the dependent variable.

$p/(1- p)$ ratio refers to the probability of belonging to a group or not (odds). The logarithm of this value is called the logit value of logodds or p. This transformation can be formulated as follows (Retherford & Choe, 1993):

$$logit\ (p) = \log\left[\frac{p}{1 - p}\right] \quad (2)$$

When the equation (2) is considered as a correlation function and *x*s are placed to indicate independent variables, the following logit model emerges (Oğuzlar, 2005):

$$\log\left[\frac{P_i}{1 - P_i}\right] = z_i = \sum_{k=0}^{P} \beta_k x_{ik} \quad (3)$$

When the data is processed in this equation, the dependent variable increases to 1 or approaches to 0 by decreasing. If the resulting value is greater than 0.5, the solution will result in 1, and 0 if it is smaller (Rençber, 2018).

Artificial Neural Networks

ANNs are algorithms that mimic how neurons and the human brain work and are based on learning information such as the biological brain and creating new information automatically from what they learn. ANNs are a great interest for researchers as a branch of the science of artificial intelligence and it allows computers to learn by massively parallel computing systems (Jain, Mao, & Mohiuddin, 1996). Kim (2017) stated that the ANNs are actually a machine learning model. They are the models that try to imitate human intelligence, and these models were created by analyzing the structure of biological nerve cells (neurons). Neurons receive, process, and transmit information via biochemical reactions (Abraham, 2005). The general structure of a biological neuron is shown in Figure 1.

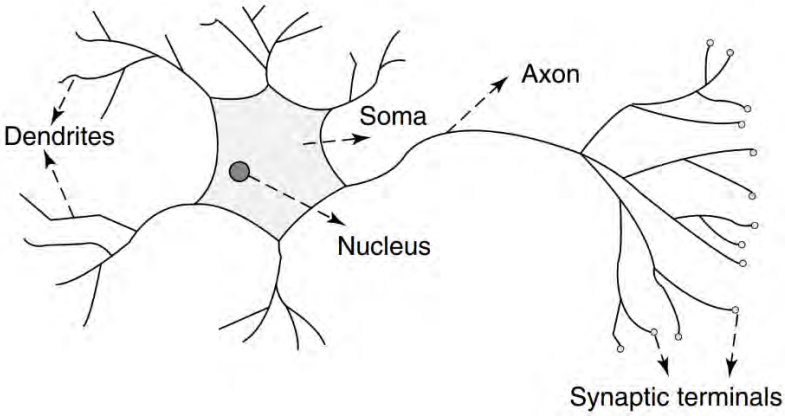

Figure 1. General Structure of the Biological Neuron (Abraham, 2005)

Dendrites in biological neurons receive input signals and transmit them to the soma. This information is processed in the soma. Axons convert these signals into output signals. Connections provide electrochemical contact between neurons. The basic principles of ANNs were first formulated by McCulloch and Pitts in 1943 based on assumptions listed as follows (Graupe, 2013):

- The activity of a neuron is all or none principle.
- The only significant delay in the nervous system is synaptic delay.
- The activity of any inhibitory synapse strictly prevents the simultaneous stimulation of the neuron.
- The structure of the connection network between neurons does not change in time.

ANNs are the structures made up of a large number of processing units connected together, similar to biological neurons. ANNs have interdependent processing elements (nodes). Each node is called an artificial nerve. An artificial neuron is shown in Figure 2.

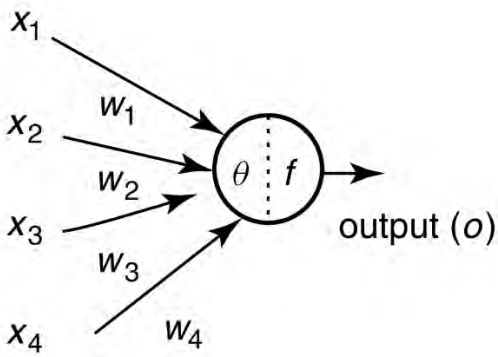

Figure 2. Artificial Neuron (Abraham, 2005)

Inputs: The data transmitted from the outside world or another neuron ($x_1, x_2 \ldots x_n$).

Weights: The data is processed by multiplying the weights. Weights are distributed randomly at the initial training phase (Henseler, 1995). Weights are called (w_1, $w_2 \ldots w_n$).

Summation function: It is the unit where weighted totals are calculated. Values from the inputs are summed up at this stage, multiplied by weights. The result is compared with the threshold value (Shanmuganathan, 2016). *w* as the weight value, *x* as the data entering the cell, *n* as the total value to indicate the number of inputs is calculated by the following formula (Graupe, 2013):

$$z_i = \sum_{i=1}^{n} w_{ij}\, x_i \tag{4}$$

Activation Function: The result of the collection phase is sent to the output being passed through the activation function. "Activation function converts the neuron input to its activation … This allows the variation of input

conditions to affect the output…" (Radi & Hindawi, 2013, p. 185). The main reason for using an activation function is to provide the non-linear feature to the neural network. Without using the activation functions, a neural network can perform a limited number of operations to learn nonlinear functions. The output we expect neural networks to learn is rarely linear (Heaton, 2012). Table 1 shows the commonly used activation functions and mathematical formulas (Rençber, 2018, p. 55).

Table 1. *Activation Functions*

Activation Function	*Mathematical Formulas*
Linear Function	$F(x) = x$
Step Function	$F(x) = \left\{ \begin{matrix} 1, & x > 0 \\ -1, & x \leq 0 \end{matrix} \right\}$
Threshold Function	$F(x) = \left\{ \begin{matrix} 1, & x > 0 \\ 0, & x \leq 0 \end{matrix} \right\}$
Sigmoid Function	$F(x) = \dfrac{1}{1 + \exp(-x)}$
Hyperbolic Tangent Function	$F(x) = \tanh(x)$
Gaussian Function	$\mathrm{F(x)} = \exp\left(-\dfrac{x^2}{2\sigma^2}\right)$

Output Function: The activation function generates an output value. This output value can be sent to another cell or it can be given out as a result (Veelenturf, 1995).

ANNs are made up of combined artificial nerve cells. Cells come together in the same direction in layers called input, hidden and output layers (Patterson & Gibson, 2017). These layers is shown in Figure 3.

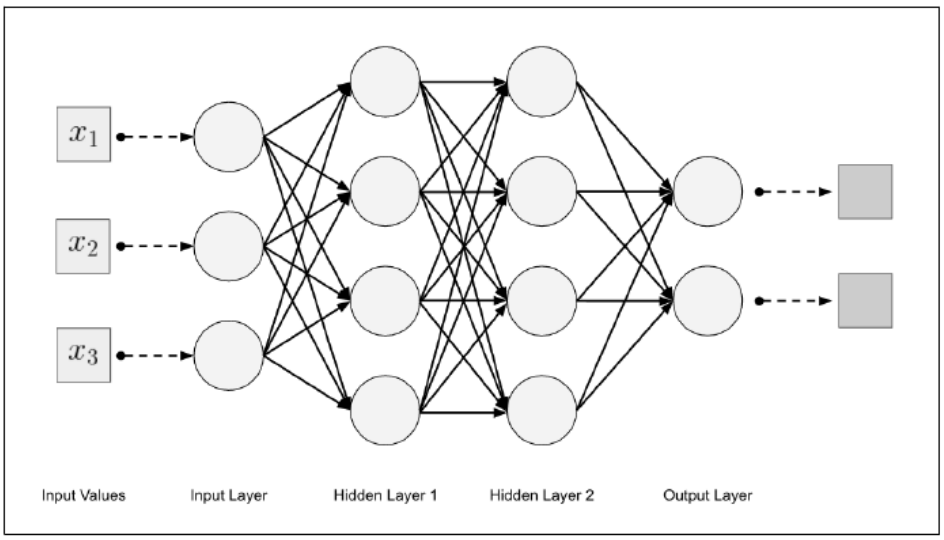

Figure 3. Artificial Neural Network Layers (Patterson & Gibson, 2017).

In the *Input Layer,* information coming from outside to the network is sent to *the Hidden Layer* which classifies the input vector and *Output Layer* retrieves the representations coming from the hidden layer (Hristev, 1998). In ANNs, the inputs are multiplied by weights and included in the function. Weight values must be determined for the network to learn correctly. This process is called *network training*. These values are randomly distributed when data first arrives in the neural network, they are recalculated as examples are given. In this way, suitable weight values are found. When the correct weight value is achieved, the artificial neural network reaches the level to generalize the problem represented by the samples. This is called network learning. The learning process of the network takes place through certain rules which are called learning rules. There are different learning rules such as Hebb rule, Hopfield rule, Delta rule, curved decline rule and Kohonen learning rule. After the completion of network training, it is tested whether it has completely learned or not. This process is called network testing. The network is asked to produce output by providing information that it does not see during training. The network produces results using the weights learned. The higher the results are, the better the network has learned (Elmas, 2018; Öztemel, 2006).

ANNs can be analyzed in three categories as supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, examples of input vectors are given, and the network learns through these examples (Shanmunagathan, 2016). Multilayer perceptron, least mean square, back propagation, and delta rule are some of the examples of supervised learning models (Akı & Karasulu, 2014).

ANNs are classified as single-layer or multi-layer according to layer architecture. Single layer networks have a single input layer and an output node. This simplest neural network is called a perceptron. Multilayer neural networks contain more than one computation layer. A perceptron contains an input and output layer. The input layer transmits data to the output layer and all calculations, and it is fully visible to the user. In multi-layered neural networks, there are additional middle layers between the input and output layers and these middle layers are called hidden layers because the calculations made are not visible to the user. The specific architecture of multi-layered neural networks is also called feedforward networks because the successive layers are supported by each other forward from the input (Aggarwal, 2018).

In the literature, there are some research which compare ANNs and LR to estimation the dependent / predictor variable in different fields (Alkan, Köklükaya & Subaşı, 2005; Biganzoli, Boracchi, Marianı & Marubini, 1998; Chiang, Zhang, & Zhou, 2006; Çakın & Özdemir, 2019; Dreiseitla & Ohno-Machadob, 2002; Eftekhar, Mohammad, Ardebili, Ghodsi & Ketabchi, 2005; Jefferson, Pendleton, Lucas & Horan, 1997; Kayri & Çokluk, 2010; Kurban, Kantar & Hocaoğlu, 2006; Kurt, Ture & Kurum, 2008; Manel, Dias, & Ormerod, 1999; Nefeslioğlu, Gökçeoğlu & Sönmez, 2008; Sarıkaya, 2019; Subaşı & Erçelebi, 2004; Tu, 1996; Vasconcelos, Silva, Tome, Alvim & Perelra, 2001; Yazıcı, 2018).

LR and ANNs are used in predicting student success as a predictor variable and there are studies on which method is more successful (Assi, Nahiduzzaman, Ratrout & Aldosary, 2018; Aydoğan & Zırhlıoğlu, 2018; Bahadır, 2016; Benzer & Benzer, 2017; Bezek Güre, Kayri & Erdoğan, 2019; Çırak & Çokluk, 2013; Delen, 2012; Gorr, Nagin, Szczypula, 1994; Hardgrave, Wilson & Walstrom, 1993; Ibrahim & Rusli, 2007; İnal & Turabik, 2017; Paliwal & Kumar, 2009; Pavlekovic, Bensic & Zekic-Susac, 2010; Teshnizi & Ayatollahi, 2015; Tosunoğlu & Apaydın, 2004). In the majority of the studies, it has been observed that ANNs create a better prediction model whereas some other studies have found no difference. This research, in which PISA 2015 data is used, is considered important in terms of its large data set to compare methods, and determining which variables are more important in predicting student success.

This study aims to compare ANNs and LR Analysis techniques at predicting the science literacy success of the 15-year Turkish students who participated in PISA research carried out in 2015. For this purpose, the following questions were considered:

1. According to the variables of students' learning time spent on science, test anxiety, environmental awareness, environmental optimism, epistemological beliefs, inquiry-based science teaching and learning practices, instrumental motivation, and disciplinary climate in science classes:
    a. What accuracy rate does the model created by LR produce in predicting the success of students' science literacy?
    b. What accuracy rate does the prediction model created by artificial neural network analysis produce in predicting students' success in science literacy?
2. Which model is better in predicting student achievement when ANNs and LR are compared?
3. What is the importance level of the variables included in the model in predicting science success?

Method

The variables used as predictors in the analysis are learning time spent on science (SMINS), test anxiety (ANXTEST), environmental awareness (ENVAWARE), environmental optimism (ENVOPT), epistemological beliefs (EPIST), inquiry-based science teaching and learning practices (IBTEACH), instrumental motivation (INSTSCIE), and disciplinary climate in science classes (DISCLISCI). LR and ANNs were used to determine the levels of predicting the success of science literacy of these variables and the estimation ratios were compared. Therefore, correlational model (Fraenkel, Wallen, & Hyun, 2012) has been adopted in the research. Predictor variables have been chosen according to their correlation with the science subtest score. The variables which have correlation above .10 are included in the analyzes. The only exception is test-anxiety which has correlation .08, yet it is considered as a potential important variable for the prediction of achievement.

Sample

The population of the students (aged 15) with which the PISA research conducted in 2015, is 1,324,089 and the population of the students that can take PISA test is 925,366 in Turkey. The sample of this study is 5895 students who participated in PISA 2015 test. In PISA, the school sample is determined using stratified random sampling. According to Statistical Regional Unit Classification in Turkey (SRE) Level 1; 5,895 students from 187 schools representing the 12 regions in 61 cities. When the number of students is compared, the result is that the most participation is from the Istanbul region and the least is from the Eastern Black Sea region (MEB, 2016).

Table 2. *Distribution of Students by Regions in PISA 2015*

Region Code	*Region Name*	*Number of Participants*	*Percentage of Participant Students*
TR1	İstanbul	1,070	18.15
TR2	Western Marmara	245	4.16
TR3	Aegean	707	11.99
TR4	Eastern Marmara	510	8.65
TR5	Western Anatolia	553	9.38
TR6	Mediterranean	817	13.86
TR7	Central Anatolia	334	5.67
TR8	Western Black Sea	303	5.14
TR9	Eastern Black Sea	194	3.29
TRA	Northeastern Anatolia	199	3.38
TRB	Central Eastern Anatolia	276	4.68
TRC	Southeastern Anatolia	687	11.65
TOTAL		*5,895*	
		100	

Data Collection Tool

In PISA research, literacy levels of students are measured in different areas such as reading, mathematics, collaborative problem solving, and science. In addition, the affective features associated with academic achievement score is assessed. Items applied to students includes the topics such as how motivated students were to learn about science, instrumental motivation towards science subjects, science self-efficacy (MEB, 2016). Computer-based tests are used in the PISA research. Each student has two hours to complete the test. Test consists of essay type and multiple-choice items. There is no fixed test form for students, i.e. different students take different item combinations. Moreover, students are asked to apply questionnaires about themselves, their families, and school environment (OECD, 2018). In this way, the students' willingness to learn science, whether they find science learning useful, how good they are at solving science problems and difficulties are investigated.

Data Analysis

First, students were ranked according to their science literacy levels, and 27% of the most successful group and 27% of the most unsuccessful group was determined. Accordingly, 1341 students were included in each group. Then, LR, via Jamovi 1.2.0.0 (The Jamovi Project, 2019) software, was used to determine whether students are in the lower or upper 27% according to their PISA 2015 science literacy scores (PV1SCIE) to be able to estimate with the variables SMINS, ENVAWARE, ENVOPT, IBTEACH, EPIST, ANXTEST, INSTSCIE and DISLISCI. To predict the same variable, an artificial neural network was created through Multi-Layer Perceptron (MLP) with the same predictor variables and which group the students belong to was estimated. *The standardized* method was used to the rescaling of continuous variables, and 90% of the data set was used for training and 10% of the data set was used for the test. This proportions are used because of two reasons. First, the data already split to 27% groups which has already reduced the amount of data, and second, generally full data set is used for LR in educational research. These proportions are used for LR and ANNs comparability purpose and both techniques used the same training data. Initial seed value was determined as 2081980 for reproducibility of the results. The lowest number of hidden layers is set to 1, and the maximum number of hidden layers to 50. In training, *batch* method was used. The artificial neural network has been tested through IBM SPSS Statistics v20. PROC v.1.16.1 (Robin et al., 2011) was used on R 3.6.0 (R Core Team, 2019) to compare the classification performance of both LR and ANNs.

Findings

According to the LR conducted to estimate whether the students are in the lower or upper 27% groups according to PISA 2015 science literacy scores, it was found that the *full-model* with all predictor variables was significant against the *constant-only-model* ($^{2}_{8}$=1169; p<.05). This shows that the predictor variables are able to distinguish between students with low and high science literacy. In addition, when the *pseudo* R^2 values of Negelkerke, McFadden, Cox and Snell, which are calculated for how much of the group membership of the students are estimated with the tested regression model, the following values are obtained respectively: .314, .353 and .471. R^2 values calculated in LR tend to be smaller than the values calculated in multiple linear regression models; therefore, it is considered sufficient to have *McFadden* R^2 value between .20 and .40 (Alpar, 2011; Tabachnick & Fidell, 2013). When the classification accuracy of the students was analyzed, it was found that a satisfactory classification was made even though no achievement variable was used. 76.5% of the students in the unsuccessful group; 79.8% of students in the successful group and 78.2% of all students in general was classified correctly. The Receiver Operating Characteristic (ROC) curve obtained for classification is given in Figure 4. The area under this curve [Area Under the Curve] (AUC) value was calculated as .856 [.842 - .871]. This value indicates that students are classified very well with predictor variables. Cantor and Kattan (2000) indicated the cut-off values for AUC as .80, good; .65, fair, and .50, poor. Regression coefficients and odds ratio values for predictor variables are given in Table 3.

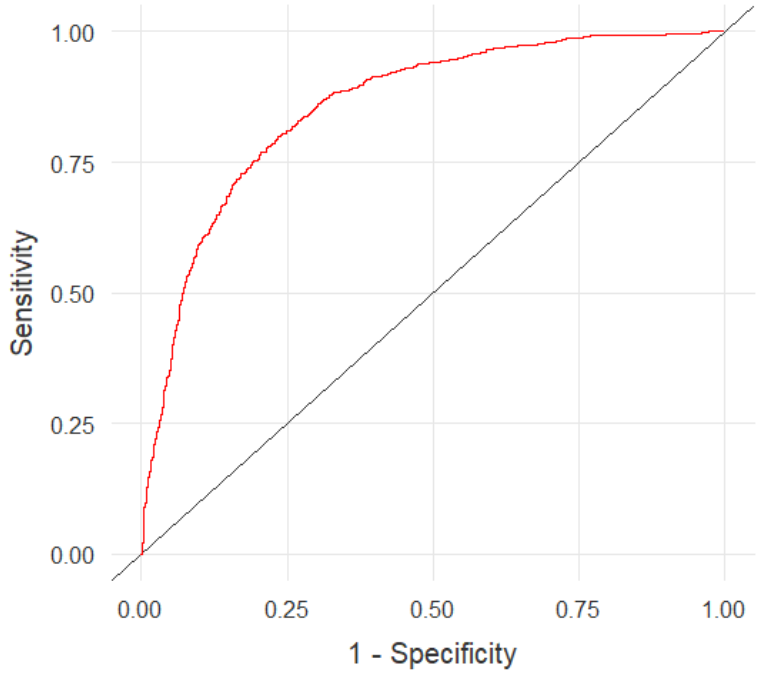

Figure 4. ROC Curve for Logistic Regression Model

Table 3. *Logistic Regression Analysis to Estimate Student Success*

					95% Confidence Interval for Odds Ratio	
Variables	B	SE	Z	Odds Ratio	Lower	Upper
Intercept	-1.896	.130	-14.54**	.150	.116	.194
SMINS	.007	5.27e-4	14.09**	1.007	1.006	1.008
ENVAWARE	.340	.037	9.22**	1.405	1.307	1.511
ENVOPT	-.538	.038	-14.18**	.584	.542	.629
IBTEACH	-.300	.044	-6.78**	.741	.679	.808
EPIST	.404	.045	9.05**	1.497	1.372	1.634
INSTSCIE	.185	.054	3.41**	1.203	1.082	1.338
ANXTEST	-.298	.048	-6.16**	.742	.675	.816
DISLISCI	.133	.054	2.47*	1.142	1.028	1.269

** p < .01; * p < .05

In Table 3, there are lower and upper limits of the 95% confidence interval for the regression coefficients, Z and Odds Ratio values for the eight predictor variables. Accordingly, all variables have a significant role in predicting students' position in the lower or upper group according to PISA 2015 science literacy success. In a one-unit increase in each predictor variable, 1 was subtracted from the e^B coefficients and multiplied by 100 to determine what change is expected in the group belonging odds of the students (Hair, Black, Babin, & Anderson, 2014). In this way, with a one-unit increase in the ENVOPT variable, a decrease of approximately 41.6% can be expected in the odds of students to be in the unsuccessful group ($[e^{-0.538} - 1]\ .100 = -41.61$). One-unit increase in the EPIST variable, an increase of approximately 49.8% can be expected in the odds of students to be in the unsuccessful group ($[e^{0.404} - 1]\ .100 = 49.78$).

According to the MLP results, it was observed that a satisfactory accuracy classification was made in both training and test sets, even any success related variable was not used when estimating the students' group. In the training set, it was found that 77% of students in the unsuccessful group, 80.7% of students in the successful group and 78.9% of all students in general can be classified correctly. In the test set, 80.4% of the students in the unsuccessful group, 88.1% of the students in the successful group and 84.3% of all students in general were classified correctly. The ROC curve in Figure 5 was obtained for the training and test sets. The area under this curve (AUC) was found to be .868 [.854 - .881] for both training and test sets. The fact that the AUC value obtained is greater than .80 indicates that the students are also classified very well with MLP.

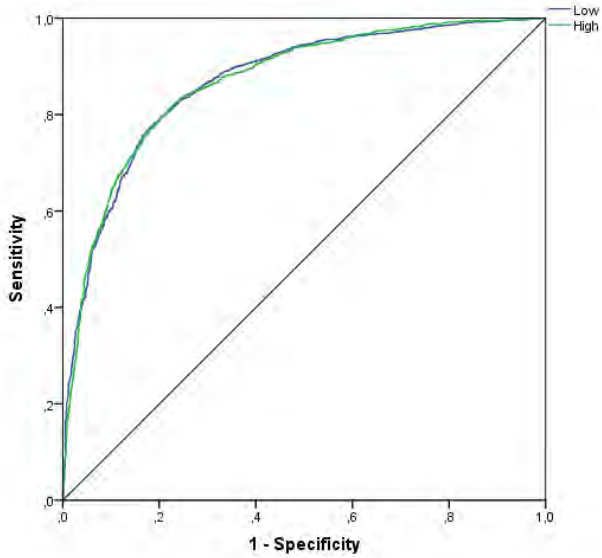

Figure 5. ROC Curve for Multi-Layer Perceptron

Correct and incorrect classification rates for estimating the lower or upper groups according to LR and ANNs according to PISA 2015 science literacy are summarized in Table 4.

Table 4. Correct Estimation Percentages for *Logistic Regression and Artificial Neural Network*

		Estimated		
	Observed	Unsuccessful	Successful	Correct Estimation Percentage
Logistic Regression	Unsuccessful	1026	315	76.5
	Successful	271	1070	79.8
	Total			78.2
ANN (Training)	Unsuccessful	923	275	77.0
	Successful	231	967	80.7
	Total			78.9
ANN (Test)	Unsuccessful	115	28	80.4
	Successful	17	126	88.1
	Total			84.3

As can be seen in Table 4, multilayer perceptron has been able to correctly classify the group that students belong to above 80%, especially in the test data. In addition, in comparing LR and classification performance of ANNs AUC values were compared as suggested by DeLong, DeLong, and Clarke-Pearson (1988) and a significant difference was found between the two AUC values (Z = 3.636; p <.05). However, when 95% probability confidence intervals of AUC values are analyzed, it is seen that confidence intervals are [.842 - .871] for LR and [.854 - .881] for artificial neural network, and therefore two confidence intervals intersect. In this context, although the classification performance of the artificial neural network is found to be significantly better than LR, the significance level is practically low due to the intersection of AUC confidence intervals. In the prediction of which group to the students belong to according to their science success with ANNs, the order of importance of the predictor variables is given in Figure 6.

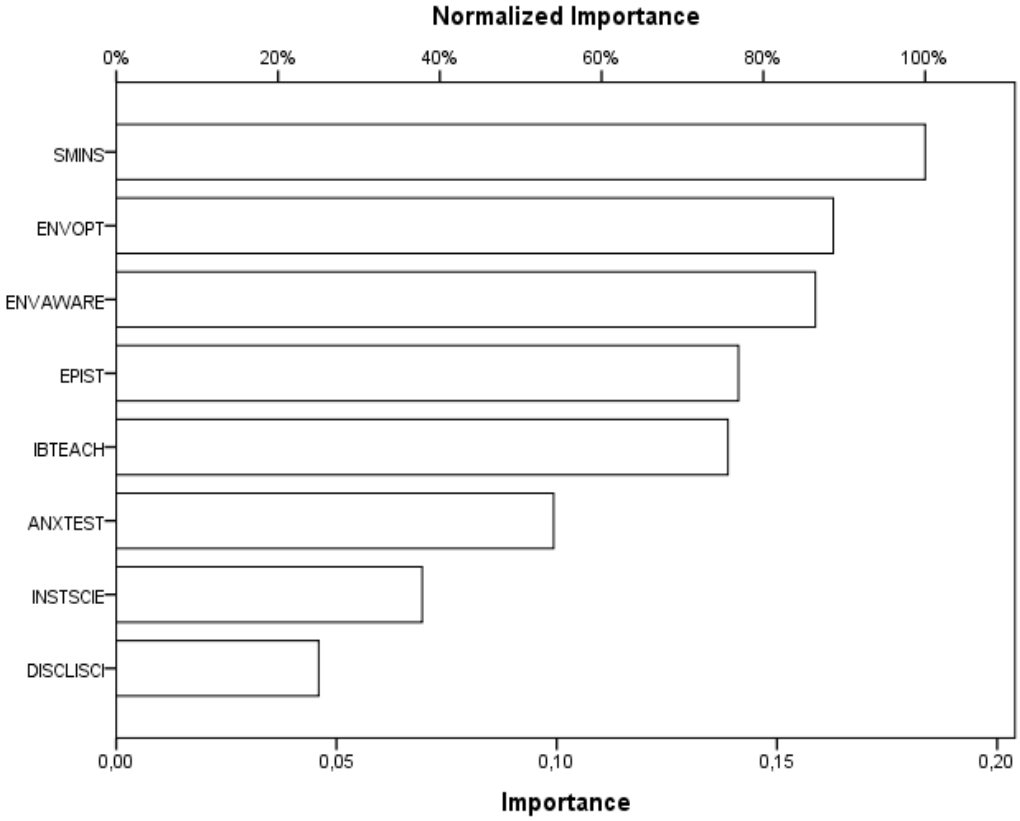

Figure 6. Order of Importance of Predictor Variables

According to the science literacy scores of the students, it is seen that SMINS is the most important variable in predicting if they are in the lower or upper groups with ANNs, and the least important variable is DISLISCI. Among the variables included in this study, it was concluded that the most important factor on student success is the time allocated to learning. In addition, students' interest in environmental problems and inquiry-based science education were found to be important variables that explain success. It is concluded that the variable that has the least important variable that explain student achievement is the disciplinary climate in the classroom.

Conclusion

LR and ANNs are two methods that can be used to predict a variable. Although it has been reported in the literature that ANNs perform better in the vast majority of studies comparing the two methods (Acar Boyacıoğlu & Kara, 2007; Alkan et.al, 2005; Assi et.al, 2018; Bahadır, 2016; Benzer & Benzer, 2017; Bezek Güre et.al, 2019; Biganzoli et.al, 1998; Budak & Erpolat, 2012; Chiang et.al, 2006; Çakın & Özdemir, 2019; Çırak & Çokluk, 2013; Delen, 2012; Flitman, 1997; Ibrahim & Rusli, 2007; Jefferson et.al, 1997; Kayri & Çokluk, 2010; Keskin Benli, 2005; Kurt et.al, 2008; Pradhan & Lee, 2010; Rençber, 2018; Sarıkaya, 2019; Subasi & Erçelebi, 2005; Teshnizi & Ayatollahi, 2015; Tolon & Güneri Tosunoğlu, 2008; Vasconcelos et.al, 2001; Walczak & Sincich, 1999; Yeşilnacar & Topal, 2005), there are also studies reporting that LR predicts at a higher level (Kurban et.al, 2007) or reaches the finding that they are equally successful (Tosunoğlu & Apaydin, 2004).

In addition, there are some studies in the literature that conclude that when two analyses are compared, there is no clear advantage, and the success of the method depends on the situation worked on (Dreiseitla & Ohno-Machadob, 2002; Eftekhar et.al, 2005; Gorr et.al, 1994; Hardgrave et.al, 1994; Manel et.al, 1999; Nefeslioğlu et.al, 2008; Pavlekovic et.al, 2010; Tu, 1996).

In the literature, it is seen that when comparing LR and ANNs' success conditions, the results show that the comparisons are carried out based on percentage values. In this study, after the percentage values were determined, the significance of the difference between the AUC values obtained for classification performance was also analyzed. Accordingly, which groups the students belong to can be estimated correctly with logistics regression at 78.2%, with ANNs at 84.3%. In addition, when comparing the classification performances, ANNs were shown to perform better. However, when the confidence intervals were considered, it was found that this difference was low in practice.

The results show that the most important variable in predicting students' science literacy was the time allocated to science education, and the least important variable was the disciplinary climate in the classroom. Other variables were environmental optimism, environmental awareness, epistemological beliefs, inquiry-based science teaching, test anxiety, and instrumental motivation in terms of the order of importance. When the variables are analyzed, the sensitivity of the student to environmental issues (ENVAWARE and ENVOPT) comes to the fore in predicting science success. In addition, it is expected that preference of inquiry-based science learning (IBTEACH) will increase the expected science literacy score in PISA.

Recommendations

This research was conducted on science literacy score with 2015 PISA results. ANNs and LR models can be used for different subtests of PISA. In addition, different variables can be included in the study to analyze their role in prediction. However, in the present study, only MLP were used as artificial neural network model. Different artificial neural network models or studies to test the predictive power of activation functions can be done. Variables such as computational cost and analysis time required were excluded from the purpose of this research. Therefore, further studies that compare LR and ANNs in terms of analysis time and required computational power are important in determining which method is more efficient to solve the same problem. It will also be useful to consider this finding in the studies to be carried out to increase student success. While developing policies on school and education system, it is especially important to pay attention to student responsibilities and effective teaching techniques, teacher competencies and classroom management.

Acknowledgements or Notes

The initial findings of this study have been presented in XI. International Congress of Education Supervision conference held in Antalya, Turkey on 5th – 7th December 2019.

References

Abraham, A. (2005). Artificial Neural Networks. In P.H. Sydenham and R. Thorn (Eds.), *Handbook of Measuring System Design* (pp. 901-908). John Wiley & Sons. doi:10.1002/0471497398.mm421

Acar Boyacıoğlu, M. & Kara, Y. (2007). A performance comparison between artificial neural networks and multivariate statistical methods in forecasting financial strength rating in Turkish banking sector. *Dokuz Eylül University Faculty of Economics and Administrative Sciences Journal*, *22*(2), 197-217

Aggarwal, C. C. (2018). *Neural networks and deep learning*. Switzerland: Springer International Publishing AG

Akı, K. & Karasulu, B. (2014). A review on the approaches and systems based on neuro fuzzy synergism in the soft computing (review). *Cumhuriyet University Faculty of Science Journal (CSJ), 35*(2), 54-86

Alkan, A., Koklukaya, E. & Subaşı, A. (2005). Automatic seizure detection in EEG using logistic regression and artificial neural network. *Journal of Neuroscience Methods*, *148*, 167–176.

Alpar, R. (2011). *Uygulamalı Çok Değişkenli İstatistiksel Teknikler [Application of Multivariate Statistics]*. Ankara: Detay Yayıncılık.

Assi, K. J., Nahiduzzaman, K. M., Ratrout, N. T. & Aldosary, A. S. (2018). Mode choice behavior of high school goers: Evaluating logistic regression and MLP neural networks. *Case Studies on Transport Policy*. doi: 10.1016/j.cstp.2018.04.006

Aydoğan, İ. & Zırhlıoğlu, G. (2018). Estimation of Student Successes by Artificial Neural Networks. *Yüzüncü Yıl University Journal of Education Faculty*, *15*(1), 577-610.

Bahadır, E. (2016). Using neural network and logistic regression analysis to predict prospective mathematics teachers' academic success upon entering graduate education. *Educational Sciences: Theory & Practice*, *16*(3), 943-964

Benzer, S. & Benzer, R. (2017). Examination of international PISA test results with artificial neural networks and regression methods. *The Journal of Defense Sciences, 16*(2).

Bezek Güre, Ö., Kayri, M. & Erdoğan, F. (2019). Predicting factors affecting PISA 2015 mathematics literacy via radial basis function artificial neural network. *Journal of Engineering and Technology, 3*(1), 1-11

Biganzoli, E., Boracchi, P., Mariani, L. & Marubini, E. (1998). Feed forward neural networks for the analysis of censored survival data: A partial logistic regression approach. *Statistics in Medicine, 17*, 1169 – 1186

Budak, H. & Erpolat, S. (2012). Comparison of artificial neural networks and logistic regression analysis in credit risk estimation. *AJIT-e: Online Academic Journal of Information Technology*, *3*(9), 24–30. doi: 10.5824/1309-1581.2012.4.002.x

Cantor, S.B. & Kattan, M.W. (2000). Determining the area under the ROC curve for a binary diagnostic test. *Medical Decision Making, 20,* 468-470. doi: 10.1177/0272989X0002000410

Chiang, W.K., Zhang, D. & Zhou L. (2006). Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression. *Science Direct Decision Support Systems*, *41*, 514 – 531. doi:10.1016/j.dss.2004.08.016

Çakın, E. & Özdemir, A. (2019). Estimation of the efficiency of technology development regions by artificial neural networks and logistics regression analysis on the basis of data envelopment analysis. *Hacettepe University Journal of Economics and Administrative Sciences*, *37*(2), 271-293

Çırak, G. & Çokluk, Ö. (2013). The usage of artificial neural network and logistic regression methods in the classification of student achievement in higher education. *Mediterranean Journal of Humanities*, *3*(2), 71-79

Çokluk, Ö. (2010). Lojistik regresyon analizi: Kavram ve uygulama. [Logistic regression analysis: Theory and practice]. *Educational Sciences: Theory & Practice*, *10*(3), 1357-1407

Delen, D. (2012). Predicting student attrition with data mining methods, *J. College Student Retention*, *13*(1), 17-35

DeLong, E., DeLong, D., & Clarke-Pearson, D. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics, 44*(3), 837-845. doi:10.2307/2531595

Dreiseitla, S. & Ohno-Machadob L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, *35*, 352–359

Eftekhar, B., Mohammad, K., Ardebili, H.E., Ghodsi, M. & Ketabchi E. (2005). Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. *BMC Medical Informatics and Decision Making*, *5*(3). doi:10.1186/1472-6947-5-3

Elmas, Ç. (2018). *Yapay zekâ uygulamaları [Applications of artificial intelligence]*. Ankara: Seçkin Yayıncılık

Flitman, A. M. (1997). Towards analyzing student failures: Neural networks compared with regression analysis and multiple discriminant analysis. *Computers Ops Res*, *24*(4), 367-377

Fraenkel, J.R., Wallen, N.E., & Hyun, H.E. (2012). *How to Design and Evaluate Research in Education.* New York: McGraw Hill.

Gorr, W.L., Nagin, D. & Szczypula, J. (1994). Comparative study of artificial neural network and statistical models for predicting student grade point averages. *International Journal of Forecasting*, *10*, 17-34

Graupe, D. (2013). *Principles of artificial neural networks*. Singapore: World Scientific Publishing Co. Pte. Ltd.

Hair,J.F., Black, W.C., Babin, B.J. & Anderson, R.E. (2014). *Multivariate Data Analysis.* Essex: Pearson.

Hardgrave, B. C., Wilson, R. L. & Walstrom, K. A. (1994). Predicting graduate student success: a comparison of neural networks and traditional techniques. *Computers Ops Res*, *21*(3), 249-263

Heaton, J. (2012). *Introduction to the math of neural networks*. Heaton Research. Inc, Chesterfield

Henseler, J. (2011). Back propagation. In P.J. Braspenning, F. Thuijsman, & A.J.M.M. Weijters (Eds.), *Artificial Neural Networks* (pp. 37-66). Springer-Verlag Berlin Heidelberg

Hristev, R.M. (1998). *The ANN Book.* Retrieved from http://www.pdg.cnb.uam.es/cursos/Complutense/Complutense2004/pages/12_NeuralNetworks/Hritsev_The_ANN_Book.pdf

Ibrahim, Z. & Rusli, D. (2007). Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression. *21st Annual SAS Malaysia Forum,* September 2007*,* Shangri-La Hotel, Kuala Lumpur

İnal, H. & Turabik, T. (2017). Determination of predictive power of some factors affecting mathematics achievement via artificial neural networks. *Uşak University Journal of Educational Research*, *3*(1), 23-50

Jain, A. K., Jianchang Mao, & Mohiuddin, K. M. (1996). Artificial neural networks: a tutorial. *Computer, 29*(3), 31–44. doi:10.1109/2.485891

Jefferson, M. F., Pendleton, N., Lucas, S.B. & Horan, M. A. (1997). Comparison of a genetic algorithm neural network with logistic regression for predicting outcome after surgery for patients with non-small cell lung carcinoma. *Cancer, 79*, 1338–1342.

Kayri, M. & Çokluk, Ö. (2010). Using multinomial logistic regression analysis in artificial neural network: an application. *Ozean Journal of Applied Science, 3*(2), 259 – 268

Keskin Benli, Y. (2005). Bankalarda mali başarısızlığın öngörülmesi lojistik regresyon ve yapay sinir ağı karşılaştırması. [Forecasting financial failure in banks: A comparison of logistic regression and artificial neural networks] *The Journal of Industrial Arts Education Faculty of Gazi University, 16*, 31-46

Kim, P. (2017). *MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence.* Apress.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: The Guildford Press

Kurban, M., Kantar, Y., & Hocaoğlu, F. O. (2007). Analysis of the power generation states of wind-solar power plant model by using logistic regression and perceptron models. *Süleyman Demirel University Journal of Natural and Applied Sciences*, *11*(1), 91-95

Kurt, I., Ture, M., & Kurum, A. T. (2008). Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert Systems with Applications*, *34*, 366 – 374

Manel, S., Dias, J. M. & Ormerod, S. J. (1999). Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: A case study with a Himalayan river bird. *Ecological Modelling*, *120*, 337 – 347

Nefeslioglu, H. A., Gokceoglu, C., & Sonmez, H. (2008). An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Engineering Geology*, *97*, 171–191

Oğuzlar, A. (2005). Lojistik regresyon analizi yardımıyla suçlu profilinin belirlenmesi [Determination of criminal profile with the help of logistic regression analysis]. *International Journal of Economics and Administrative Sciences, 19*(1), 21-34

Organisation for Economic Co-operation and Development [OECD]. (2018). *PISA 2015 PISA Results in Focus*. doi: https://doi.org/10.1787/aa9237e6-en

Öztemel, E. (2006). *Yapay Sinir Ağları [Artificial Neural Networks]*. İstanbul: Papatya Yayıncılık

Paliwal, M. & Kumar, U. A. (2009). A study of academic performance of business school graduates using neural network and statistical techniques. *Expert Systems with Applications*, *36*, 7865–7872

Patterson, J. & Gibson, A. (2017). *Deep Learning a Practitioner's Approach*. Sebastopol: O'Reilly Media

Pavlekovic, M., Bensic M. & Zekic-Susac M. (2010). Modeling children's mathematical gift by neural networks and logistic regression. *Expert Systems with Applications*, *37*, 7167–7173

Pradhan, B. & Lee, S. (2010). Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling & Software*, *25*, 747–759

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Radi, A. & Hindawi, S.K. (2013). Applying artificial neural network hadron – Hadron collisions at LHC. In K. Suzuki (Ed.) *Artificial Neural Networks: Architectures and Applications* (pp. 183-201). Rijeka: InTech.

Rençber, Ö. F. (2018). *Sınıflandırma Problemlerinde Çoklu Lojistik Regresyon, Yapay Sinir Ağları ve ANFIS Yöntemlerinin Karşılaştırılması: İnsani gelişmişlik endeksi üzerine uygulama [Comparison of Multinomial Logistic Regression, Artificial Neural Network, and ANFIS Methods in Classification Problems: An application on human development index]*. Ankara: Gazi Kitapevi

Republic of Turkey Ministry of National Education [MEB], (2016), *PISA 2015 Ulusal Raporu [PISA 2015 National Report]*, Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü, Ankara.

Retherford, R. D. & Choe M. K. (1993). *Statistical Models for Causal Analysis*. New York: A Wiley-Interscience Publication

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 77. doi: 10.1186/1471-2105-12-77

Sarıkaya G. (2019). Forecasting Bist national-100 index by using artificial neural network and regression models. *Education and Society in the 21st Century*, *8*(23), 325-340

Shanmuganathan, S. (2016). Artificial neural network modeling: An introduction. In S. Shanmuganathan & S. Samarasinghe (Eds.) *Artificial Neural Network Modeling* (pp. 1-14). Switzerland: Springer.

Subasi, A. & Erçelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine*, 78, 87-99

Tabachnick, B.G. & Fidell, L.S. (2013). *Using Multivariate Statistics.* Pearson: Boston.

Teshnizi, S. H. & Ayatollahi, S. M. T. (2015). A comparison of logistic regression model and artificial neural networks in predicting of student's academic failure. Acta Inform Med, 23:5, 296-300

The Jamovi Project (2019). *Jamovi (Version 1.0)* [Computer Software].

Tolon, M. & Güneri Tosunoğlu, N. (2008). Analyzing consumer satisfaction data: A comparison of artificial neural networks and regression analysis. *Gazi University Journal of the Faculty of Economics and Administrative Sciences*, *10*(2), 247-259

Tosunoğlu, N. & Apaydin, A. (2004). Öğrenci başarılarının sınıflandırılmasında lojistik regresyon analizi ve sinir ağları yaklaşımı [Logistic regression and artificial neural networks on the classification of student success]. *The Journal of Gazi University Faculty of Tourism. 1*, 170-188.

Tu, V. J. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemid*, *49*(11), 1225 -1231

Vasconcelos, M. J. P., Sllva, S., Tome, M., Alvim, M. & Perelra J. M. C. (2001). Spatial prediction of fire ignition probabilities: Comparing logistic regression and neural networks. *Photogrammetric Engineering & Remote Sensing*, *67*(1), 73-81.

Veelenturf, L.P.J. (1995). *Analysis and Applications of Artificial Neural Networks.* London: Prentice Hall.

Walczak, S. & Sincich, T. (1999). A comparative analysis of regression and neural networks for university admissions. *Information Sciences*, *119*, 1-20

Yazıcı, M. (2018). Comparison of discriminant analysis, logistic regression, and artificial neural networks in credit risk analysis. *Journal of Finance Letters*, *109*, 91-106

Yesilnacar, E. & Topal, T. (2005). Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Engineering Geology*, *79*, 251–266