



RESEARCH ARTICLE

# Modeling Automobile Sales in Turkiye with Regression-Based Machine Learning Algorithms

Merve BABAOĞLU<sup>1</sup> , Ahmet COŞKUNÇAY<sup>2</sup> , Tolga AYDIN<sup>2</sup> 

## ABSTRACT

The automobile sector is the locomotive of industrialized countries. The employment opportunities it creates are of great value because of its interconnectedness with other industries and the value it adds. Demand forecasting studies in such an important sector are one of the main drivers for the provision of raw materials and services needed in the future. In this study, 10 independent variables are used that directly or indirectly affect the level of car sales, which is our dependent variable. These variables are gross domestic product, real sector confidence index, capital expenditures, household consumption expenditures, inflation rate, consumer confidence index, percentage of one-year term deposits, and oil barrel, gold, and dollar prices. The dataset used consists of annual data between 2000 and 2021. To examine the sales forecast model, two variables that affect minimum sales are first extracted from the model using the least squares method. Linear Regression, Decision Tree, Random Forest, Ridge, AdaBoost, Elastic-net, and Lasso Regression algorithms are applied to build a predictive model with these variables. The Mean Squared Error (MSE), Mean Absolute Error (MAE), and coefficient of determination ( $R^2$ ) are used to compare the performance of the predictive models. This study proposes an approach for sectors affected directly or indirectly by automotive sales to gain foresight on this issue.

**Keywords:** Automobile Sales, Regression, Demand Forecasting



DOI: 10.26650/JODA.1242645

<sup>1</sup>Artuklu University, Vocational Higher School of Mardin, Department of Computer Technologies, Mardin, Turkiye

<sup>2</sup>Ataturk University, Faculty of Engineering, Department of Computer Engineering, Erzurum, Turkiye

ORCID: M.B. 0000-0003-3030-8690;  
A.C. 0000-0002-7411-310X;  
T.A. 0000-0002-8971-3255

## Corresponding author:

Merve BABAOĞLU,  
Artuklu University, Vocational Higher School of Mardin, Department of Computer Technologies, Mardin, Turkiye  
E-mail: mervebabaoglu@artuklu.edu.tr

**Submitted:** 26.01.2023

**Revision Requested:** 15.02.2023

**Last Revision Received:** 13.03.2023

**Accepted:** 15.03.2023

**Citation:** Babaoglu, M., Coskuncay, A. & Aydin, T. (2023). Modeling automobile sales in Turkiye with regression-based machine learning algorithms. *Journal of Data Applications*, 1, 19-33.  
<https://doi.org/10.26650/JODA.1242645>



## Introduction

The automotive sector is one of the sectors that has contributed the most to the transformation of the economy around the world since the twentieth century. This sector is also one of the main consumers of many branches of the manufacturing industry, from iron and steel to petrochemicals, and is an important trigger for their development. Especially for Turkiye, it is an indispensable sector as it creates high added value, brings high income, has a positive impact on employment and has a structure that allows bringing developed technologies to our country (Görener & Görener, 2008).

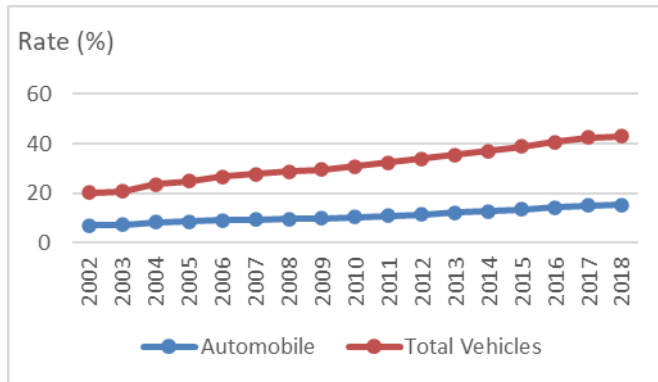
Although the automotive sector in Turkiye has a 50-year history, the sector, which started only with assembly in the mid-1950s, has reached its current position with increased investments after the Customs Union Agreement with the EU in 1996 (Dikmen, 2016).

Automobile production accounts for about 70% of the total mass production of motor vehicles in the world. This ratio is also true for Turkiye. Automobile production takes place on a much larger scale than the production of other motor vehicles. For this reason, automobile production also supports the production of other vehicles by creating a strong ancillary industry (Mmpvizyon, 2023).

Turkiye is a country with a majority young population.. The number of motor vehicles per capita of this population is increasing. The players in the automotive sector are also aware of this great strength. It is imperative for every player in this field to estimate how many sales there will be in the industry. Sales forecasting plays an important role in many areas, from purchasing raw materials to providing the necessary personnel for the industry, and from R&D investments to determining advertising expenditures.

The per capita rates of cars and motor vehicles in Turkiye are shown in Figure 1 (Euronews, 2022).

Forecasts are referred to when conclusions can be drawn about the future based on data from the past. These conclusions can be drawn using mathematical methods as well as subjective interpretations. Demand forecasting is the ability to predict how many sales will be made of a product using various methods. Demand forecasting is a very challenging problem because it is influenced by many factors. GDP, population, inflation, imports, exports, consumer confidence index, fuel prices, and gold and foreign exchange prices are just a few of these factors that affect automobile sales.



**Figure 1.** Rates of cars and motor vehicles per capita in Türkiye.

Machine learning methods, which are becoming increasingly popular in multivariate forecasting studies, continue to produce successful results. Choosing the right algorithms for a forecasting study is as important as choosing the right variables. The data must be properly analyzed, the algorithms must be understood, and the appropriate model must be selected for the area of focus.

The aim of the study is to make inferences about the future by looking at the history of automobile sales, which directly or indirectly affect many sectors. Thus, these sectors will be able to get ideas from these inferences to determine the steps they will take.

In this study, 10 different variables that directly or indirectly affect sales are used to model automotive demand. It is also noteworthy that so many variables are used in this study to predict automobile sales because the more variables used for the prediction, the greater the amount of explained variance in the dependent variable.

In order to process numerical data using machine learning methods and create the most accurate predictive models possible, regression algorithms are preferred in this study. Linear, Ridge, Lasso, ElasticNet, Random Forest, Decision Trees, and AdaBoost regression algorithms are preferred for this study. Models are built for sales forecasting and comparisons are made with some commonly used performance criteria.

In the second part of the study, studies in the field of automotive sales forecasting are presented from the literature. In the third part, regression algorithms, which are the most appropriate machine learning methods for the multivariate structure preferred for predictions, are described. The fourth part is about the methods used for prediction.

In the fifth part of the study, the findings obtained by the variables selected for research and machine learning methods are shown. In the sixth section, the findings obtained and the comparison of their results are carried out.

## Literature Review

Numerous studies have been conducted in Turkiye and around the world on the future sales of motor vehicles. While mathematical or statistical methods have often been used for this purpose in the past, nowadays machine learning methods, artificial neural networks, or meta-heuristic algorithms are preferred, which provide faster and more reliable results.

Karaatlı et al. (2012), created a prediction of car sales using artificial neural network methods. Monthly data between January-2007 and June-2011 were used in the study. Gross domestic product, real sector confidence index, investment spending, consumer spending, consumer confidence index, dollar exchange rate, and time were used as independent data. Since the MAPE value in the study was 16.82%, the estimate made falls into the class of “correct estimates”.

Sharma and Sinha (2012), used fuzzy neural back sales of cars of one brand in India. They tried to predict the results using a propagation algorithm and compared the results with multiple regression algorithm results.

Hulsmann et al. (2012), conducted a study to evaluate German and US car market prediction models. In their study, they argue that decision trees should be used as the most accurate and explainable method.

Kuvvetli et al. (2015), aimed to estimate monthly vehicle sales for different segments and brands considering economic and environmental parameters. The Levenberg-Marquadt algorithm was chosen to train the model. The results were compared with the linear regression results.

Topal (2019), attempted to estimate the sales volume of a specific car brand using online consumer integration and search engine data. The data was analyzed using artificial neural networks and the Bayesian backpropagation method. The correlation found was 74%, which is above the acceptable value.

Kaya and Yıldırım (2020), proposed an 8-layer Deep Neural Network (DSA) model for predicting automobile sales. The inputs of the model were composed of various economic indicators such as exchange rate, gross domestic product, consumer confidence index, and consumer price index. Forecasts for vehicle sales were made based on the model’s outputs. Between 2011 and 2018, a total of 90 data were collected and analyzed on a monthly basis.

Civelek (2021), used artificial neural networks to predict the sales of tractors.

## Forecasting Models

Machine learning, a subfield of artificial intelligence, is a discipline that aims to make the best decision by learning from available information or experience.

The more important it is to find the data that most affects the dependent variable in prediction studies, the more important it is to find the most appropriate algorithm for that data. The optimal algorithm shows the most successful performance.

Multiple regression is a statistical tool used to predict the outcome that depends on several other independent predictors or variables. It combines several factors to find out how and to what extent they affect a particular outcome (Lin & Wu, 1999).

Using regression to establish a relationship between the dependent variable and many independent variables is a suitable method for prediction. However, the selection of independent variables is a crucial first step to obtaining accurate predictions (Shahabuddin, 2009). Ridge, lasso, and elastic net regression methods can also be good options in such cases (Bulut, 2018).

In our study, the machine learning methods that are best suited for our preferred multivariate structure are regression algorithms. Regression algorithms are controlled algorithms used to find out how much independent variables affect the dependent variable and to find out the possible relationships between these variables.

### Multiple Linear Regression

Multiple linear regression describes the linear relationship between two or more independent variables and a dependent variable. In this regression, there is a relationship between dependent and independent variables. Eq. 1 is used for linear regression. The independent variables are denoted by  $x$  and the dependent variables are denoted by  $Y$ .

$$Y = \beta_0 + \beta_1 x_1 + \epsilon \quad \text{Eq. 1}$$

$Y$ : Dependent variable observation vector.

$x$ : The independent variable observation matrix.

$\beta$ : The vector of coefficients.

$\epsilon$ : The random error vector.

Representation of Eq. 1 by matrix:

$$\begin{bmatrix} y_1 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & \dots & x_1 \\ \dots & \dots & \dots \\ 1 & \dots & x_n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \dots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{bmatrix} \tag{Eq. 2}$$

In this case  $\beta$ :

$$\beta = (X'X)^{-1}X'Y \tag{Eq. 3}$$

Classical regression analysis is performed with the above Eq. 3.

### Decision Tree

Decision Trees have a form that can be configured for both classification and regression.

If our data are numerical, the regression structure is used, if categorical, the classification structure is used. By decoding the data we have, calculations based on the variables are made and a rule tree is created by establishing a relationship between these variables.

This structure is used to divide a large amount of data into very small groups.

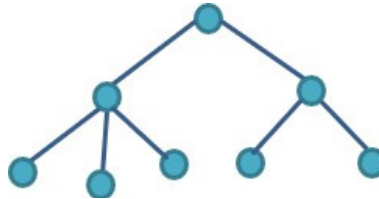


Figure 2. Simple Decision Tree.

### Random Forest

Random Forest is a supervised learning algorithm. As the name implies, it creates multiple decision trees, translates them into a forest, and combines them to obtain a more stable prediction.

Random Forest adds additional randomness to the model as the trees grow. Instead of looking for the most important feature when decoupling a node, it looks for the best feature from a random subset of features. This creates a large variety that often leads to a better model (Devhunter, 2022).

Random Forest has two main processing parameters. These parameters are the number of randomly selected estimators at each node ( $m$ ), and the number of trees in the forest ( $J$ ). In classification, the default value of  $m$  is  $m=\sqrt{p}$ , where  $p$  denotes the total number of predictors (Cutler et al., 2012).

## Ridge

Ridge regression has a similar structure to the least squares method. It creates a model with all variables in the data set but provides a solution by bringing the coefficients of the unrelated variables closer to zero.

The Ridge estimator formula is as shown in Eq. 4:

$$\hat{\beta}(\text{ridge}) = \arg \min ||y - x\beta||^2 + \gamma ||\beta||^2 \quad \text{Eq. 4}$$

As can be seen in Eq. 4, ridge regression applies a quadratic correction in addition to classical regression. Here  $\gamma \geq 0$  is expressed as a penalty correction or complexity coefficient. At the same time, the magnitude of this value means that the correction will also be large.

## Lasso

Lasso regression attempts to minimize the error by using the least squares method, as Ridge does. Unlike Ridge, however, it equates the coefficients of unrelated variables to zero. Thus, it has the great property of excluding irrelevant variables.

$$\hat{\beta}(\text{lasso}) = \arg \min ||y - x\beta||^2 + \gamma ||\beta|| \quad \text{Eq. 5}$$

In this algorithm, the correction is made according to the absolute value. Therefore, it is important that the margin of error determined by the least squares method is kept as small as possible.

## Elastic-Net

Elastic-Net is a middle ground between Ridge and Lasso regression. Elastic-Net performs a punishment operation like Ridge regression and makes a variable selection like Lasso regression. Penalization is in the style of Ridge regression and variable selection is in the style of Lasso regression.

$$\beta(\text{elasticnet}) = (\text{argmin} ||y - x\beta||^2 + \gamma_2 ||\beta||^2 + \gamma_1 ||\beta||) \quad \text{Eq. 6}$$

When  $\gamma=0$ , Elastic-Net corresponds to Ridge regression and  $\gamma=1$  corresponds to Lasso regression.

## AdaBoost

In the AdaBoost algorithm, the training process continues by increasing the relative weight of the training data belonging to the incorrectly estimated data in the first regression while the next regression operation is performed. The regression process continues until the weights are updated and the stop condition is created (Freund & Schapire, 1997).

## Method

### Dataset

The study attempts to estimate the level of automobile sales in Turkiye using machine learning methods. The automobile sales data used in this study include all domestic and imported automobile sales. Some commonly used independent variables that directly or indirectly affect automobile sales are: Real Sector Confidence Index, Capital Expenditure, Consumer Confidence Index, and Oil Barrel (Karaatlı et al., 2012), GDP, Dollar Rate (Lin & Wu, 1999), Household Consumption Expenditure, Inflation Rate, Percentage of One-Year Time Deposits (Alper & Mumcu, 2000), and Gold Price. The sources from which these data are obtained are listed in Table 1.

**Table 1.** *Parameters measured in the experiments.*

Variable	Source
Car sales units	ODD
GDP	World bank
Oil barrel prices	Investing.com
Real sector confidence index, dollar price, percentage of 1 year term deposits	TCMB
CPI	Legal bank
Investment expenditures	Department of strategy and budget
Household consumption expenditures, Consumer confidence index,	TUIK
Gold Prices	Altinpiyasa.com

In total, we have 10 independent variables. To determine which of these variables is the one that most affects the level of auto sales, the least squares method is used.

The results of this method are shown in Figure 3. According to these results, the variables GDP, Real sector confidence, CPI and Maturity rate with a large p-value (significance) are excluded from the model. This is because the smaller the p-value, the more power these independent variables have to explain the dependent variable.

In Figure 3, you can see the p-value values of the variables.



	coef	std err	t	P> t
GDP	24.8794	267.360	0.093	0.928
Oil Barrel	-2844.3759	1203.190	-2.364	0.038
Real Sector Confidence	-3878.3620	5272.276	-0.736	0.477
CPI	1353.4736	4053.247	0.334	0.745
Investment Expenditures	0.0067	0.002	3.139	0.009
Household Consumption Expenditure	-56.6254	60.220	-0.940	0.367
Consumer Confidence Index	1.71e+04	5589.799	3.059	0.011
USD	-2.852e+05	8.14e+04	-3.504	0.005
Maturity Rate	-5381.1566	5904.687	-0.911	0.382
Gold	2743.8419	740.760	3.704	0.003
const	-2.528e+05	4.16e+05	-0.608	0.556

**Figure 3.** *p-values of independent variables.*

After the variables to be used in the model are determined, the data are preprocessed and the average of the series is taken and it is assigned to the lost data.

In the next section, the preferred methods for comparing the performance of the algorithms in the study will be described.

### Model Performance Measures

Performance evaluation criteria  $R^2$ , MSE, and MAE are preferred to compare predictive modeling studies.

The  $R^2$  value determines how well the data will fit into the regression model. MSE gives you an absolute number of how much your predicted results differ from the actual number. Finally, MAE sums the absolute error value, a more direct representation of the sum of the error terms.

The formulas of these criteria are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (o_i - p_i)^2}{\sum_{i=1}^n (o_i - \sigma)^2} \quad \text{Eq. 7}$$

$$MSE = \frac{1}{n} \sum_{j=1}^n (o_i - p_i)^2 \quad \text{Eq. 8}$$

$$MAE = \frac{1}{n} \sum_{j=1}^n |o_i - p_i| \quad \text{Eq. 9}$$

In the above formulas,  $n$  is the number of samples,  $o_i$  is the true value of the observation,  $p_i$  is the estimated value of the observation, and  $\sigma$  is the average of the true observation values.

### Findings

For Random Forest and Decision Tree Algorithms, 75% training and 25% test data are selected. For all other algorithms, 80% training and 20% test data are selected. In this way, the

best  $R^2$  values are obtained. When data is divided in this way, Random Forest and Decision Tree regression algorithms are tested for 6 years and the others are tested for 5 years. In addition, since Ridge, Lasso and Elastic-Net regression algorithms give a better  $R^2$  value when the random\_state value is determined and constant years are tested, the graph of these three algorithms is created together and it is observed that the prediction results are very close to each other.

### Forecasting with Linear Regression Model

Figure 4 shows the actual sales values and the forecasted values that the algorithm has given us.

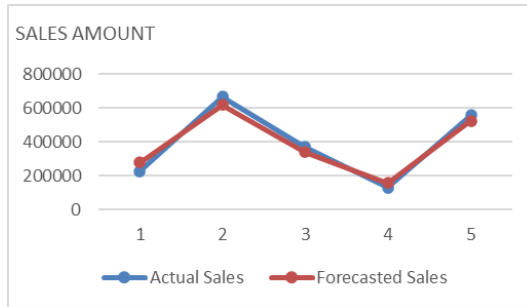


Figure 4. Forecasted and actual sales by Linear Regression.

In the Linear Regression algorithm, the performance criterion values are realized as follows:

$$R^2_{linear} = 96\%$$

$$MSE_{linear} = 0.001794$$

$$MAE_{linear} = 0.02651$$

### Forecasting with Decision Tree Regression Model

Figure 5 shows the actual sales values and the predicted values that the algorithm gave us.

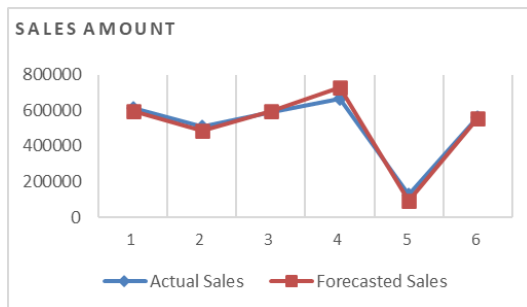


Figure 5. Forecasted and actual sales by Decision Tree regression.

The values of the performance criteria in the Decision Tree regression algorithm are realized as follows:

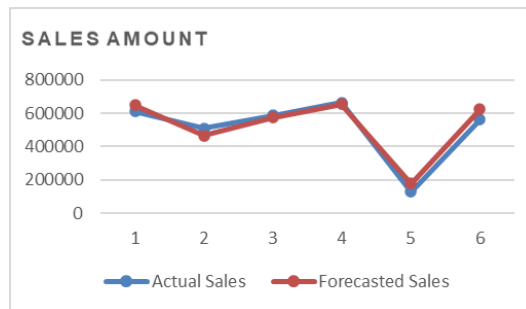
$$R^2_{tree} = 97\%$$

$$MSE_{tree} = 0.004795$$

$$MAE_{tree} = 0.052606$$

### Forecasting with Random Forest Regression Model

Figure 6 shows the actual sales values and the predicted values that the algorithm gave us.



**Figure 6.** Forecasted and actual sales by random forest regression.

The values of the performance criteria in the Random Forest Regression Algorithm are realized as follows:

$$R^2_{rf} = 94.67\%$$

$$MSE_{rf} = 0.004058$$

$$MAE_{rf} = 0.044487$$

### Forecasting with AdaBoost Regression Model

This algorithm is found to give better results when  $n\_estimators=10$  is chosen. Figure 7 shows the actual sales values and the predicted values that the algorithm gave us.

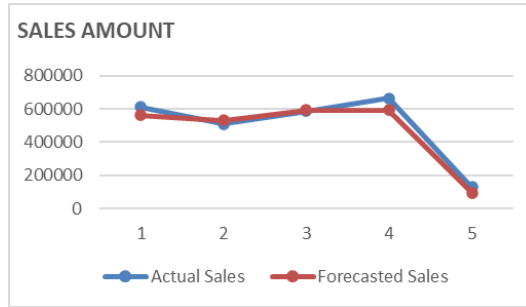


Figure 7. Forecasted and actual sales by AdaBoost regression.

The values of the performance criteria in the AdaBoost Regression Algorithm are realized as follows:

$$R^2_{adaboost} = 95.14\%$$

$$MSE_{adaboost} = 0.009622$$

$$MAE_{adaboost} = 0.06866$$

### Forecasting with Lasso, Ridge and Elastic-Net Regression Models

Figure 8 shows the actual values of car sales and the estimated values obtained from the Ridge, Lasso, and Elastic Net algorithm models.

In the Lasso algorithm, it has also been observed that it gives better results when the alpha value in the algorithm is set to 0.05. In the Lasso regression algorithm, the values of the performance criteria are realized as follows:

$$R^2_{lasso} = 90.04\%$$

$$MSE_{lasso} = 0.01452$$

$$MAE_{lasso} = 0.089073$$

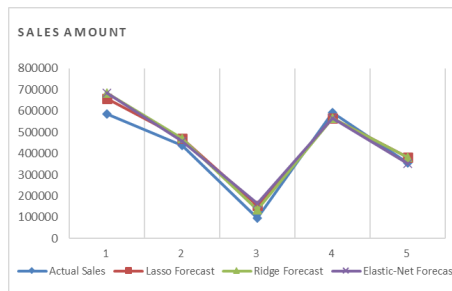


Figure 8. Forecasted and actual sales by Lasso, Ridge, and Elastic-Net regression models.

The Ridge algorithm was also found to give better results when the alpha value was chosen to be 0.01. The values of the performance criteria in the Ridge regression algorithm are realized as follows:

$$R^2_{ridge} = 94.78\%$$

$$MSE_{ridge} = 0.014396$$

$$MAE_{ridge} = 0.088564$$

The Elastic-Net algorithm was also found to perform better when the alpha value is 0.01,  $l1\_ratio=0.5$ , and  $normalized=False$ . The values of the performance criteria in the Elastic-Net regression algorithm are realized as follows:

$$R^2_{elastic} = 94.85\%$$

$$MSE_{elastic} = 0.014414$$

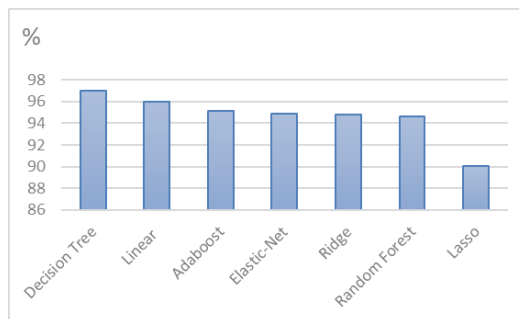
$$MAE_{elastic} = 0.088639$$

After the prediction studies of the models, the performances of the algorithms are compared using the values of  $R^2$ , MSE, and MAE. The summaries of these results can be found in Table 2.

**Table 2.** Performance measurement values of algorithms.

Algorithms	$R^2$	MSE	MAE
Decision Tree	97	0.004795	0.052606
Linear	96	0.001794	0.02651
AdaBoost	95.14	0.009622	0.06866
Elastic-Net	94.85	0.014414	0.088639
Ridge	94.78	0.014396	0.088564
Random Forest	94.67	0.004058	0.044487
Lasso	90.04	0.01452	0.102048

Figure 9 shows the  $R^2$  values of the results of the algorithms.



**Figure 9.**  $R^2$  values of algorithms.

## Conclusion

For this study, annual automobile sales data for 2000-2021 in Turkiye are taken from the ODD (Automotive Distributors Association) website. In order to perform prediction studies, modeling is performed using regression algorithms from machine learning methods and predictions are made. In performing this modeling, 10 different independent variables are used that most influence car sales. The variables are preprocessed and a value is assigned to the missing data.

Of these data, 4 variables that least affect the dependent variable using the least squares method are excluded from the model.

$R^2$  is a value that indicates how well the predicted values obtained explain the actual values. In other words, the obtained observed values are around or away from the regression line. The distance to this line is best determined by the value  $R^2$ . An  $R^2$  value of 90% or more indicates that the predictive model gives a good result. Figure 9 shows the  $R^2$  values of the regression algorithms whose performance is from best to worst.

Accordingly, the Decision Tree algorithm performed best with  $R^2$  values of about 97% in the car sales model estimation study. Moreover, the MSE value of this algorithm is 0.004795. The worst modeling algorithm is Lasso regression with 90.04%.

There may be many variables that affect car sales. The fact that the vast majority of these variables are used in a study complicates predictive studies. In this study, the use of 10 separate independent variables is preferred. However, the least squares method is used to facilitate modeling and remove unnecessary variables. As a result, four variables with a large p-value are excluded from the study. Estimation studies are conducted with the remaining 6 variables. Despite the large number of independent variables selected, the  $R^2$  values of 90% and higher obtained with all regression algorithms indicate that this study is successful and achieved its objective.

Looking at the automobile sales data, it is seen that the highest sales are made towards the end of the year. There may be different reasons for this. Thinking that prices will increase even more in the new year or entering the winter season may be the reasons that increase these sales. If automobile companies take advantage of this situation and make year-end campaigns, sales will increase substantially.

---

**Ethics Committee Approval:** Authors declared that this study does not require ethics committee approval.

**Peer Review:** Externally peer-reviewed.

**Author Contributions:** Conception/Design of Study- M.B., A.C., T.A.; Data Acquisition- M.B.; Data Analysis/ Interpretation- M.B.; Drafting Manuscript- M.B.; Critical Revision of Manuscript- A.C., T.A.; Final Approval and Accountability- M.B., A.C., T.A.

**Conflict of Interest:** The authors have no conflict of interest to declare.

**Grant Support:** The authors declared that this study has received no financial support.

## References

- Alper, C. E., Mumcu, A. (2000). *Türkiye’de otomobil talebinin tahmini*. Research Report, Boğaziçi University. [http://ideas.econ.boun.edu.tr/content/wp/ISS\\_EC\\_05\\_01.pdf](http://ideas.econ.boun.edu.tr/content/wp/ISS_EC_05_01.pdf).
- Bulut, H. (2018). *R uygulamaları ile çok değişkenli istatistiksel yöntemler*. Nobel Academy.
- Civelek, Ç. (2021). Tractor Sales Forecasting for Turkey Using Artificial Neural Network. *European Journal of Science and Technology*, 31 (1), 375-381.
- Cutler, A., Cutler, D. R. And Stevens, J. R. (2012), Random forests, *Ensemble Machine Learning Methods and Applications*, 45(1), 157-176.
- Devhunter. (2022, May 05), *Rastgele Orman (Random Forest) Algoritması*, <https://Devhunteryz.Wordpress.Com/2018/09/20/Rastgele-Ormanrandom-Forest- Algoritmasi/Comment-Page-1/>
- Dikmen, I. (2006). Otomotiv Sektörü Ve Rekabet Değerlendirme. Access Date: 10.12.2011, [http://www.kalder.org.Tr/Genel/15kongre/Sunumlar/Isik\\_Dikmen.Doc](http://www.kalder.org.Tr/Genel/15kongre/Sunumlar/Isik_Dikmen.Doc).
- Euronews. (2022, June 04). *Türkiye’de 100 kişiye düşen otomobil sayısı 14, AB’de ise 51*, <https://tr.euronews.com/2019/12/30/Turkiye-De-100-Kisiye-Dusen-Arac-Say-S-28-Ab-De-Ise-51>.
- Freund, Y., Schapire, R.E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119-139.
- Görener A., Görener Ö. (2008). The contributions of automotive industry in Turkish economy and sectoral-expectances about the future. *Journal Of Yaşar University*, 7(26), 306-319.
- Hulsman, M., Borscheid, D., Friedrich, C.M., Reith, D. (2012). General sales forecast models for automobile markets and their analysis. *Transactions On Machine Learning and Data Mining*, 5(2), 65-86.
- Karaatlı, M., Helvacıoğlu, Ö., Ömürbek, N., Tokgöz, G. (2012). An artificial neural network based automobile sales forecasting . *International Journal of Management Economics and Business*, 8(17), 87-100.
- Kaya, K.S., Yıldırım, Ö. (2020). A prediction model for automobile sales in Turkey using deep neural networks. *Journal Of Industrial Engineering*, 31(1), 57-74.
- Kuvvetli, Y., Dağsuyu, C., Oturakçı, M. (2015). A prediction approach based on artificial neural networks with consideration of environmental and economic indicators for car sales in Turkey. *Journal of Industrial Engineering*, 26(3), 23-31.
- Lin, Z.C., Wu, W.J. (1999). Multiple linear regression analysis of the overlay accuracy model zones. *IEEE Trans. On Semiconductor Manufacturing*, 12(2), 229 – 237.
- Mmpvizyon 2023 Otomotiv Sektör Raporu,” [https://tubitak.gov.tr/tubitak\\_content\\_files/vizyon2023/mm/ek3.pdf](https://tubitak.gov.tr/tubitak_content_files/vizyon2023/mm/ek3.pdf), 22.05.2022.
- Shahabuddin S. (2009). Forecasting automobile sales, *Management Research News*, 32(7), 670-682.
- Sharma, R., Sinha, A.K. (2012). Sales forecast of an automobile industry. *International Journal of Computer Applications*, 53(12), 25-28.
- Topal, İ. (2019). Çevrimiçi tüketici bütünleşmesi ve arama motoru verileri kullanılarak yapay sinir ağları ile otomobil satış tahmini, *Nevşehir Hacı Bektaş Veli Üniversitesi Journal of ISS*, 9(2), 534-551.

