Research Article

Estimation Of Turkey's Carbon Dioxide Emission with Machine Learning

Ahmet ÇOŞGUN

Akdeniz University, Faculty of Engineering, Department of Mechanical Engineering Department of Thermodynamics, Main Campus, Antalya/Turkey
Corresponding Author Email: acoskun@akdeniz.edu.tr - **ORCID:** 0000-0002-0243-5476

Carbon dioxide emissions are an important factor in the increase of greenhouse gases in the atmosphere and climate change. Controlling and reducing carbon dioxide emissions plays an important role in combating global warming and climate change. Various national and international efforts are being carried out to reduce greenhouse gas emissions and switch to sustainable energy sources. For this reason, estimating carbon dioxide emissions in the coming years is important for determining the measures to be taken.
In this study, Turkey's carbon dioxide emissions are successfully estimated using two different machine learning models. The success of the study was evaluated using three different statistical measures: R2, MSE and MAE. The R2 of decision trees was 89.4%, MSE was 0.013 and MAE was 0.011; the R2 of artificial neural networks was 92.7%, MSE was 0.009 and MAE was 0.006. When we compare the two models, it is seen that ANN is more successful than decision trees and predicts with less error.

1. Introduction

One of the biggest problems facing humanity today is air pollution. As the population increases, urban development and growth increases, and industry develops, air pollution becomes increasingly important. In general, the harmful effects of air pollutants on humans, living organisms and the environment show complex distribution patterns depending on time, space, exposure time, concentration and other characteristics. This complexity makes it difficult to model or measure the patterns and trends of pollutants and to estimate the levels to which people are exposed. One of the most important steps to prevent air pollution is to model and assess the level of pollution [1]. If air quality is measured regularly, it is possible to determine the level of pollution in the region. In this way, air pollution control plans can be developed, air pollution maps can be created and distribution models can be developed. In this context, solutions to improve air quality and standards can be created according to the results obtained from air quality measurements in a healthier, more realistic and simpler way [2].

Machine learning, one of the latest artificial intelligence technologies, has produced objective and more precise results on air pollution. Machine learning is the general name for computer algorithms that can independently learn the solution to the problem in question through complex pattern recognition and data-driven decision making. The model created using existing data sets and machine learning methods is based on achieving maximum performance. When the literature is examined; it is seen that forecasting methods for air pollution have been developed using various combinations of NO2, NO, O3, CO, SO2, PM2.5, PM10 data sets [3,4,5,6,7,8].
In this study, unlike the literature, machine learning models were created to estimate Turkey's carbon footprint between 1990 and 2023. For this purpose, decision tree regression and artificial neural network methods were used.

2. Material and Methods

In this study, machine learning models were created to estimate Turkey's carbon footprint. For this purpose, decision tree regression and artificial neural network methods were used. In the study, 33 annual data points between 1990 and 2023 were used.

2.1. Data Set

The independent variables of the study are population, gross domestic product ($), electricity consumption per capita (kwh), renewable energy in total energy (%), coal in total energy (%), natural gas

in total energy (%), liquid fuels (gasoline, diesel, etc.) in total energy (%), number of internal combustion engine vehicles, and forest area (hectare). While designing the model for estimating carbon emissions, variables that can affect the change in carbon emissions for Turkey were selected as input units. First of all, demographic and economic growth are thought to increase carbon emissions. The increase in population increases both energy consumption and consumption over time, thus increasing production in parallel. Especially countries with high populations, such as Turkey, China, India, etc., cannot slow down carbon emissions because they cannot fully utilize renewable energy sources yet. Therefore, population, gross domestic product, and electricity consumption per capita are included as input units in the model. Whether the energy used in the country is obtained from fossil-based or renewable energy sources also has a positive relationship with carbon emissions. The use of coal, which is a fossil-based energy source and one of the highest sources of carbon emissions, especially for electricity generation, industry, and residential heating, causes countries to emit significant carbon emissions. Increasing the use of renewable energy sources, which are defined as clean energy sources, is included in the model as it is thought to cause countries to reduce carbon emissions over the years. Another variable that affects carbon emissions is the rate of increase in the number of vehicles using internal combustion fossil fuels over time, depending on the population, the growth of the logistics sector using heavy vehicles due to the growth of the economy, and the level of prosperity. Although the engine technology of these internal combustion engine vehicles has improved over time and the use of electric motorized cars has increased, the increasing number of internal combustion engine vehicles in traffic is included in the model as it is thought to increase carbon emissions. Forests are known to be the largest carbon storage areas. The stored carbon is emitted into the atmosphere due to man-made or natural events. Therefore, an increase in forest areas over time slows down the release of carbon into the atmosphere, while a decrease increases it. For this reason, the area covered by forest areas was added to the model as the last output unit.

2.2. Data pre-processing

In machine learning, the correct processing and preparation of data is one of the most important factors determining the quality of the results. Data collected from multiple sources is often found in an unorganized form. This affects the prediction performance of the models. Therefore, the raw data must be modified before training, evaluating, and using machine learning models. Data pre-processing is a series of operations for cleaning, transforming, organizing, and preparing data. In raw data, there is a situation of having incorrect values, such as misspelling, corruption, duplication, etc. Data cleaning is defined as the process of detecting and correcting corrupt or incorrect data, especially in large data sets [9]. It makes the resulting dysfunctional data useful by correcting it. After the scattered, noisy, corrupted, or incorrect observations are identified, the following processes are applied [10, 11]

- Use summary statistics to identify normal data and identify outliers.
- Identification and removal of columns with the same value or variance
- Identify duplicate data rows and remove them.
- Marking empty values as missing
- Fill in the missing values using statistics or a trained model.

2.3. Normalization

The normalization process is applied to the raw data and has an effect on the preparation of a suitable data set for training. The scaling of the inputs and outputs of the model (normalization) closely affects the performance of the network. Without applying the normalization method to the raw data set, it can be very slow. Because normalization regularizes the distribution of values in the data set, excessively large or small values may appear in the data set. When calculating the net inputs, these values may mislead the network by causing excessively large or small values. The fact that some values on the same data set have values less than 0 and some have larger values shows that these distances between the data, especially the extreme data, will be more effective on the results. By normalizing the data, it is ensured that each parameter in the training input set contributes equally to the prediction process of the model [12,13].

Scaling all inputs in a certain range (mostly in the range 0-1) causes both the information coming from different environments to be reduced to the same scale and the effect of incorrectly entered very large and small values to be eliminated. Different techniques can be used in normalization processes. There are many types of data normalization in the literature. These can be listed as rules such as the Min-Max rule, median, sigmoid, and Z-score. In this study, the data were normalized between 0 and 1 using the min-max method [14].

$$x' = \frac{x_1 - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

Where x^1 is the normalized data, x_1 is the input value, x_{min} is the smallest value in the data set, and x_{max} is the largest value in the data set.

2.4. Machine Learning Techniques

Machine learning makes predictions from the learned data based on known features using mathematical and statistical methods. Prediction and classification operations can be performed using machine learning. If the system output of the input data is quantitative, it is called prediction, and if it is qualitative, it is called classification. Machine learning algorithms are divided into two groups: supervised learning and unsupervised learning. Supervised learning algorithms apply what they have learned to new data to predict the value of the dependent variable by learning the relationship and rules between events using labelled data. An unsupervised machine learning algorithm is used when the data is not labelled. In this algorithm, the data is analysed and focuses on finding a relationship between them. Unsupervised machine learning has a system structure that focuses on data [15, 16].

2.5. Decision Trees Regression

Tree-based learning algorithms are among the most widely used supervised learning algorithms. A decision tree is a structure used to partition a dataset containing a large number of records into smaller sets by applying a set of decision rules. Decision tree algorithms are used in both classification and regression. Regression is used on numerical target data, while classification is used on categorical data [17]. A decision tree is a flowchart-like tree structure where each internal node represents a feature (or attribute), each branch represents a decision rule, and each leaf node represents the result. The top node in a decision tree is known as the root node. Nodes that do not have any child nodes, i.e., nodes where the data cannot be further divided, are called leaf nodes. Leaves contain the final prediction for the observations that reach them. The decision tree algorithm starts from the root node and proceeds along the tree, making a decision based on the input feature values, until it reaches a leaf node. The value at the leaf node represents the predicted output value [18].
They have a predefined dependent variable. It categorizes the independent variables into intervals according to the knowledge gained. When asked for a value from this interval during prediction, it answers with the average value in this interval

(learnt during training). For this reason, decision tree regression is not continuous like other regression models but discrete. Decision trees try to maximize information gain by making choices that reduce the entropy (degree of randomness) of the current situation. For this purpose, it recalculates the error function for each question (node or node) and selects the question or situation with the lowest error [19].

2.6. Artificial Neural Networks

Artificial neural networks are an artificial intelligence technique inspired by the information processing technique of the human brain. It has been developed with the aim of automatically realizing abilities such as deriving new information, creating new information, and discovering new information through learning, which are the characteristics of the human brain, without any help [20].
The model inspiration of artificial neural networks is expressed as neuron structures and activities in the human brain. Just as neuron cells come together to form the human nervous system, artificial neurons come together to form artificial neural networks with the connections they establish [21]. Artificial neural networks consist of an input layer, one or more hidden layers, and an output layer, as shown in Figure 1.
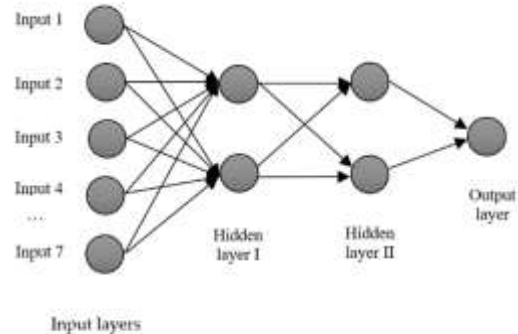

Figure 1. *ANN model*

The ability to learn from an information source is one of the most important features of ANN. Artificial neural networks are trained on appropriate data sets for the problem to be solved. The goal of this training is to find the most appropriate values for the parameters in the ANN model. These parameters are the weight values and threshold values of the neurons. Our parameters are continuously updated throughout the training. In artificial neural networks, information is stored in the weights of the connections of the neurons in the network. Therefore, it is important to determine the weights. Since the information is stored in the whole network, the weight value of a node alone does not mean

anything. The weights in the whole network should have optimal values. The process to reach these weights is called "training the network.". Accordingly, for a network to be trainable, the weight values must be dynamically changeable within a certain rule [22,23].

The activation function provides the curvilinear matching between input and output units (layers). The correct choice of the activation function significantly affects the performance of the network. The activation function can generally be chosen as unipolar (0–1), bipolar (-1–1), or linear. It is the component that enables the network to learn the non-linear structure [24].

Artificial neural networks are divided into two categories: feed-forward and feedback networks, according to their learning methods. Feed-forward networks are artificial neural network models in which one-way signal flow occurs; incoming data signals are processed on the input and output layers and then transferred to the output. A feedback neural network model is a network structure in which the outputs in the output and intermediate layers are fed back to the input units or previous intermediate layers. Thus, inputs are transferred in both forward and backward directions [25].

2.7. Evaluation of the models

In the study, the three most commonly used metrics in the literature were preferred for the evaluation and comparison of the models. The determination coefficient (R^2), MSE, and MAE values of the models were analyzed, and successful and ideal models were decided.

The R^2 metric shows how appropriate the independent variables in the data set are in predicting the value of the dependent variable. Therefore, it expresses the predictive power of the regression. It is the ratio of the sample variance of the dependent variable to the part explained by the independent variable. The R^2 is between 0 and 1. For the R^2, being equal to 1 indicates that all of our prediction values are equal to the actual values. However, a situation such as the R^2 being equal to 1 is very likely to indicate that the model is memorized. If the dependent variable prediction values are given as the average of the actual dependent variable values, the R^2 value is equal to 0. An R^2 of 0 appears as the worst prediction condition. In other words, giving the average of the actual dependent variables to all dependent variable predictions is the worst possible prediction [21]. R^2 is shown in Equation 2. MSE (mean squared error) is a measure of success evaluation. It provides the opportunity to evaluate the differences between actual values and predicted

$$R^2 = 1 - \sum_i^n \frac{\left(y_{i_{actual}} - y_{i_{prediction}}\right)}{\left(y_{i_{actual}} - Average(y_{i_{actual}})\right)^2} \quad (2)$$

values. Especially when using optimization methods, it is a function that is tried to be optimized to build better models. It is the average error between the predicted values predicted by a machine learning model and the actual values. It is always positive, and it can be said that predictors with MSE values close to zero perform better [26]. MSE is shown in Equation 3.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} e_i \quad (3)$$

MAE (mean absolute error) is the mean of the absolute values of the errors between actual values and predicted values. This metric works directly on the absolute values of the errors. It is a more direct representation of the sum of the error terms. A low MAE is expected. As the MAE approaches 0, the performance of the model is interpreted as better [27]. MAE is shown in Equation 4.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |e_i| \quad (4)$$

3. Results and Discussions

The data set used in the study covers the period between 1990 and 2023. The data set contains 33 data points, consisting of annual data. Machine learning models consisting of 8 independent variables and 1 dependent variable were created to estimate Turkey's annual carbon footprint. Two machine learning methods, decision trees and artificial neural networks, were used in the study. The basic working structure of the study is shown in Figure 2.

Figure 2. *Main stages of the study*

In the first stage of the study, data pre-processing was carried out for better performance of the raw data. At this stage, inconsistent, incomplete, and noisy data were reviewed, and corrections were made. Thus, the performance of the system has increased, including the success and speed of the system.

In the second stage of the study, the effect of independent variables on the dependent variable was analysed. The correlation relationship between the variables was analysed. Correlation analysis is a

statistical method that provides information about the relationship between variables and the direction and severity of this relationship. The correlation value is between -1 and 1. A value of 0 means that there is no correlation between the variables. As this value moves towards -1, it means 'Negative Perfect Correlation', and as it moves towards +1, it means 'Positive Perfect Correlation'. In the study, it was observed that the relationship between the variables is medium- to high-severity and has a positive correlation [28].

In the next stage of the study, a normalization process was performed. In order to increase the performance and success of the models, the data were normalized using the min-max method. By normalizing the data between 0 and 1, the training and testing times were reduced to the optimum time. In the study, the data set is divided into two parts: training and testing. The training data set is the data set on which the model is trained. Here, the most optimal parameters (coefficients and weights) are set for the models. The test set is the data set used to evaluate the model developed in the training set. The larger the training set, the better the model will learn, and the larger the test set, the more reliable the evaluation metrics will be. In the study, training and test data were separated at different rates (60%−40%, 70%−30%, 80%−20%), and the most successful rate was achieved when 70% training and 30% testing were separated. In order to increase the accuracy of the model, training and test data were selected according to the random sampling method.

In the study, the parameters of the ANN model were decided as a result of different trials in the training phase. It was observed that the feedback ANN model with two hidden layers and three neurons in each layer gave the most successful result. The structure of the developed model is shown in Figure 3.
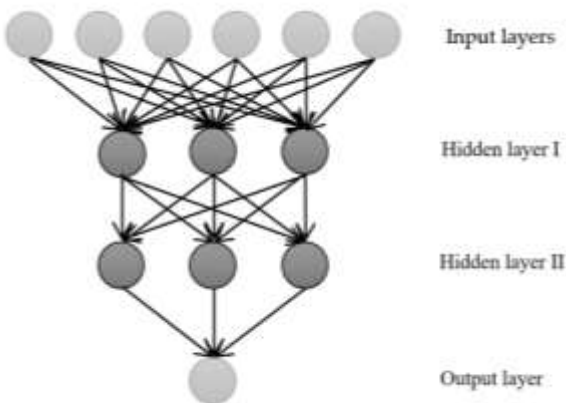

Figure 3. *ANN model*

The most important step in the creation of decision trees is the criteria, or criteria according to which the branching in the tree will be made or according to

which attribute values the tree structure will be formed. There have been various approaches developed to solve the problem. The most important ones are information gain and information gain ratio, Gini index, towing rule, and Chi-Square contingency table statistics. According to the information gain and information gain ratio methods, information theory, including entropy rules, is used to determine which feature to branch according to in the decision tree. Entropy is a measure of disorder or uncertainty in a system [29]. Entropy:

$$\text{Entropy}(T) = \sum_{i=1}^{n} p_i log_2(p_i) \quad (5)$$

Information gain is based on partitioning a data set based on a feature and then subtracting it from the overall entropy. As it approaches 1, the importance of the feature decreases. In terms of information gain, however, the opposite is the case [30].

$$\text{Gain}(T, X) = \text{Entropy}(T) - \text{Entropy}(T, X) \quad (6)$$

In the study, R^2 (Coefficient of determination), MSE (Mean square error), and MAE (mean absolute error) metrics were used for model evaluation, success, and error analyses. Table 1 shows the evaluation of the models.

Table 1. *Evaluation of models*

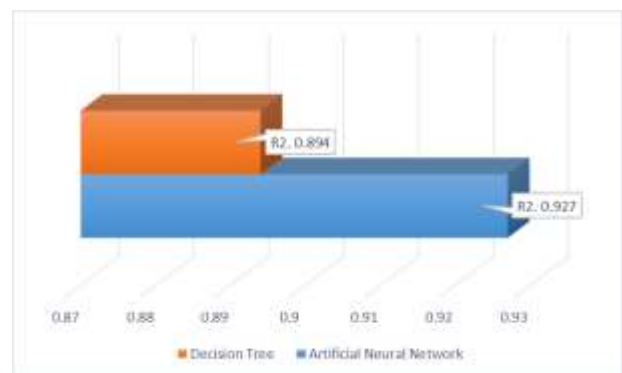|	Artificial Neural Network	Decision Tree
R^2	0.927	0.894
MSE	0.009	0.013
MAE	0.006	0.011

Figure 4. *R^2 of correlation*

Figure 4 shows the (R^2) graphs. As a result of the study, the R^2 was 89.4% for the decision tree and 92.7% for the ANN. According to these results, it can be concluded that the independent variables in the models have high explanatory power for the dependent variable and provide ideal values. Figure 5 shows the MSE graphs. In the study, the MSE metric was 0.013 for the decision tree and 0.009 for the ANN. These values are very close to
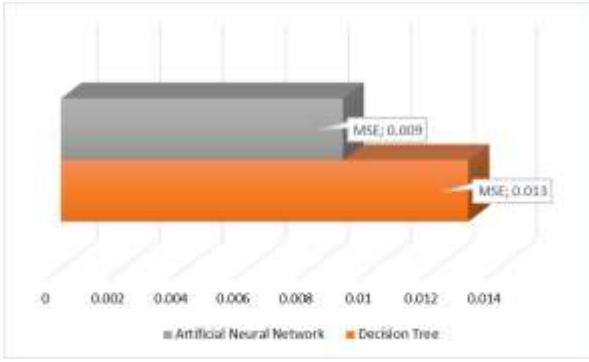
***Figure 5.** MSE of the model*

zero, which are ideal values. These numerical values show that the model has a low error rate and successfully predicts the dependent variable.
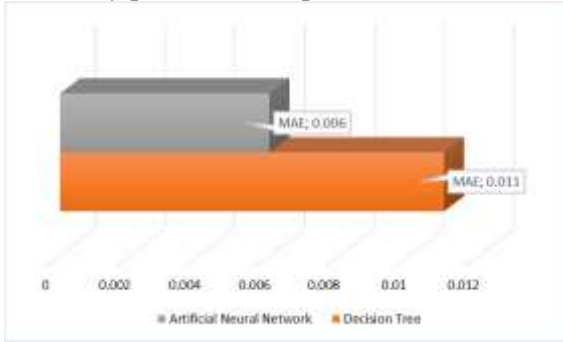

***Figure 6.** MAE evaluation*

Figure 6 shows the MAE graphs. In the study, the MAE metric was 0.006 for the decision tree and 0.011 for the artificial neural networks. These values are very close to the ideal MAE and show that the model is successful.

4. Conclusions

Carbon dioxide emissions are an important factor causing an increase in greenhouse gases in the atmosphere and climate change. Controlling and reducing carbon dioxide emissions plays an important role in combating global warming and climate change. Various national and international efforts are directed towards reducing greenhouse gas emissions and switching to sustainable energy sources.An estimation of carbon dioxide emissions in the coming years is important to determine the measures to be taken. Carbon dioxide emissions are often associated with specific industrial activities, energy production, transport, and other human interactions. Estimating these emissions is important to provide information on future climate change and environmental impacts. However, making an accurate and reliable forecast is a complex process and involves a number of factors.

In this study, Turkey's carbon dioxide emissions were successfully predicted using two different machine learning models. The success of the study was evaluated using three different statistical metrics: R2, MSE, and MAE. The R2 of decision trees was 89.4%, MSE 0.013, and MAE 0.011; the R2 of artificial neural networks was 92.7%, MSE 0.009, and MAE 0.006. According to these results, it can be said that the success of both models is high. When we compare the two models, it is seen that ANN is more successful and predicts with less error than decision trees. This result is similar to the literature [31,32]. According to these results, Turkey's carbon dioxide emissions can be successfully estimated according to machine learning methods. Thus, the measures to be taken for the reduction of carbon dioxide emissions and their effects can be determined. It will be an important source of information, especially for policymakers. In further studies, different parameters that will affect carbon dioxide emissions can be added to the data set, and different machine learning models can be tested to increase the success rate. In addition, the study can be adapted to different countries. Thus, it will be useful for the measures to be taken by other countries.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** No conflict of interest was declared by the authors.
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** I declare that there is no contribution from any other author
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Kunt, F. (2007). Hava Kirliliğinin Yapay Sinir Ağları Yöntemiyle Modellenmesi ve Tahmini, Selçuk University Graduate School of Natural and Applied Sciences, M.Sc. Thesis, *Environmental Engineering* Department, Konya.

[2]Aydınlar, B., Güveni H. ve Kırksekiz, S. (2009). Hava Kirliliği ve Modellenmesi, *Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Çevre Mühendisliği Bölümü Yüksek Lisans Rapor.*

[3]Alimissis, A., Philippopoulos, K., Tzanis, C.G., and Deligiorgi, D. (2018). Spatial estimation of urban air pollution with the use of artificial neural network

models, *Atmospheric Environment*, 191, 205-213, 2018. DOI: 10.1016/j.atmosenv.2018.07.058

[4]Hu, K. & Rahman, A. (2017). HazeEst: Machine Learning Based Metropolitan Air Pollution Estimation From Fixed and Mobile Sensors, *IEEE Sensors,* 17(11): 3571-3525. DOI: 10.1109/JSEN.2017.2690975

[5]Huang, C-J., & Kuo, P-H. (2018). A Deep CNN-LSTM Model for Particulate Matter (PM2.5) *Forecasting in Smart Cities, Sensors*

[6]Martínez-Espaňa, R., Bueno-Crespo, A., Timón, I., Soto, J., Muňoz, A. & Cecilia, J.M. (2018). Air-Pollution Prediction in Smart Cities through Machine Learning Methods: *A Case of Study in Murcia*, Spain.

[7]Tamas, W., Notton, G., Paoli, C., Nivet, M-L. & Voyant, C. (2016). Hybridization of Air Quality Frecasting Models Using Machine Learning and Clustering: An Orginal Approach to Detect Pollutant Peaks, *Aerosol and Air Qaulity Research,* 16: 405-416.

[8]Zaree, T. & Honarvar, A.R. (2018). Improvement of Air Pollution Prediction in a Smart City and its Correction with Weather Conditions using Metrological Big Data, Turkish *Journal of Electrical Engineering & Computer Sciences*, 26: 1302-1313.

[9]Kuhn, M., Johnson, K., Kuhn, M., & Johnson, K. (2013). Data pre-processing. *Applied predictive modeling*, 27-59.

[10]Onoz, B., & Oguz, B. (2003). Assessment of outliers in statistical data analysis. *Integrated technologies for environmental monitoring and information production*, 173-180.

[11]Kwak, S. K., & Kim, J. H. (2017). Statistical data preparation: management of missing values and outliers. *Korean journal of anaesthesiology,* 70(4), 407-411.

[12]Nayak, S. C., Misra, B. B., & Behera, H. S. (2014). Impact of data normalization on stock index forecasting. *International Journal of Computer Information Systems and Industrial Management Applications,* 6(2014), 257-269.

[13]Yu, L., Wang, S., & Lai, K. K. (2005). An integrated data preparation scheme for neural network data analysis. *IEEE Transactions on Knowledge and Data Engineering*, 18(2), 217-230.

[14]Mazziotta, M., & Pareto, A. (2022). Normalization methods for spatio- temporal analysis of environmental performance: Revisiting the Min–Max method. *Environmetrics*, 33(5), e2730.

[15]Kabas, O., Kayakus, M., Ünal, İ., & Moiceanu, G. (2023). Deformation Energy Estimation of Cherry Tomato Based on Some Engineering Parameters Using Machine-Learning Algorithms. *Applied Sciences*, 13(15), 8906.

[16]Kayakuş, M., Terzioğlu, M., Erdoğan, D., Zetter, S. A., Kabas, O., & Moiceanu, G. (2023). European Union 2030 carbon emission target: The case of Turkey. *Sustainability,* 15(17), 13025. DOI: 10.3390/su151713025

[17]Geurts, P., Irrthum, A., & Wehenkel, L. (2009). Supervised learning with decision tree-based methods in computational and systems biology. *Molecular Biosystems,* 5(12), 1593-1605. DOI: 10.1039/b907946g

[18]Thomas, T., P. Vijayaraghavan, A., Emmanuel, S. (2020). Applications of decision trees. *Machine learning approaches in cyber security analytics*, 157-184.

[19]Kayakuş, M., & Açikgöz, F. Y. (2022). Classification of News Texts by Categories Using Machine Learning Methods. *Alphanumeric Journal,* 10(2), 155-166.

[20]Kriegeskorte, N. (2015). Deep neural networks: a new framework for modeling biological vision and brain information processing. *Annual review of vision science,* 1, 417-446.

[21]Ünal U., G. F., Terzioğlu, M., Kayakuş, M., Tutcu, B., Çoşgun, A., Tonguç, G., & Kaplan Yildirim, R. (2023). Estimation of Methane Gas Production in Turkey Using Machine Learning Methods. *Applied Sciences,* 13(14), 8442.

[22]Kayakus, M., Tutcu, B., Terzioglu, M., Talaş, H., & Ünal Uyar, G. F. (2023). ROA and ROE Forecasting in Iron and Steel Industry Using Machine Learning Techniques for Sustainable Profitability. *Sustainability,* 15(9), 7389. DOI:10.3390/su15097389

[23]Huang, Y., Kangas, L. J., & Rasco, B. A. (2007). Applications of artificial neural networks (ANNs) in food science. *Critical reviews in food science and nutrition*, 47(2), 113-126.

[24]Yağmur, A., Kayakuş, M., & Terzioğlu, M. (2022). House price prediction modeling using machine learning techniques: a comparative study. *Aestimum*, 81.

[25]Piepho, H. P. (2019). A coefficient of determination (R2) for generalized linear mixed models. *Biometrical Journal,* 61(4), 860-872.

[26]Wang, Z., & Bovik, A. C. (2009). Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE signal processing magazine,* 26(1), 98-117.

[27]Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific model development discussions,* 7(1), 1525-1534.

[28]Xu, X., He, X., Ai, Q., & Qiu, R. C. (2015). A correlation analysis method for power systems based on random matrix theory. *IEEE Transactions on smart grid*, 8(4), 1811-1820.

[29]Hu, Q., Che, X., Zhang, L., Zhang, D., Guo, M., & Yu, D. (2011). Rank entropy-based decision trees for monotonic classification. *IEEE Transactions on Knowledge and Data Engineering,* 24(11), 2052-2064.

[30]Kingsford, C., & Salzberg, S. L. (2008). What are decision trees?. *Nature biotechnology*, 26(9), 1011-1013.

[31]Ağbulut, Ü. (2022). Forecasting of transportation-related energy demand and CO2 emissions in Turkey with different machine learning algorithms. *Sustainable Production and Cons*. 29, 141-157.

[32]Kayakuş, M., & Açikgöz, F Y. (2023). Fake News Detection on Twitter with Machine Learning Methods. *Journal of Abant Social Sciences,* 23(2), 1017-1027.