

Research Article

Effects of Economic Growth on Employment: Comparison of Artificial Neural Network and Multilinear Regression as an Forecasting Model in the Light of GDP and Employment Data

Merve ÜNLÜ ASLAN

Lecturer, Department of Business Administration, Vocational Collage of Social Sciences, Kayseri University, Kayseri

merveunluaslan@kayseri.edu.tr, www.orcid.org/0000-0001-5677-1830

Erhan KAVUNCUOĞLU

Lecturer, Department of Computer Technology, Gemerek Vocational School, Cumhuriyet University, Sivas.

ekavuncuoglu@cumhuriyet.edu.tr, www.orcid.org/0000-0001-6862-2891

Esmâ UZUNHİSARCIKLI

Associated Professor, Department of Biomedical Device Technology, Kayseri Vocational Collage, Kayseri University, Kayseri, Turkey.

uzunhise@kayseri.edu.tr, www.orcid.org/0000-0003-2821-4177

Abstract

The purpose of this study was to develop a model to estimate employment data (ED) by using the data of Gross Domestic Product (GDP), which is an indicator of economic growth, and by focusing on Turkey. Two expert systems based on GDP and ED was designed and built using Artificial Neural Networks (ANN) and Multiple Linear Regression (MLR) models. Through these expert systems that used GDP data, it would be possible to predict how economic growth will affect employment. For this purpose, the ED was estimated by using the GDP data through these two models. The estimated ED that was obtained through both models was compared with the actual ED. As a result of the comparison, it was found that the expert system, which was formed by using ANN, produced more realistic results.

Keywords: Artificial Neural Network, Multiple Linear Regression, Data of Employment Regarding the Sectors, Gross Domestic Product, Economic Growth.

Received: 01.08.2020

Accepted: 18.09.2020

Suggested Citation:

Ünlü Aslan, M., Kavuncuoğlu, E., Uzunhisarcıklı, E. (2020). Effects of Economic Growth on Employment: Comparison of Artificial Neural Network and Multilinear Regression as an Forecasting Model in the Light of GDP and Employment Data, *Journal of Business Academy*, 1 (3): 235-260.

© 2020 Journal of Business Academy.

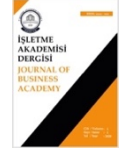


İşletme Akademisi Dergisi

2020, 1 (3): 235-260

DOI:10.26677/TR1010.2020.544

Dergi web sayfası: www.isakder.org



[Araştırma Makalesi](#)

Ekonomik Büyümenin İstihdama Etkileri: GSYİH ve İstihdam Verileri Işığında Tahminleme Modeli Olarak Yapay Sinir Ağı ve Çoklu Doğrusal Regresyonun Karşılaştırılması

Öğr. Gör. Dr. Merve ÜNLÜ ASLAN

Kayseri Üniversitesi, Sosyal Bilimler Meslek Yüksekokulu, İşletme Yönetimi Bölümü, Kayseri, Türkiye.

merveunluaslan@kayseri.edu.tr, www.orcid.org/0000-0001-5677-1830

Öğr. Gör. Erhan KAVUNCUOĞLU

Cumhuriyet Üniversitesi, Gemerek Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Sivas, Türkiye.

ekavuncuoglu@cumhuriyet.edu.tr, www.orcid.org/0000-0001-6862-2891

Doç. Dr. Esmâ UZUNHİSARCIKLI

Associated Professor, Department of Biomedical Device Technology, Kayseri Vocational Collage, Kayseri University, Kayseri, Turkey.

uzunhise@kayseri.edu.tr, www.orcid.org/0000-0003-2821-4177

Öz

Bu çalışmanın amacı, ekonomik büyümenin bir göstergesi olan Gayri Safi Yurtiçi Hasıla (GSYİH) verilerini kullanarak ve Türkiye'ye odaklanarak istihdam verilerini (İV) tahmin etmek için bir model geliştirmektir. GSYİH ve İV'ye dayalı iki uzman system Yapay Sinir Ağları (YSA) ve Çoklu Doğrusal Regresyon (ÇDR) modelleri kullanılarak tasarlanmış ve inşa edilmiştir. GSYİH verilerini kullanan bu uzman sistemler sayesinde ekonomik büyümenin istihdamı nasıl etkileyeceğini tahmin etmek mümkün olacaktır. Bu amaçla, İV buiki model üzerinden GSYİH verileri kullanılarak tahmin edilmiştir. Her iki modelden elde edilen tahmini İV, gerçek İV ile karşılaştırılmıştır. Karşılaştırma sonucunda YSA kullanılarak oluşturulan uzman sistemin daha gerçekçi sonuçlar verdiği görülmüştür.

Anahtar Kelimeler: Yapay Sinir Ağı, Çoklu Lineer Regresyon, Sektörlere Göre İstihdam Verileri, Gayri Safi Yurtiçi Hasıla, Ekonomik Büyüme.

Makale Gönderme Tarihi: 01.08.2020

Makale Kabul Tarihi: 18.09.2020

Önerilen Atıf:

Ünlü Aslan, M., Kavuncuoğlu, E., Uzunhisarcıklı, E. (2020). Ekonomik Büyümenin İstihdama Etkileri: GSYİH ve İstihdam Verileri Işığında Tahminleme Modeli Olarak Yapay Sinir Ağı ve Çoklu Doğrusal Regresyonun Karşılaştırılması, *İşletme Akademisi Dergisi*, 1(3): 235-260.

© 2020 İşletme Akademisi Dergisi.

1. INTRODUCTION

When unemployment problems emerge in the economy causing social concerns, policy-makers focus firstly on economic growth. With the increasing availability of data in an academic circle in recent years, significant increases have been observed in the number of studies which investigate the interaction between economic growth and other variables of growth (ErtugrulandUcak, 2013:662).

Parallel to the studies in the literature, this study examined the interaction between GDP and employment variables. For this purpose, two expert systems were designed by using GDP and ED variables from previous years. These expert systems used Artificial Neural Networks (ANN) and Multiple Linear Regression (MLR) as a model. Through these systems, it would be possible to predict how economic growth will affect employment. Also, in order to design these expert systems, summary tables were utilized. These summary tables consisted of GDP and ED variables which were divided into 26 regions according to the level 2 classification by Turkey Statistical Institute (TSI) (TSI, 2018a). At the same time, these data were separated into three according to the economic activity branches (agricultural, industry and services) by TSI.

Subsequently, the expert system that gives more realistic results was identified by comparing the real values with the estimated values of expert systems. As a result, it was determined that the Expert System, which was created by using the ANN, made more realistic estimations.

Briefly, the goal of this study was to examine the correlation between GDP variables and ED variable. Another goal was to design a specialized expert system which can predict employment rates (EA, EI, and ES) through GDP data using the relevant correlation ratios.

To summarise, in the first part of the study, the purpose of the study and the method used in the study are briefly mentioned. In the second part of the study, the concept of economic growth, GDP and employment, and the relationship of these concepts with each other was mentioned. In addition, theoretical viewpoints on these concepts were located. Finally, in the same part of the study the results of the previous empirical studies that analyzed the relation between GDP and ED were given. In the third part of the study, variables used in the study and classification of related variables were mentioned. In the fourth part of the study, the two separate expert systems, which were designed specifically for Turkey to estimate the effect of GDP on ED, were explained. Additionally, information was provided on the methods and processes used in relevant systems. In the fifth part, the results that were obtained from the two Expert Systems formed in the study were compared. As a result, it was determined which Expert System revealed more realistic ED in the same part of the study.

2. CONCEPTUAL EXPLANATIONS AND EMPIRICAL STUDIES IN THE LITERATURE

Economic growth and employment are two of the most important macroeconomic variables showing the success of the economy of a country and constitute indispensable elements of economic policies of many countries, especially the developed ones (Murat andYılmazEser, 2013: 94).

Among economists, there is a general consensus that employment will increase if economic growth happens (Murat andYılmazEser, 2013: 94). In other words, economic growth leads to an increase in employment services (Fuchs,1980: 94) At this point, if we would define the “employment” concept, which is used to refer to the labor factor; employment, briefly, is meant to be the use of labor in order to make production possible (Murat andYılmazEser, 2013: 94).

According to Word Bank, “employment” is defined as persons of working age who were engaged in any activity to produce goods or provide services for pay or profit, whether at work during the reference period or not at work due to temporary absence from a job, or to working-time arrangement (Wordbank, 2018a). In addition, the definition of “employment level” is the number of people engaged in productive activities in an economy (EUROSTAT, 2018).

As it is understood from the employment and employment level definitions, it is possible to define economic growth as leading to an increase in employment. If we are to define the economic growth by means of GDP, which is an indicator of national income measurement, “economic growth may be defined as an increase in the GDP, which is the monetary expression of all the goods and services that are produced in a particular country within a given period” (Muratoglu, 2011: 167). As it is understood from the definition, in recent years GDP is accepted as the real measure of the production power of the economy in national income computations due to reasons like facilitation of measurement, concentration of international economic integrations, and economic borders’ not recognizing political limits (Murat and Yılmaz Eser, 2013: 95).

According to the definition of Word Bank, Gross Domestic Product (GDP) is the sum of value added by all resident producers plus any product taxes (less subsidies) not included in the valuation of output (Wordbank, 2018b). According to another definition of the TSI, GDP is a standard measure of the value added to goods and services produced in a given period of time in a country (TSI, 2018b). Defining GDP as the total market value of all end-products and services produced within a given time period within the borders of a country, if the income obtained from these goods and services rises, the number of goods and services that are produced, and accordingly, the demand on the labor (i.e. the amount of employment) will also be reduced. On the contrary, if the income obtained from goods and services increases, there will also be an increase in the number of goods and services produced, and accordingly, an increase in the demand on labor and employment rates.

Briefly, there is no doubt that there is a need for labor to produce goods and services. The primary determinant of employment is the demand for domestic goods and services. The other factors that determine employment are export and import levels. When these factors are considered together, and when the difference between the imports-exports is added to the demand for domestic goods and services, GDP is obtained. Therefore, the first determinant of employment is GDP (Muratoglu, 2011: 167).

As it may be understood from all these explanations, the change in the GDP of a country affects employment at a considerable level. For this reason, it is important to explain the relationship between GDP and ED, and it is also important to anticipate how the change in GDP can affect the employment.

For this purpose, the studies which were been focusing on the relationship between GDP and employment data in the literature, have been investigated. As a result of the literature review, empirical studies in which the effect of economic growth on employment is examined through the GDP data given under this heading.

1. In his pioneering study, by using the data of the period 1947-1960, Okun (1963) reported that the relationship between growth and unemployment rates for the economy of the USA was a reverse relationship. In this respect, he determined that the unemployment rates decreased during periods when national income increased. However, unemployment rates increased during periods when national income decreased. Furthermore, he concluded that every 1% increase that exceeded the natural unemployment rate resulted in the reduction of real GDP at a rate of 3%.

2. Swane and Vistrand (2006) investigated the relationship between GDP and employment-growth in Sweden. By using the "employment-population rate" as the measurement of the employment creation dimension, he found a significant and positive relationship between GDP and the increase in employment.
3. Sawtelle (2007) estimated the employment elasticities in each of the 14 USA industry sectors as based on the changes in real GDP between the years 1991 and 2001, and then compared them with each other. In the same study, two *Employment Determination Models* were estimated for each industry sector and for the total economy. While one of the models associated employment with real GDP, the other associated employment with many other macroeconomic variables together with real GDP. Since the demand for the labour force is a derived demand, it was concluded in the study that the expansion of real GDP creates an increasing demand for workers.
4. Korkmaz and Yilgor (2010) analyzed the relationship between growth and employment in Turkey after the 2001 crisis. For this purpose, they applied the Granger Causality Test Method by using the quarterly-annual data for the 1997-2009 period. As a result of the Granger Causality Test Method, they found a causality relationship between GDP and employment within the given time period they examined. In other words, they concluded that the increase in GDP will also increase employment rates.
5. Muratoglu (2011) examined the relationship between employment and growth in the economy of Turkey for the 2000-2011 period by using econometric methods. He found a relationship between GDP and the employment series in the short term. In the long-run, on the other hand, he concluded that there was no relationship between the two variables. He also determined that there was no causality relationship between the two series for the analyzed period.
6. Ertugrul and Ucak (2013) examined the effects of economic growth on employment in Turkey for the 2000-2012 period. In this study, the effects of economic growth in Turkey on employment growth were examined with the Kalman Filter Method, which is one of the changing parameters methods (CPM). According to the results of the CPM model, it was determined that the effect of GDP growth on employment growth was stable during the 2003-2005 period however, there was a significant decrease in this effect in the 2006-2008 period. In the 2008-2009 period, where there was the global financial crisis period, it was determined that there was a large decline in the effect of GDP growth on employment growth, and in the 2010-2011 period, there was a recovery in the effect of the GDP growth on employment growth.
7. Murat and Yilmaz Eser (2013) examined the relationship between employment and economic growth and dealt with the growing phenomenon that economic growth did not create employment from the perspective of Turkey. In this context, they evaluated the validity of the growing phenomenon that did not create employment for Turkey. As a result, it was determined that both GDP and the number of employed people had an upwards trend for the 1970-2011 period over. In the same study, it was also determined that the years 1993, 2000, 2002, 2003 and 2004 were the "the years of growth without employment" and it was revealed that the most important reason for this was the increase in labour productivity.
8. Sarica et al. (2015) examined the long-term relationship between employment, real wages, and GDP. In their study, the quarterly industrial data covering the period 2005-2013 for Turkey were subjected to Border Co-Integration and Granger Causality tests. According to the empirical findings they obtained, they found a long-term relationship between employment, real wages, and GDP. According to the findings obtained from the Granger Causality Test, a one-way causality relationship was detected from GDP to employment. In the light of these data, it was concluded that the increase in GDP resulted in an increase in employment.

9. Abdioglu and Albayrak (2017) investigated the "growth phenomenon that does not create employment" in Turkey in terms of employment elasticity coefficients for the 1988-2015 period. According to the findings obtained in their study, the highest employment elasticity coefficient belongs to the construction sector. In addition to all these findings, it was also determined that positive/negative GDP deficits did not have different effects on employment for many sectors (except for the mining and construction sectors).

As a result of the literature review, studies focusing on the relationship between GDP and ED data and at the same time which can estimate employment rates using GDP data were also examined. Among these studies, a study which used the ANN model was not found.

3. VARIABLES USED IN THE STUDY AND CLASSIFICATION OF RELATED VARIABLES

Our work was designed to determine the impact of economic growth on employment. Therefore, it aimed to develop a model which can estimate the ED based on GDP in Turkey. For this purpose, GDP was calculated using the production method and ED were utilized for Turkey between the years 2004-2014. GDP and ED from summary tables published by the TSI were utilized in this study (TSI, 2018c; TSI, 2018d). These data are summarized under two separate classifications through the summary tables provided by the TSI.

The range, variety, and classification of GDP and ED are explained in Table 1.

As demonstrated in Table 1, the first classification is the Level-2 classification of the TUIK. In the Level-2 classification, Turkey is divided into 26 regions. The geographical areas that include these 26 regions are also explained in Table-1.

In the second classification, the GDP and ED are separated into three different economic activity branches, which are agriculture, industry, and services. In addition, maximum and minimum GDP and ED values are demonstrated in Table-1 according to three different economic activity areas. Furthermore, the abbreviations of GDP and ED were given (GDPA, GDPI, GDPS, EA, EI, ES) in Table-1.

Also, in the present study, two expert models were designed to predict the effect of GDP on the ED in Turkey by performing ANN and MLR models.

4. METHODS

4.1. Flowchart of Developed Models

The purpose of this study was to develop two proficient models that estimate volumes of Employment in the Agricultural Sector (EA), Employment in the Industry Sector (EI), and Employment in the Services Sector (ES) based on GDP data. The outline of the models consists of three steps as demonstrated in Figure 1, (1) obtaining and arranging survey data; (2) evolving MLR and ANN models and performing testing processes after the first training; (3) performing a statistical analysis to models and determining the most accurate estimated model.

In the first step, GDP and ED were selected as models in order to create input and output parameters and these parameters were used to create the models. Summary tables published by the TSI in certain periods were utilized to determine the parameters. In the modelling process, while GDP data were used as an input, ED was used as an output. In addition, the data were normalized at peak points (0 to 1 interval) that reached a certain range between the data.

In the second step, the MLR and ANN models were evolved. The number of neurons in the layers and learning algorithm were described for the ANN model. In the third step, MLR and

ANN, which were performed to test the data, were compared and the best-estimated model was decided.

4.2. Use of Variables in the Models

GDP was used for the model to be created in the research. For this reason, GDP data were created with the independent variables and inputs necessary to constitute the model in the study. These independent variables were Region, GDPA, GDPI and GDPS. In addition, three dependent variables and outputs were used which were EA, EI and ES. Consequently, the study used 286 research data (inputs and outputs).

4.3. Model of ANN

Feed-Forward Neural Network (FFNN) have a wide range of application fields of ANN (Wang et al., 2015: 282; Azadi and Sepaskhah, 2011: 178). Figure 2 shows one hidden layer FFNN.

Basically, three types of layers can be mentioned for FFNN. The input layer is the first layer where data is accepted into the network. (Azadi and Sepaskhah, 2011: 179). In the input layer, there are as many nodes as the number of attributes belonging to the data. The data is processed in the hidden layer, which is the layer after the input layer, and trained according to the desired output values. Each node in the input layer connects to each node in the hidden layer. Each link has a weight. Data from the input layer is multiplied by these weights and collected at each node in the hidden layer. Thus, a net input value is obtained. The output of each node in the hidden layer is obtained by transforming the total value with an activation function. Thus, the relationship between input and output values can be adjusted. (Zhang et al., 1998: 38,46,47; Wang et al., 2015: 285). Eq. (1) shows the output y_j of each knot j (Kocadagli, 2015: 53, 54).

$$y_j = f \left(\sum_{i=1}^m W_{ji} X_i + b_j \right) \quad (1)$$

Where X_i is the output of the i th connecting knot in the previous layer; W_{ji} is the weight of the connection between the i th knot of the previous layer. b_j is the knot threshold as bias input. Consequently, the output layer is the layer in which the desired values are shown as output. The values obtained in the activation function are multiplied by their weight as they are sent to the output layer and converted with a linear activation function. The number of outputs in the output layer is related to the research topic (Zhang et al., 1998: 45; Wang et al., 2015: 285; Beale et al., 2010: 22, 42).

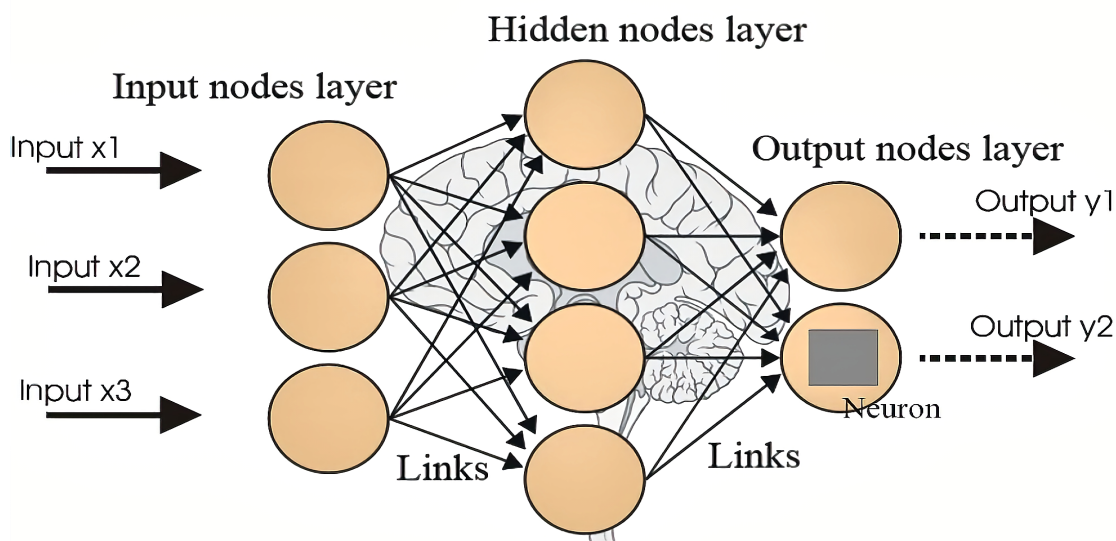


Figure 2. A Typical FFNN.

In this study, four variables Region, GDPA, GDPI and GDPS were used as input parameters; EA, EI and ES were used as target parameters; and these variables were introduced to ANN as inputs and outputs. There are studies in the literature showing that networks with a hidden layer are sufficient to solve problems with non-linearly distributed data (Wang et al., 2015: 285; Azadi and Karimi-Jashni, 2016: 17; Zhang et al., 1998: 44). A single hidden layer for this study was determined for estimating EA, EI and ES's volumes.

The number of nodes in the hidden layer is related to the complexity of the problem to be solved, and the number of nodes is determined by trials and errors obtained as a result of these trials (Wang et al., 2015: 284; Azadi and Karimi-Jashni, 2016: 19; Azadi and Sepaskhah, 2011: 179, 180). In the study, many attempts have been made to obtain the minimum error and the optimum number of nodes has been found.

Although many activation functions are used in the literature, practically few functions such as logsig, tansig and pure are used (Wang et al., 2015: 285; Zhang et al., 1998: 47). In this study, activation functions are used in different variations in output and hidden layer. Among these, the variation with the lowest error is the tansig and logsig activation functions in the hidden layer and the output layer, respectively.

During the training of the network, the weights are randomly determined and changed with a chosen optimization algorithm to minimize the error between the estimated values and the original output values. In practice, no optimization algorithm guarantees to find the global minimum in non-linear problem solutions and the main problem of optimization algorithms is fitting to local minimum points (Zhang et al., 1998: 48; Kocadagli 2015: 57, 62; Azadi and Karimi-Jashni, 2016: 18). Therefore, if the algorithms used cannot find the general minimum point, it aims to find the local minimum point. In this study, one of the algorithms using backpropagation determines the weights at which the minimum error will be obtained (Fukuoka et al., 1998: 1059 - 1062, Azadi and Karimi-Jashni, 2016: 18).

During the training of the network, the whole network has been trained between 500 and 1000 times and according to the error obtained, the whole network is started with new weight values.

Thus, the network tends to find the global minimum point rather than adhering to local minimum points. In the study, suitable weight values were determined to obtain the lowest error rate on the test data.

It is necessary to try many variations on both the layer and the number of neurons to achieve the minimum error. In this study, the optimum layer and number of neurons were determined to predict EA, EI and ES with minimum error. In addition, Levenberg-Marquardt optimization algorithm was used in the study to look at the error value of the network and update the weights.

Artificial intelligence techniques work more efficiently with normalized data. For this reason, in order to facilitate the learning process of the network, the data is first passed through the normalization process (Wang et al., 2015: 284, 285). This process squeezes the extreme peak values in the data into a certain range and meets the requirements that make the algorithm work efficiently. There are different methods for input normalization. Eq. (2) was used for data normalization

$$k_n = 0.1 + \left[\frac{k - k_{min}}{k_{max} - k_{min}} \right] \times 0.8 \quad (2)$$

in which k is the observed dataset; k_n is the normalized dataset; k_{min} and k_{max} are the minimum and maximum original dataset.

Network overload is one of ANN's biggest problems. In case of overload, the network is trained. However, the forecasting performance is low on the data that the trained network cannot see. When such a situation is encountered, early stopping technique is applied during training (Kocadağlı, 2015: 59; Azadi and Karimi-Jashni, 2016: 19).

All data is divided into 3 parts to implement the early stop technique. These are training set, test set and verification set. The training set is used to train the network and establish the relationship between inputs and outputs. The test set is used to measure the predictive performance of the developed model (Zhang et al., 1998: 50; Kocadağlı, 2015: 59). The number of iterations determined for the network was determined between 500 and 1000. The training process was stopped when the number of verification errors reached 6 iterations. In the study, 286 data were divided into education, verification, and testing, respectively, at 70%, 15%, and 15%. The data included in this ratio were taken randomly from all data, not sequentially. In the study, the network was first trained using the optimization algorithm with the separated training data, and then the predictive performance was measured with the test data. In addition, the t-test application was performed to look for reliability between the estimated EA, EI and ES and actual EA, EI and ES data for the ANN model ($\alpha=0.05$).

4.4. Model of MLR

MLR is a technique that models the relationship between two or more independent variables and was discovered by Francis Galton in the nineteenth century. This technique models the relationship between independent variables and dependent variables with a linear equation (Jahandideh et al., 2009: 2875; Azadi and Karimi-Jashni, 2016: 19; Shu and Lam, 2011: 635 - 637). In equation 3, n gives the number of observations for MLR.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon \text{ for } i = 1, 2, \dots, n. \quad (3)$$

y_i = dependent variable

x_i = independent variable

β_p = slope coefficients for independent variable

$\beta_0 = \text{constant term}$

$\varepsilon = \text{residuals (error term)}$

In this study, four independent variables Region, GDPA, GDPI and GDPS for 286 data were used. The volumes of EA, EI, and ES were determined as the dependent variable. 286 data were employed in the MLR model. Following this, 70% of all data (200 data) were used to obtain MLR equivalence. Then, the remaining 15% of data (43 data) was used for estimation.

4.5. The Method used to Evaluate the Estimate Performance

In assessing the performance of the forecast, the difference between the estimated values and the original values should be looked at. The following equation (Eq. 4) is used to measure the estimate ability of the models developed in this study.

Normalized Mean Absolute Error (NMSE) is an estimator of the overall deviations between estimated and measured values. It is defined as contrary to the bias, and the deviations (absolute values) are summed instead of the differences.

$$NMSE(x, y) = MSE(x, y) / MSE(x, 0) = \frac{\|x - y\|_2^2}{\|x\|_2^2} \quad (4)$$

where y is the approximation to x .

In addition, in the study, the t-test application was performed to look for reliability between estimated EA, EI and ES and actual EA, EI and ES data for the MLR model ($\alpha=0.05$).

5. RESULTS AND DISCUSSION

5.1. Assessment of Correlation Coefficient

The independent variables were defined as Region, GDPA, GDPI, and GDPS; and the dependent variables were defined as EA, EI and ES. Firstly, the correlation coefficient is a value used to show the relationship between the variables. The variable takes a value between -1 and 1 when calculated. Negative values indicate a negative relation and positive values indicate a positive relation. There is a perfect relationship when values are 1 or -1. Thus, the correlation between dependent variables and independent variables was investigated. As seen in Table 2, there were negative and positive correlation coefficients. These correlation coefficients between the data were examined for the evaluation of a relationship between these variables.

When the correlation coefficients were examined in Table 2, it was seen that the weakest correlation was between Region and EA ($r=0,332$) and the highest correlation was between GDPS and ES ($r=0,957$).

After observing the correlations between variables, MLR and ANN models were developed, and their performances were checked and compared.

5.2. Examination of ANN's Estimated Performance of EA, EI and ES's Volumes

Three layers were used to obtain the most correct estimation of EA, EI and ES volumes. The layers include one input layer, one hidden layer with different numbers of neurons, and one output layer. Two activation functions and one training algorithm were utilized in the designed network layers. To analyze the performance of the network, NMSE as a correctness scale, was calculated for each model. According to Table 3, the network that makes the estimate using the LM algorithm and tansig function which has 10 neurons in the hidden layer (4-10-3).

All of the data were separated into three divisions to perform learning methods. These were training data, validation data and test data. The training data was utilized to define the best-estimated model by modifying the network weights. With the validation data, each iteration was tested for how much the network learns and uses the error values of the validation data as a learning criterion, and the test data was utilized to scale the network's estimation efficiency.

The distribution chart is utilized to display the relationship between Estimated Data and Original Data as two variable and to display the correlation coefficient (Flott, 2012: 33, 34). If the distribution of the data in the chart resembles a straight line and the correlation coefficient is close to -1 or 1, this indicates that the correlation between the estimated data and original data as the two variables is high. Fig. 3, Fig. 4 and Fig. 5 display the distributions of the relationship between the original and the estimated values obtained from the ANN model. As demonstrated, the correlation values between the estimated EA, EI and ES values and the original EA, EI, and ES values were high.

As seen in Table 4, Table 5 and Table 6, the pairwise comparisons between the estimated EA, EI and ES values and the original EA, EI, and ES values were evaluated by using t-tests. There was no significant difference between the estimated EA, EI and ES values and the original EA, EI, and ES values ($p > 0.05$). As a result, this demonstrated the high estimated performance of the ANN models.

Table 4. The Relationship Between the Estimated EA and the Original EA is Demonstrated by Using T-Test as the Statistical Technique for ANN.

Group Name	N	Missing	Mean	Std Dev	SEM	P Value
Original EA	43	0	236441.861	156548.532	23873.427	0.875
Estimated EA	43	0	231377.186	140564.312	21435.857	

Table 5. The Relationship Between the Estimated EI and the Original EI is Demonstrated by Using T-Test as the Statistical Technique for ANN.

Group Name	N	Missing	Mean	Std Dev	SEM	P Value
Original EI	43	0	196627.907	239177.561	36474.236	0.953
Estimated EI	43	0	199597.099	224700.587	34266.518	

Table 6. The Relationship Between the Estimated ES and the Original ES is Demonstrated by Using T-Test as the Statistical Technique for ANN.

Group Name	N	Missing	Mean	Std Dev	SEM	P Value
Original ES	43	0	388395.349	365368.121	55718.116	0.908
Estimated ES	43	0	397569.033	366673.116	55917.126	

Relationships between the original EA, EI and ES values and estimated EA, EI, ES values are given in Fig. 3, Fig.4, and Fig.5. There were significant relationships between these variables (R:0,9524 for EA; R:0,9805 for EI; R:0,9897 for ES).

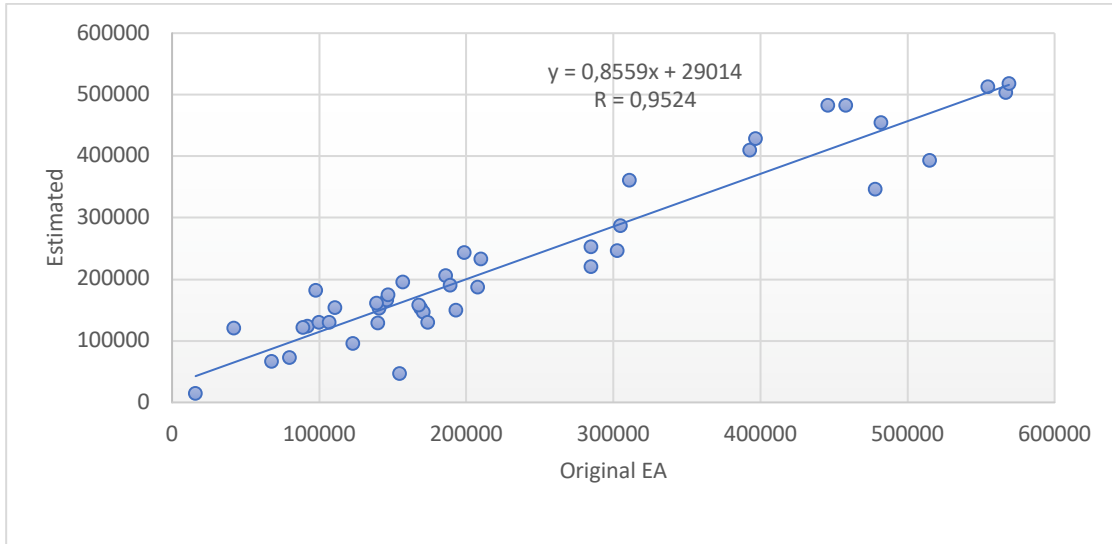


Figure 3. Scatter Plot of Original Output Versus Estimated EA by ANN.

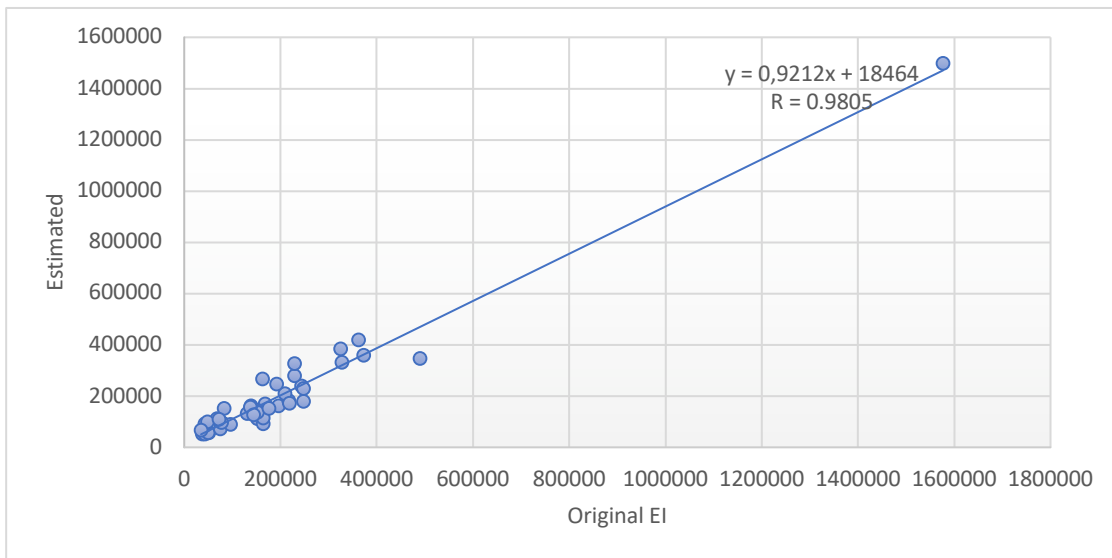


Figure 4. Scatter Plot of Original Output Versus Estimated EI by ANN.

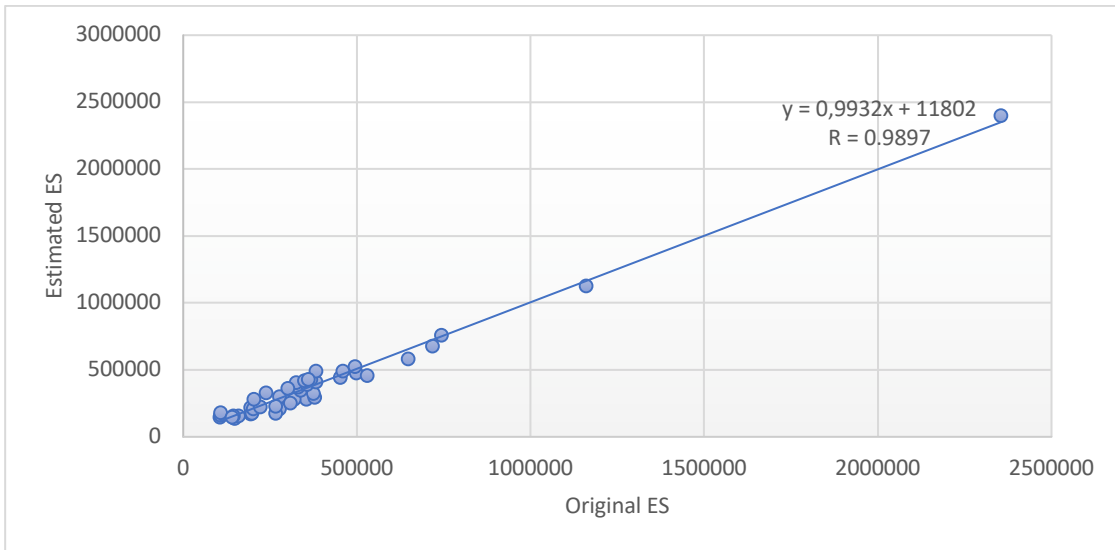


Figure 5. Scatter Plot of Original Output Versus Estimated ES by ANN.

According to Fig. 3, Fig. 4 and Fig. 5 the weakest relationship between the original employment values and estimated employment values were observed in the agriculture sector ($R=0,9524$). Furthermore, the strongest relationship was observed in the services sector ($R=0,9879$).

In Figures 6,7 and 8, line charts of EA, EI and ES were given. In these charts, the harmony of the original and estimated values with each other was demonstrated. According to these three charts, there was a high degree of harmony between the original values and the estimated values. This high degree of harmony indicated the success of the ANN model that was produced in the study.

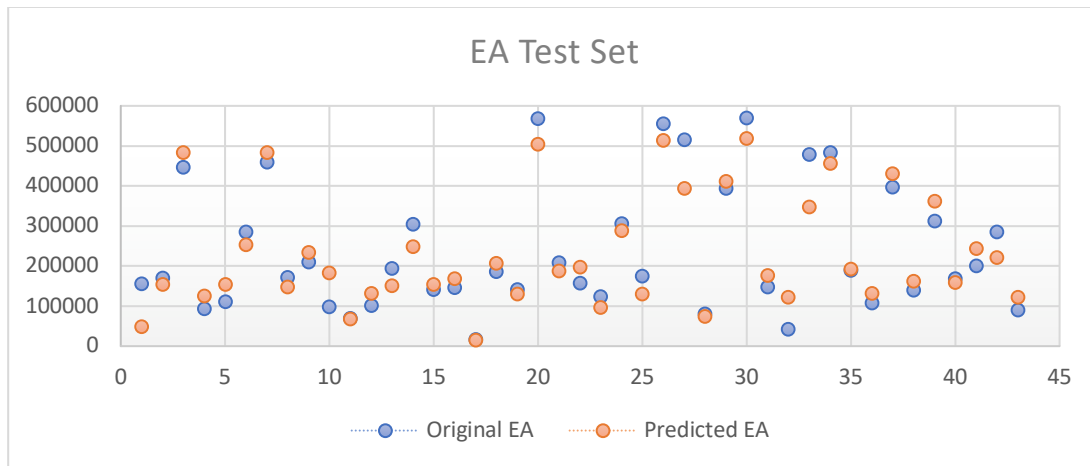


Figure 6. Estimated and Original Output EA in the Test Set for the LM Algorithm for ANN.

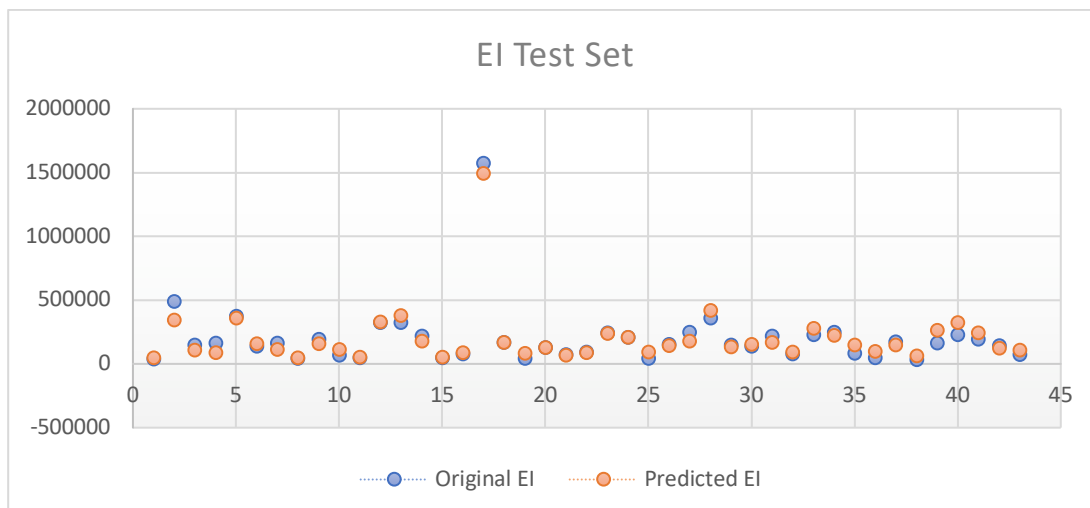


Figure 7. Estimated and Original Output EI in the Test Set for the LM Algorithm for ANN.

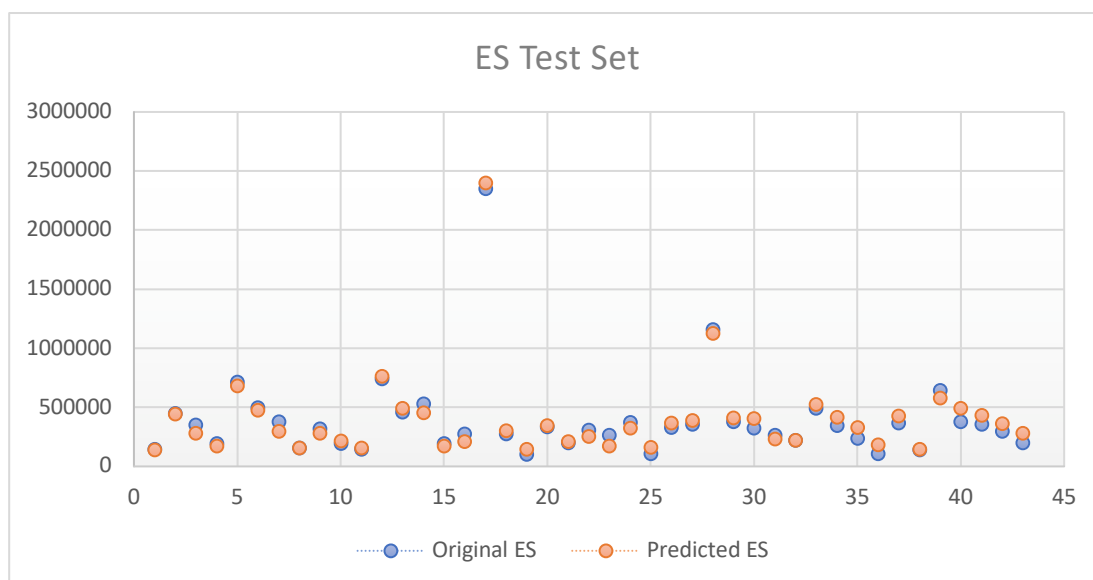


Figure 8. Estimated and Original Output ES in the Test Set for the LM Algorithm for ANN.

5.3. Analysis of the MLR's Predictive Performance of EA, EI and ES volumes

To perform a statistical technique estimation using the MLR model, firstly all data is divided into ratios of 15% and 70%. The statistical parameters of the MLR model are calculated as demonstrated in the following formula which is based on the number of independent variables.

$$EA = 0,200 + (0,00114 * Region) + (0,551 * GDPA) - (0,702 * GDPI) + (0,418 * GDPS)$$

$$EI = 0,0955 - (0,00338 * Region) - (0,0942 * GDPA) + (0,841 * GDPI) + (0,271 * GDPS)$$

$$ES = 0,0690 - (0,00268 * Region) - (0,0149 * GDPA) - (0,191 * GDPI) + (1,297 * GDPS)$$

The independent variables of the 15% division were applied as input data to the MLR model and estimated data were obtained as the MLR model output. The errors (NMSE) and correlation coefficient between the original values and the estimated values are 0.0197 and 0.681 for EA; 0.0011 and 0.947 for EI; 0.00034 and 0.980 for ES, respectively.

As seen in Table 7, Table 8 and Table 9 the pairwise comparisons between the estimated EA, EI and ES values and the original EA, EI, and ES values were evaluated using t-tests. There was no

significant difference between the estimated EA, EI and ES values and the original EA, EI, and ES values ($p>0.05$). This result demonstrates the high estimation performance of the MLR models.

Table 7. The Relationship Between the Estimated EA Values and the Original EA Values is Demonstrated by Using the T-Test as the Statistical Technique for MLR.

Group Name	N	Missing	Mean	Std Dev	SEM	P Value
Original EA	43	0	236441,860	156548,532	23873,427	0,959
Estimated EA	43	0	235059,573	75612,076	11530,733	

Table 8. The Relationship Between the Estimated EI and the Original EI is Demonstrated by Using the T-Test as the Statistical Technique for MLR.

Group Name	N	Missing	Mean	Std Dev	SEM	P Value
Original EI	43	0	196627,907	239177,561	36474,236	0,743
Estimated EI	43	0	180652,780	210266,255	32065,303	

Table 9. The Relationship Between the Estimated ES and the Original ES is Demonstrated by Using the T-Test as the Statistical Technique for MLR.

Group Name	N	Missing	Mean	Std Dev	SEM	P Value
Original ES	43	0	388395,349	365368,121	55718,116	0,942
Estimated ES	43	0	382718,910	354066,703	53994,666	

Relationships between the original EA, EI and ES values and estimated EA, EI, and ES values were given in Fig. 9, Fig. 10 and Fig. 11. There was asignificant relationshipbetween these variables ($R=0.681$ for EA; $R=0.947$ for EI; $R=0.980$ for ES).

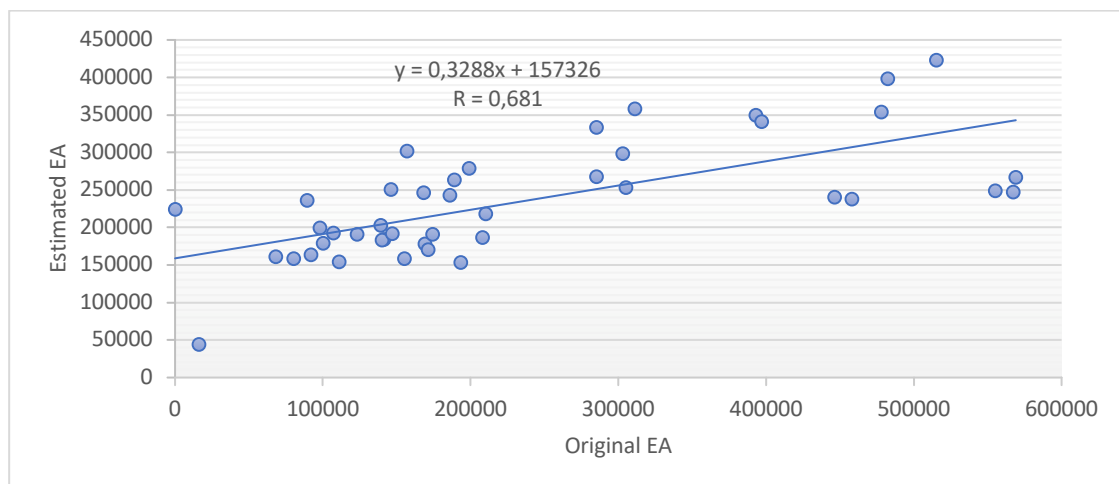


Figure 9. Scatter Plot of Original Output Versus Estimated EA by MLR.

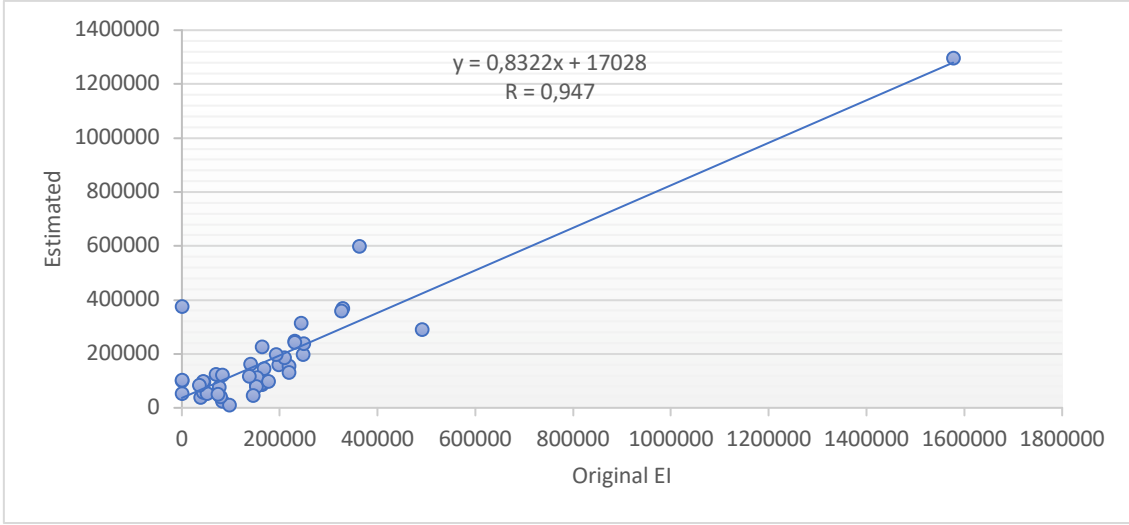


Figure 10. Scatter Plot of Original Output Versus Estimated EI by MLR.

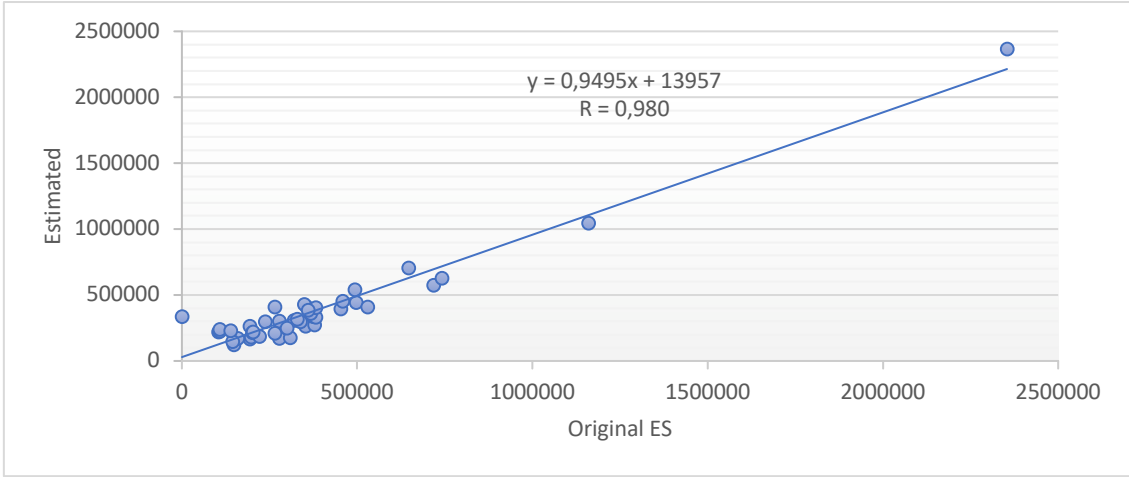


Figure 11. Scatter Plot of Original Output Versus Estimated ES by MLR.

According to Fig. 9, Fig. 10 and Fig. 11 the weakest relationship between the original employment values and estimated employment values were observed in the agriculture sector ($R=0,681$), and the strongest relationship was observed in the services sector ($R=0,980$).

In Figures 12, 13 and 14 the linechart of EA, EI and ES was given. In these charts, the harmony of the original and estimated values with each other was demonstrated. According to these three charts, there was a high degree of harmony between the original values and the estimated values. This high degree of harmony indicated the success of the MLR model that was produced in the study.

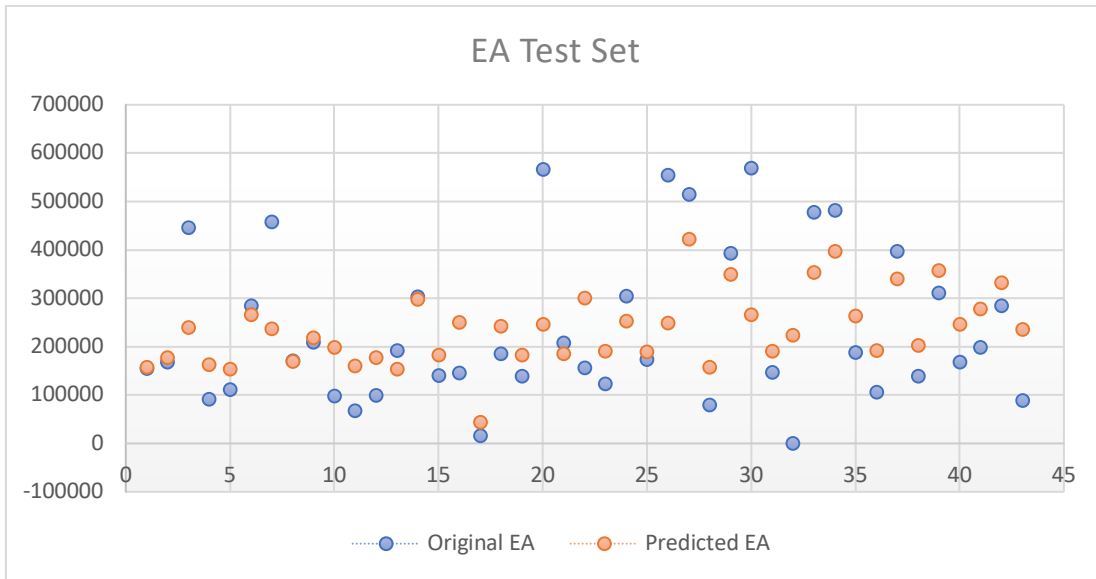


Figure 12. Estimated and Original Output EA in the Test Set for MLR.

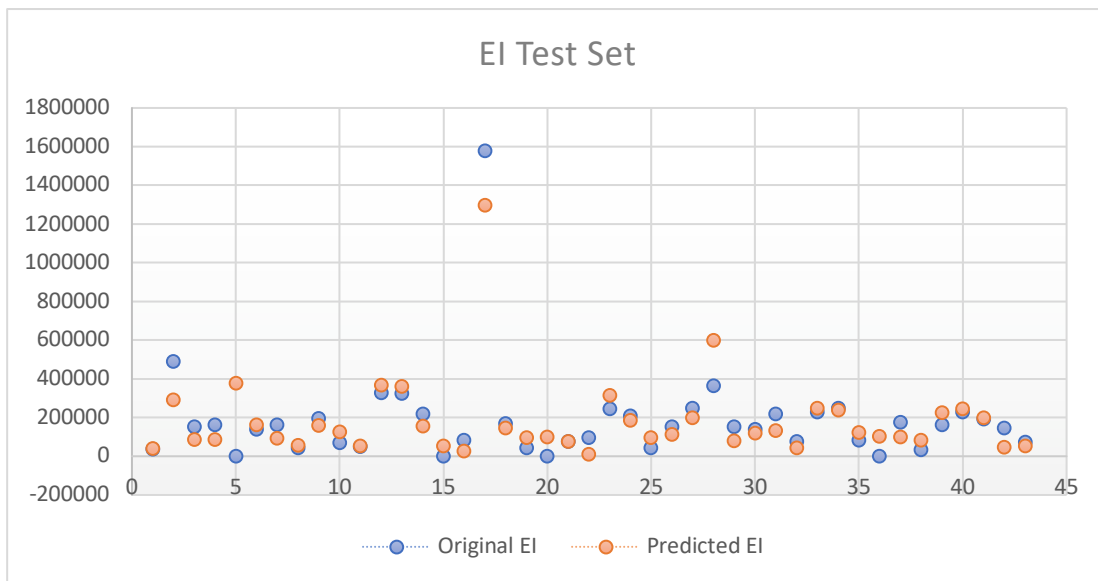


Figure 13. Estimated and Original Output EI in the Test Set for MLR.

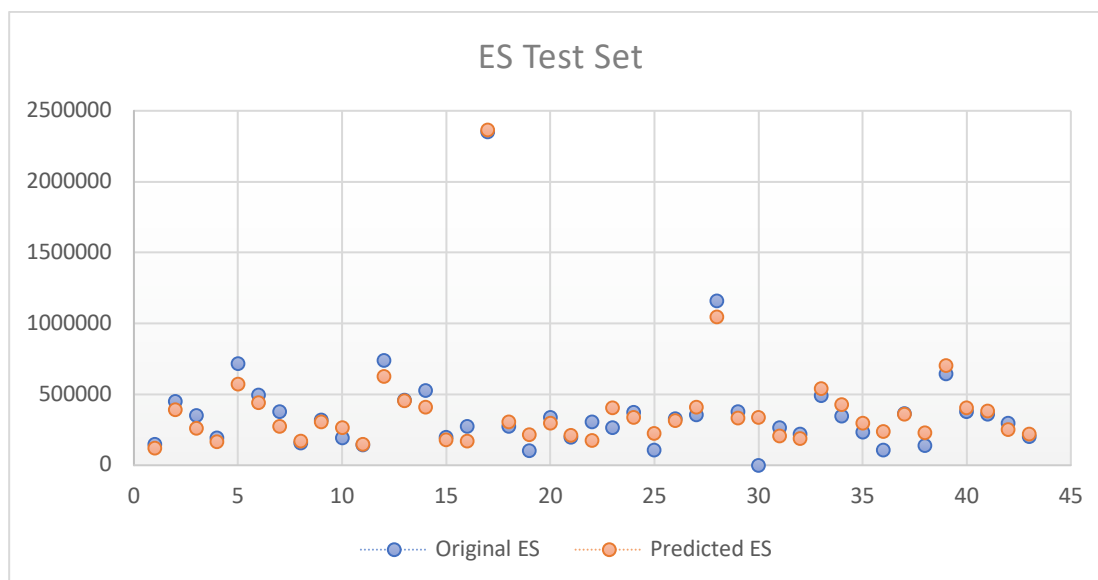


Figure 14. Estimated and Original Output ES in the Test Set for MLR.

5.4. Comparison of Predicting Capabilities of the MLR and ANN Models

When Table 10 was examined for the performance of MLR and ANN (4-10-3, Tansig-Tansig, LM) models, the NMSE values of ANN (4-10-3, Tansig-Tansig, LM) models were 0.003302, 0.0004195 and 0.0001828 for EA, EI, and ES, respectively. The NMSE values of the MLR (4-10-3, Tansig-Tansig, LM) models were 0,0197, 0,0011, 0,00034 for EA, EI and ES, respectively. Likewise, the R-values of the ANN models were 0.9524, 0.9805 and 0.9897 for EA, EI, and ES, respectively. The R-values of the MLR models were 0.681, 0.947 and 0.980 for EA, EI and ES, respectively. These results put forward that ANN had better predictability of EA, EI, and ES than MLR.

6. CONCLUSIONS

In this study, the ANN and MLR models were used to develop expert systems. These systems aim at predicting employment data using GDP data. The estimated performances of MLR and ANN on EA, EI, and ES have been examined.

When the findings obtained by ANN and MLR methods are compared, it is clearly seen that ANN application is superior to MLR application in terms of classification success. While the classification accuracy of ANN is over 95% for each of EA, EI and ES, the application accuracy of MLR method is not high for all EA, EI and ES. Unlike the classical linear modeling approach, ANN, which handles all factors and effectively performs analysis and classification in the prediction process after data training, has been obtained as an ideal model (Table 10).

One of the most important factors in ANN success can be the training process using the data and the ability to adapt the input values to the output values by constantly renewing itself after this process and trying to reduce the error rate.

However, in ANN, the behavior of the network cannot be explained especially in hidden layers. There is no clear information about how the solution to a problem is realized. However, it is clearly seen in the results obtained that this situation does not cause any restrictions in the approach to the global optimum point.

The biggest limitation of MLR is its search for linearity relationship between data. If there is a linear relationship between the data, a less complex equation model will be created with dependent and independent variables, and this will enable the estimation process to be completed in a shorter time.

As a result, statistical analyses were performed on the obtained data, which indicated that ANN showed better adaptability and estimation performance than MLR when working with non-linear variables. It would be possible to design a specific system using a model developed by ANN. In this way, it would be possible to estimate the employment figures closest to the actual data by using GDP data, thereby saving time and money.

When the results with employment rates closest to reality obtained by using artificial neural networks in this study are evaluated, the use of artificial intelligence-based methods in new studies in the field of macroeconomics is recommended for both researchers and practitioners.

Since the GDP variable is closely related to the many macroeconomic variables, through the GDP variables, making a successful forecast of employment data is quite important in terms of predictability of other macro economic variables. In this way, methods that give them more realistic results can be developed and these methods added to the literature. Simultaneously, it will be possible to support policy makers in the economy.

REFERENCES

- Azadi, S. and Sepaskhah, A. R. (2011). Annual Precipitation Forecast For West, Southwest, And South Provinces of Iran Using Artificial Neural Networks. *Theoretical and Applied Climatology*, 109 (1-2), 175–189.
- Azadi, S. and Karimi-Jashni, A. (2016). Verifying The Performance Of Artificial Neural Network And Multiple Linear Regression In Predicting The Mean Seasonal Municipal Solid Waste Generation Rate: A Case Study Of Fars Province, Iran. *Waste Management*, (48), 14–23.
- Abdioglu, Z. and Albayrak, N. (2017). Jobless Growth: A Research on Sub-Sectors. *Karadeniz Technical University Institute of Social Sciences Journal of Social Sciences*, 7 (13), 215-228.
- Beale, M. H., Hagan, M. T. and Demuth, M. H. (2010). *Neural Network Toolbox™ 7 User's Guide*. Natick, MA: MathWorks Inc.
- Ertugrul, H. M. and Ucak, A. (2013). The Relationship Between Economic Growth and Employment: A Dynamic Empirical Analysis For Turkey. *International Journal of Human Sciences*, 10 (1), 661-675.
- EUROSTAT (2018). *Definition of Employment Level: European Statistical Office (EUROSTAT) Web*. [Online] <https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Employment>, (Accessed Date: 13.11.2018).
- Fukuoka, Y., Matsuki, H., Minamitani, H. and Ishida, A. (1998). A Modified Back-Propagation Method To Avoid False Local Minima. *Neural Networks*, 11 (6), 1059–1072.
- Fuchs, V.R. (1980). *Economic Growth and The Rise of Service Employment*. NBER Working Paper No. 486.
- Flott, L.W. (2012). Using The Scatter Diagram Tool to Compare Data, Show Relationship. *Metal Finishing*, 110 (8), 33–35.
- Jahandideh, S., Jahandideh, S., Asadabadi, E.B., Askarian, M., Movahedi, M.M., Hosseini, S. and Jahandideh, M. (2009). The Use Of Artificial Neural Networks and Multiple Linear

- Regression To Predict Rate Of Medical Waste Generation. *Waste Management*, 29 (11), 2874–2879.
- Korkmaz, S. and Yilgor, M. (2010). The Relationship Between Growth and Employment After 2001 Crisis In Turkey. *Journal of Management and Economics Research*, 8 (14), 169-176.
- Kocadagli, O. (2015). A Novel Hybrid Learning Algorithm For Full Bayesian Approach Of Artificial Neural Networks. *Applied Soft Computing*, (35), 52–65.
- Murat, S. and Yılmaz Eser, B. (2013). The Relationship Between Economic Growth and Employment in Turkey: The Validity Of Jobless Growth Phenomenon. *HAK-İŞ International Journal of Labour and Society*, (2) 3, 92-123.
- Muratoglu, Y. (2011). *The Relationship Between Growth and Employment: The Case of Turkey: Eurasian Economies Association Web.* [Online] <https://www.avekon.org/proceedings/avekon02.pdf>. (Accessed Date: 22.11.2019).
- Okun, A. M. (1963). *Potential GNP: its measurement and significance*. New Haven: Yale University.
- Wang, Y., Li, J., Gu, J., Zhou, Z. and Wang, Z. (2015). Artificial Neural Networks For Infectious Diarrhea Prediction Using Meteorological Factors in Shanghai (China). *Applied Soft Computing*, (35), 280–290.
- Wordbank (2018a). *Definition of Employment.* *WordbankWeb.*[Online] <https://datacatalog.worldbank.org/employment-services-total-employment-modeled-ilo-estimate-0>. (Accessed Date: 23.11.2019).
- Wordbank (2018b). *Definition of Gross Domestic Product (GDP).* *WordbankWeb.*[Online] <https://wits.worldbank.org/CountryProfile/Metadata/es/Indicator/Development>. (Accessed Date: 18.11.2019).
- Sarica, S., Çetin, A. K. and Kutluturk, M. M. (2015). The Impacts on Employment of Real Wage And GDP: Time Series Analysis. *Journal of Academic Social Science*, 2 (14), 123-138.
- Sawtelle, B. (2007). Analysing the Link Between Real GDP And Employment: An Industry Sector Approach. *Business Economics*, 42 (4), 46-54.
- Swane, A. And Vistrand, H. (2006). *Jobless Growth In Sweden? A Descriptive Study: Stockholm School Of Economics web.* [Online] <http://arc.hhs.se/download.aspx?mediumid=151>. (Accessed Date: 5.11.2019).
- Shu, Y. and Lam, N.S.N. (2011). Spatial Disaggregation of Carbon Dioxide Emissions From Road Traffic Based on Multiple Linear Regression Model. *Atmospheric Environment*, 45 (3), 634–640.
- TSI (2018a). *Income and Living Conditions Survey Micro Data Set (Cross Sectional) 2014: Turkey Statistical Institute (TSI) Web.* [Online] http://www.tuik.gov.tr/MicroVeri/GYKA_2014/turkce/metaveri/siiniiflamalar/index.html. (Accessed Date: 20.03.2018).
- TSI (2018b). *Definition of Gross Domestic Product (GDP): Turkey Statistical Institute (TSI) Web.* [Online] http://www.tuik.gov.tr/PreTablo.do?alt_id=1105. (Accessed Date: 5.11.2018).
- TSI (2018c). *Employment Data of Turkey Between 2004 and 2014: Turkey Statistical Institute (TSI) Web.* [Online] http://www.tuik.gov.tr/PreTablo.do?alt_id=1007. [Accessed Date: 10.03.2018].
- TSI (2018d). *Gross Domestic Product Data of Turkey Between 2004 and 2014: Turkey Statistical Institute (TSI) Web.* [Online] http://www.tuik.gov.tr/PreTablo.do?alt_id=1108. (Accessed Date: 10.03.2018).
- Zhang, G., Eddy Patuwo, B. and Hu, M. Y. (1998). Forecasting with Artificial Neural Networks: The State of the Art. *International Journal of Forecasting*, 14 (1), 35–62.

Appendix 1.**Table 1.** Classifications and the Variables That Constitute the Classification in The Study.

Classifications	Subclassifications	The Range and Variety of Variables	
First Classification	TÜİK Level-2 Classification (Regions)	TR10	Istanbul
		TR21	Tekirdag, Edirne, Kırklareli
		TR22	Balikesir, Canakkale
		TR31	Izmir
		TR32	Aydın, Denizli, Mugla
		TR33	Manisa, Afyon, Kutahya, Usak
		TR41	Bursa, Eskisehir, Bilecik
		TR42	Kocaeli, Sakarya, Duzce, Bolu, Yalova
		TR51	Ankara
		TR52	Konya, Karaman
		TR61	Antalya, Isparta, Burdur
		TR62	Adana, Mersin
		TR63	Hatay, Kahramanmaraş, Osmaniye
		TR71	Kırıkkale, Aksaray, Nigde, Nevsehir, Kırsehir
		TR72	Kayseri, Sivas, Yozgat
		TR81	Zonguldak, Karabük, Bartın
		TR82	Kastamonu, Çankırı, Sinop
		TR83	Samsun, Tokat, Çorum, Amasya
		TR90	Trabzon, Ordu, Giresun, Rize, Artvin, Gümüşhane
		TRA1	Erzurum, Erzincan, Bayburt
		TRA2	Ağrı, Kars, Iğdır, Ardahan
		TRB1	Malatya, Elazığ, Bingöl, Tunceli
		TRB2	Van, Mus, Bitlis, Hakkari
		TRC1	Gaziantep, Adıyaman, Kilis
		TRC2	Şanlıurfa, Diyarbakır
		TRC3	Mardin, Batman, Sırnak, Siirt

Second Classification	GDP	GDP of the Agriculture Sector (GDPA)	494436758-11894286316 (Minimum-Maximum Values)
		GDP of the Industry Sector (GDPI)	310246133-166397951487 (Minimum-Maximum Values)
		GDP of the Services Sector (GDPS)	1857014512-383491347677 (Minimum-Maximum Values)
	ED	Employment in the Agriculture Sector (EA)	11000 – 680000 (Minimum-Maximum Values)
		Employment in the Industry Sector (EI)	15000 – 1870000 (Minimum-Maximum Values)
		Employment in the Services Sector (ES)	78000 – 3198000 (Minimum-Maximum Values)

Appendix 2.

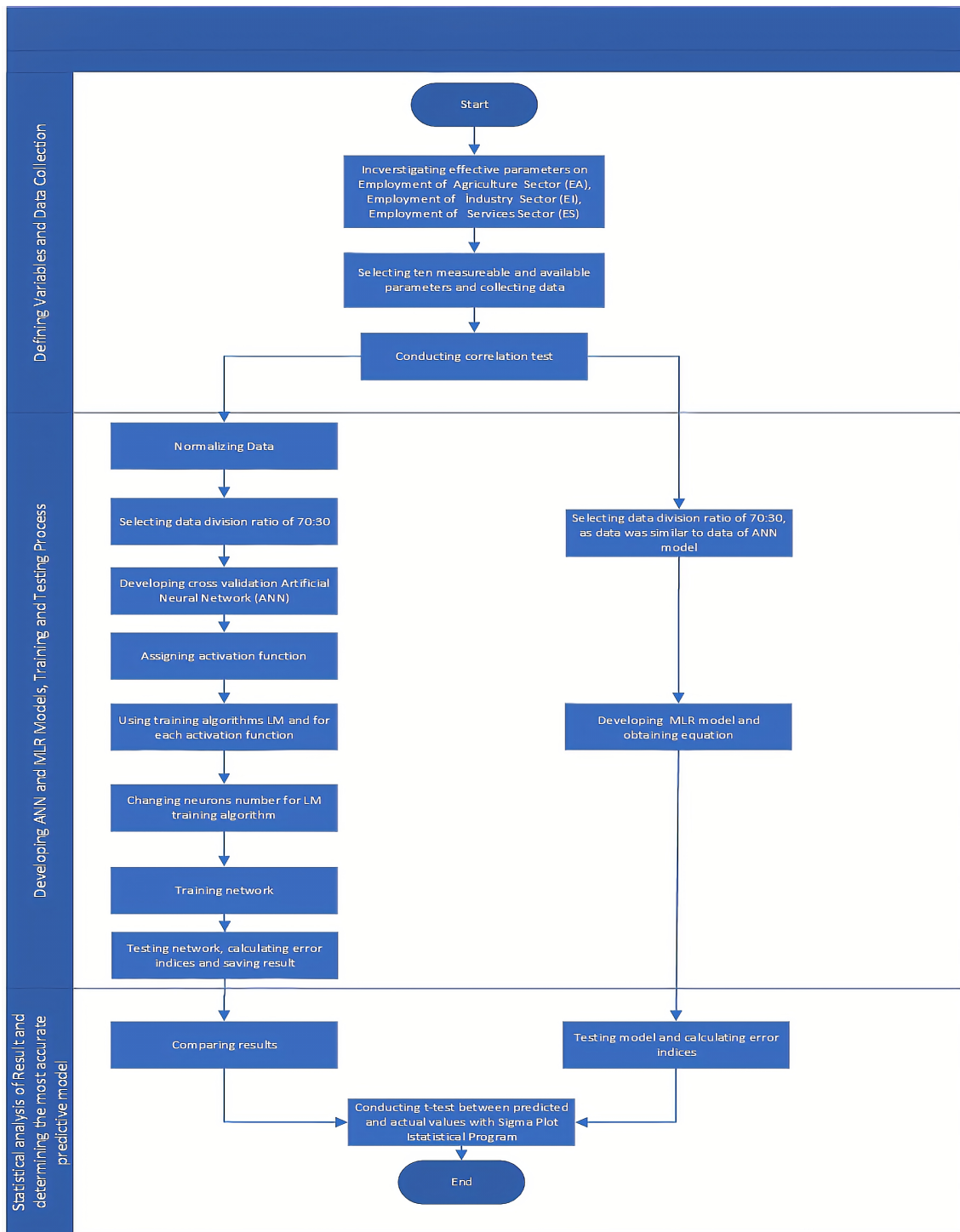


Figure 1. The Flow Process was Determined to Predict EA, EI, and ES (Azadi and Karimi-Jashni, 2016: 16).

Appendix 3.

Table 2. Calculated Correlation Coefficients Between EA, EI, ES and Four Independent Variables.

	Region	GDPA	GDPI	GDPS
EA	0,0332	0,564	-0,284	-0,274
EI	-0,536	-0,127	0,924	0,903
ES	-0,505	-0,0867	0,93	0,957

*The Sigma Plot Statistical Program software was used to determine the correlation coefficients.

Appendix 4.

Table 3. Performance of ANN Models in Training and Testing Situations.

Training Algorithms	Layers and Number of Neurons	Training Correlation Coefficient(R)	Testing Correlation Coefficient (R)	MSE (Mean Squared Error)	Epoch	Activation Function
Levenberg–Marquardt (LM) (trainlm) for Employment of Agriculture Sector	4-10-3	0.9589	0.9524	0.003302	1000	Tansig-Tansig-
Levenberg–Marquardt (LM) (trainlm) for Employment of Industry Sector	4-10-3	0.9926	0.9805	0.0004195	1000	Tansig-Tansig
Levenberg–Marquardt (LM) (trainlm) for Employment of Services Sector	4-10-3	0.9925	0.9897	0.0001828	1000	Tansig-Tansig-

Appendix 5.

Table 10. Performance Measurements for Predictive Capabilities of the MLR and ANN Models.

Models	R	NMSE
Levenberg– Marquardt (LM) (trainlm) for Employment of Agriculture Sector	0.9524	0.003302
Levenberg– Marquardt (LM) (trainlm) for Employment of Industry Sector	0.9805	0.0004195
Levenberg– Marquardt (LM) (trainlm) for Employment of Services Sector	0.9897	0.0001828
MLR for Employment of Agriculture Sector	0,681	0,0197
MLR for Employment of Industry Sector	0,947	0,0011
MLR for Employment of Services Sector	0,980	0,00034