BŞEÜ Fen Bilimleri Dergisi
7(1), 282-296, 2020

BSEU Journal of Science
DOI: 10.35193/bseufbd.611749

e-ISSN: 2458-7575 (http://dergipark.gov.tr/bseufbd)

Araştırma Makalesi – Research Article

Akaryakıt İstasyonu Yakıt Satış Otomasyonu Veriler Üzerinden Bir Veri Madenciliği Çalışması

İlhan Tarımer[1*], Buse Cennet Karadağ[2]

Geliş / Received: 27/08/2019 ***Revize / Revised: 08/05/2020*** ***Kabul / Accepted: 21/05/2020***

ÖZ

Bu makale benzin istasyonlarında satılan benzin, otogaz gibi petrol ürünlerinin satış tahminlerine ilişkin bir çalışmadır. Akaryakıt tiplerinin gelecekteki satışlarını tahmin etmek için bazı veri madenciliği teknikleri kullanıldı ve sonuçlar karşılaştırmalı olarak sunuldu. Elde edilen sonuçların akaryakıt istasyonlarında tank yönetimi stratejisini etkilediği görülmektedir. Bu amaçla belirlenen bir akaryakıt istasyonu deposuna giren ürün miktarı ve istasyondan satılan akaryakıt miktarı verilerinden, istasyonun periyodik satış tahminleri çıkartılmıştır. Önce C# ve .NET dillerinde web tabanlı bir yazılım geliştirilmiş ve bununla ilişkili bir veritabanı kurulmuştur. Geliştirilen yazılımın admin ve kullanıcı adlı 2 ayrı girişi üzerinden anlık satış verilerinin kaydı tutulmuştur. Ardından bu veritabanından çekilen depo ve satış verileri üzerinde veri madenciliği yapılmıştır. Bu çalışma sonuçlarının turizm güzergâhında bulunan akaryakıt istasyonlarına depo yönetimi ve satış tahminleri konusunda destek vermesi ve istasyon için bir satış denetim düzeni kurulmasına yardımcı olması beklenmektedir.

Anahtar Kelimeler- Yazılım, Akaryakıt satışı, Veri tahmini

[1]*Sorumlu yazar iletişim: itarimer@mu.edu.tr (https://orcid.org/0000-0002-7274-5680)
Bilişim Sistemleri Mühen, Muğla Sıtkı Koçman Üniversitesi, Teknoloji Fakültesi, Marmaris Yolu, Menteşe, 48000, Muğla
[2]İletişim: busecennet5@gmail.com (https://orcid.org/0000-0002-2488-1047)
Bilişim Sistemleri Mühendisliği ABD, Muğla Sıtkı Koçman Üniversitesi, Fen Bilimleri Enstitüsü, Menteşe, 48000, Muğla

BŞEÜ Fen Bilimleri Dergisi
7(1), 282-296, 2020

BSEU Journal of Science
DOI: 10.35193/bseufbd.611749

e-ISSN: 2458-7575 (http://dergipark.gov.tr/bseufbd)

A Data Mining Case Study Over Fuel Sales Automation Data

ABSTRACT

This paper is devoted to the sales estimations of petroleum products that are are sold at fuel stations such as auto gas and diesel. Some techniques of data mining were used to predict the future sales of fuel oils types and the results were presented comparatively. It is also seen that the results obtained effect the strategy of tank management at fuel stations. For this purpose, the periodic sales estimations of the station are deducted from the data of the quantity of products entering a fuel station depot and the amount of fuel sold from the station. First, a web-based software was developed in C # and .NET languages and a related database was established. Instant sales data was recorded through 2 separate entries called as admin and user of the developed software. Then, data mining was performed on the warehouse and sales data drawn from an associated database. The results of this study are expected to support fuel stations on the tourism route in terms of warehouse management and sales estimates and help to establish a sales control scheme for the station.

Keywords- Software, Fuel sales, Data estimation

BŞEÜ Fen Bilimleri Dergisi
7(1), 282-296, 2020

BSEU Journal of Science
DOI: 10.35193/bseufbd.611749

e-ISSN: 2458-7575 (http://dergipark.gov.tr/bseufbd)

I. INTRODUCTION

Turkey fuel sector together with more than 13000 fuel station enterprises has an important place within Turkish Economy. The vital issues of fuel station enterpreneurs, which are important actors of the fuel sector, is the identification and prevention of wastages [1]. In this study, it is aimed to make a warehouse audit of fuel products such as gasoline, diesel and LPG sold from a fuel station, to eliminate the wastes occurring, to make forecasts from the sales and to obtain information about the amount of fuel that the company should have in the future. ASP.NET is that is a web technology which gives multiple language supports, and together with the C# programming platform have been used to develop the software that is quick, secure and expandable in the study [2]. Thus, it is considered that the database operations are to do easily.

In the study, the software is designed that is a web based program, then the software's codes are created by using C# and ASP.NET platforms [3]. A database is created by using SQL Server 2014. This database is used to record the sales done by the fuel company. The software platform which is ASP.NET is a server-sided technology that it is used to develop Microsoft .NET Framework based applications [4]. Why the automation software is created by .NET platform is to make analyze by an interface created in this platdorm without set-up any other program [5]. Besides this, the .NET platform brings an advantage that an application is to be developed in different programming languages.

Operations can be made real in many computers at same time since an application prepared by means of .NET platform can be operated as web based one. Therefore, its usage is becoming increasingly common [5]. The reason of using C# language as parallel to .NET platform when creating the software is that it is a powerful, modern and object-oriented programming language [6].

As a definition, data mining [7, 8, 9] is a procedure that it is to access to information amongst big data [10].Models used in data mining are to be examined beneath two separate categories that are named definer and estimator. At estimator models, it is aimed that it is to develop a modal via the data that their results are known. It is also targeted that the result values are estimated for data sets which are unknown by using the created model [11]. On the other hand, at definer modals, the patterns found already at current data to be used at at guiding to guide is used [11]. Neural network and regression modals are estimator [12]. As it is seen from the chapter 2 that it is used a program named IBM SPSS Modeler to make data mining tasks. SPSS Modeler is a leading visual data science and machine learning solution. It helps organizations shorten the time to value and get the desired results by speeding up operational tasks for data scientists. SPSS Modeler allows organizations to take advantage of data assets and modern applications with complete algorithms and models that are ready for immediate use. Model evaluation and visualization operations can be performed after obtaining models with data mining algorithms with the related software. [13]. When analyzing the data within the SPSS Modeler, two algorithms of machine learning were used: artificial neural network and multiple linear regression. It is true that the earlier-mentioned papers make an infrastructure to the study. Besides these, based on the results obtained by data mining analysis of a fuel station, it is aimed to make estimations about the types of fuels that similar ones should have.

In the literature, there are many forecasting studies made in different fields with different data mining methods. Prediction studies with made multiple linear regression algorithms and artificial neural network algorithms are given below.

Marvuglia and Messineo (2012) ,with variables such as weather data, air conditioner or not, made the household electricity consumption estimation using the Radial Based Artificial Neural Network method. Carlson and Umble (1980) estimated car types using multiple linear regression analysis in this study, which used disposable income as independent variables, varying according to car type, gasoline prices, gasoline shortage, and the strike of American automotive workers. Businger and Read (1999) used regression analysis in the demand forecast of maintenance and repair parts in the American navy.

With ANN estimation methods, Alho (2014) demographic structure, Kouwenberg & Zwinkels (2014) housing market, Lima et al. (2014) examined the flow of water to the hydroelectric power plant. Turanlı & Güneren (2003) made demand forecasts for the tourism sector.

BŞEÜ Fen Bilimleri Dergisi
7(1), 282-296, 2020

BSEU Journal of Science
DOI: 10.35193/bseufbd.611749

e-ISSN: 2458-7575 (http://dergipark.gov.tr/bseufbd)

The flow plan of the paper is as following: The method followed has been emphasized in section 2. In this section, the data are withdrawn to the database created through the software automation; several subtractions have been made by making data mining over these data. The results are given at the end of this section. The discussions and conclusions are took part in the 3rd section.

II. METARIAL AND METHOD

For the study, a software is designed for a desktop use and its codes are written in C# platform. Figure 1 shows the interface of the software created. The following figures give the use of the interface how to run.

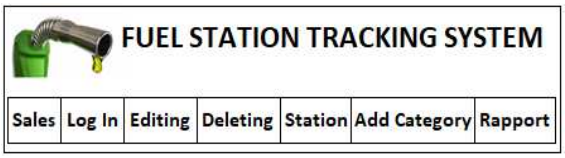

Figure 1. The fuel station tracking program

After login to the program as admin or user, the program interfaces direct the flow to the relevant pages. When user wants to make sales, to check product stock and to update product quantity, the program interfaces are redirected to the related pages.

A. User Login and New User

When the developed software is operated first, User Login is seen on the screen. If correct information is written, the flow is directed to the admin or user panels regarding to the type of user. In case of not affiliated to the program, new membership can be opened and it is able to enter to the system by clicking ”New Log” button.

The figure 2 shows the user login form. From this interface, it is connected to the database. Whether user name and its password are in the database system is checked. If it is admin, then it is directed to the admin panel otherwise to the user panel. In case of user name and/or its password are incorrect or aren’t being in the system, the interface responds you with an error message. If one is wanted to register to the system, firstly a form is opened where information can be filled. When information is entered and clicked on the ‘Save’ button, new user information is added to the database. Afterwards, it can be login to the system with the new user infos (Figure 1).

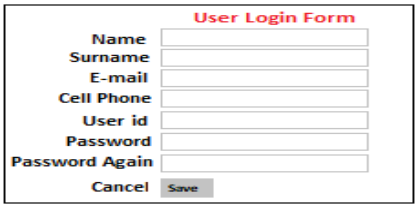

Figure 2. New user form interface

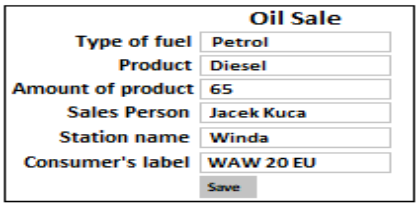

Figure 3. Product sales interface

B. Regulations on sales, recording and its infos for fuel

The Figure 2 and 3 show the fuel sales interface. From screen seen in the Figure 2, type of fuel is selected. The oil types are changed regarding to the oil selected. When the fuel oil type is selected, amount of product (liter or pcs) is entered; the name which is selling the oil and the active sales person’s name are determined. The consumer's car label is entered, thus sales process is completed (Figure 3).

BŞEÜ Fen Bilimleri Dergisi
7(1), 282-296, 2020

BSEU Journal of Science
DOI: 10.35193/bseufbd.611749

e-ISSN: 2458-7575 (http://dergipark.gov.tr/bseufbd)

In terms of the codes of the software, when the interface page is loaded, database connection is activated, the product type is taken from the database as well. Then the fuel products are listed according to the selected product type. The sales persons and names of the stations are listed through the activating database.

When to click on login button in the interface screen, name of the entered product name, sales person's name, station's name, amount of the product sold and consumer info are assigned to the defined variables. Thus, the database connection is opened and the sales price of the product is drawn according to the ID of the selected product. Besides these, sales price is assigned to the defined variables, and the entered data is added to the database. In the product registration interface name of the product, type of the product, wholesaler purchasing price and the amount of the product bought are entered and registered (Figure 4).

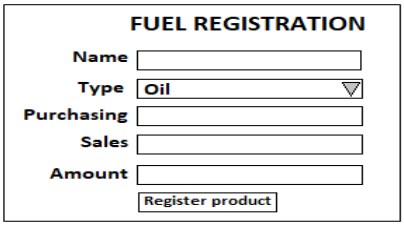

Figure 4. Product registration interface

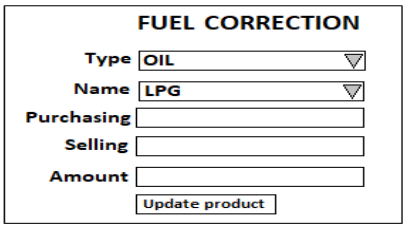

Figure 5. Product info updating interface

In the product registeration and updating interfaces, some incorrectness and/or lacks on purchasing and selling prices of the product entered are corrected or updated. The type and the name of product are chosen. The purchasing price that it wants to correct, the selling price and the amount are entered. The price of product and the amount of product are corrected as clicking on the "Update product" button (Figure 5).

C. Sales deletion, storing stations and add categories

This is the section that the one is included uncorrected, lacked and/or the data needed to be deleted. When the "select" button on the left side of the data line to be deleted is clicked on, the deletion is performed (Figure 6).

SALES DELETION

	Name of Product	Sales Person	Station Name	Sales Price	Amount of Product	Total Sales	Label of Vehicle	Date
Choose	LPG	Berat Gider	Merkez 1	3,78	5,29	19,9962	48ABC 372	10.06.2019 / 02:04:11
Choose	UnLeaded	Feride Bayrak	Akova 2	7,16	10,33	73,9628	48 GA 333	10.06.2019 / 02:04:43
Choose	LPG	İsmail Kirpi	Bezcey 2	3,78	13,23	50,0094	48 AAD 280	10.06.2019 / 02:04:59
Choose	Diesel	İrfan Sarsar	Merkez 2	6,61	40,85	270,0185	06 AGE 456	10.06.2019 / 02:05:59
Choose	LPG	İrfan Sarcan	Merkez 3	3,78	13,23	50,0094	48 KJ 347	10.06.2019 / 02:06:52
Choose	Diesel	Mert Korkar	Akova 1	6,61	7,56	49,9716	17 AAJ 024	10.06.2019 / 02:07:17
Choose	Diesel	Abdullah Velibey	Bezcey 1	6,61	32,53	215,0233	48 P 1345	10.06.2019 / 02:08:48
Choose	Diesel	Berat Gider	Central 1	6,61	6,06	40,0566	48 PE 1356	10.06.2019 / 02:09:50
Choose	LPG	Feride Bayrak	Akova 2	3,78	7,34	27,7452	48 AAN 907	10.06.2019 / 02:08:50
Choose	UnLeaded	İsmail Kirpi	Bezcey 2	7,16	3,49	24,9884	20 LG 673	10.06.2019 / 02:09:58
Choose	UnLeaded	İrfan Sarsar	Merkez 2	7,16	5,55	39,738	59 TKD 028	10.06.2019 / 02:10:23
Choose	UnLeaded	İrfan Sarcan	Merkez 3	7,16	7,75	55,49	18 HV 150	10.06.2019 / 02:12:31
Choose	LPG	Mert Korkar	Akova 1	3,78	10,24	38,7072	45 HVC 156	10.06.2019 / 05:01:18
Choose	LPG	Abdullah Velibey	Bezcey 1	3,78	9,87	37,3086	48 YKT 670	10.06.2019 / 06:17:37
Choose	Diesel	Berat Gider	Merkez 1	6,61	25,42	168,0262	48 UYH 560	10.06.2019 / 06:27:41

Figure 6. Sales deleting page

As shown in Figure 7, regarding to open a new branch or a new station, the station recording interface is used. To do this, it is more than enough to enter the station name and to click on "Station record" button.

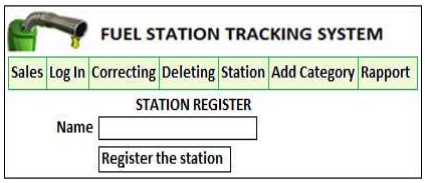

Figure 7. Station registering page

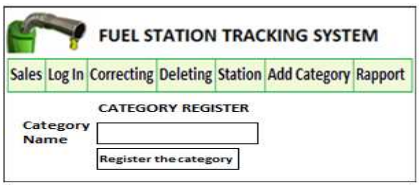

Figure 8. Category recording page

The page seen in the Figure 8 is an interface page that it would assist when it is wanted to add a new category or type of product. In this screen, if type of the product is entered, and clicked on the "Save category", then the recording procedure is realized (Figure 8).

D. Processing of the obtained data

In this chapter, data mining is performed on the fuel sales data of the relevant fuel station processed monthly and annually in Excel. Data mining is the acquisition of previously unknown, valid and applicable information from large databases and the use of this information in making business decisions [14]. The steps that make up the data mining process are given in the flowchart in Figure 9.

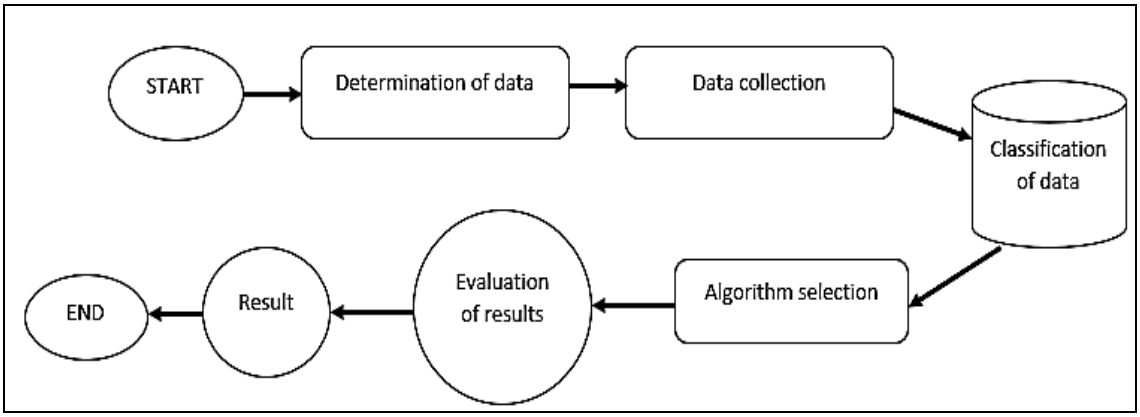

Figure 9. Data Mining Process Flowchart

Fuel purchase estimation analysis has been carried out in order to estimate the amount of fuel purchases that will take place in the coming months, taking into account the monthly fuel sales quantities made at the station. The flow chart of the operations to be performed for the estimation of the data is given in Figure 10.

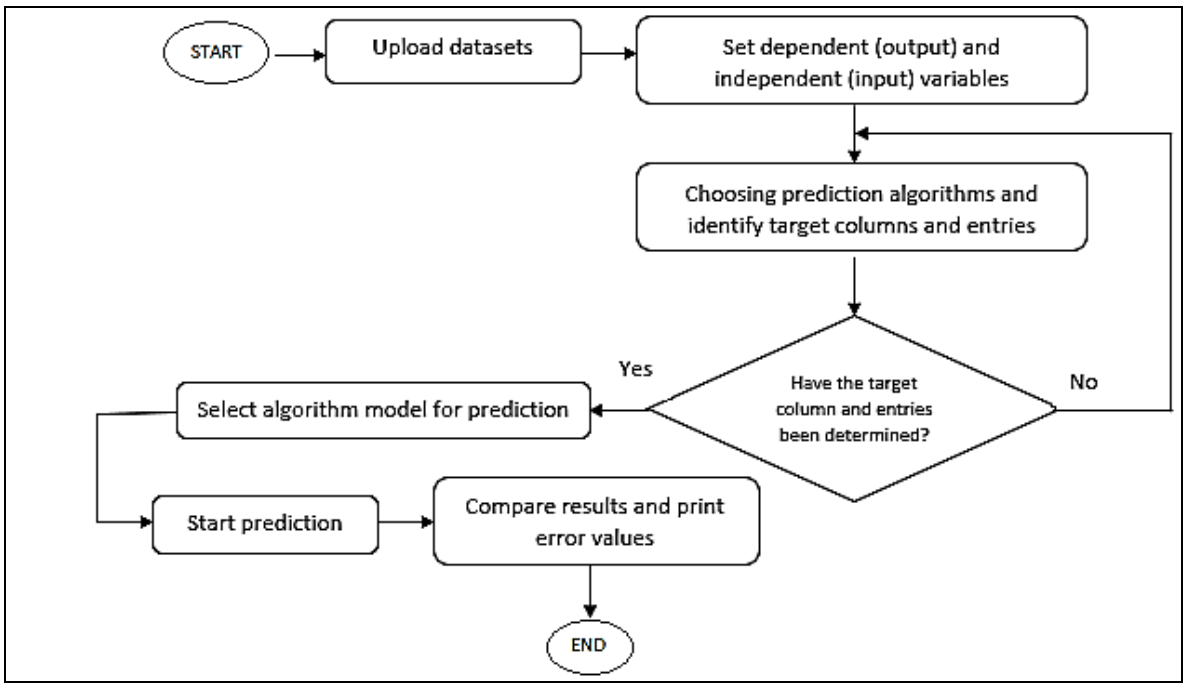

Figure 10. Data Mining Flowchart

1) Sales Forecasting with Artificial Neural Networks: In here in order to estimate the amount of fuel purchases in the coming months, 48 monthly data and artificial neural networks and algorithms of multiple linear regression analysis used. By examining the annual and monthly sales amounts, it was analyzed in which seasons and months the most sales occurred in fuel oil types. A table respecting a type of fuel used in the analysis is given below (Table 1).

Table 1. Sales data of diesel fuel type

Months	2015	2016	2017	2018	2019
January	21.917,19	45.657,91	65.892,63	103.569,87	88.212,00
February	19.968,72	50.305,15	71.443,32	76.829,65	105.957,00
March	20.376,41	51.106,92	94.104,43	106.523,59	103.705,00
April	23.726,87	59.496,61	84.669,22	128.424,21	90.698,40
May	24.291,98	63.687,92	99.587,14	123.373,86	105.315,00
June	23.322,94	56.844,83	96.641,39	113.353,00	89.905,00
July	22.661,80	55.429,23	107.426,69	121.493,00	93.807,39
August	22.662,96	62.286,94	112.578,00	115.138,00	89.671,00
September	22.676,70	56.020,36	107.042,64	108.690,00	92.349,00
October	29.663,51	67.513,49	129.314,35	115.000,00	85.277,00
November	34.593,90	65.732,01	126.860,65	107.470,00	70.872,00
December	29.384,37	71.890,41	112.033,42	94.728,00	77.969,00

Artificial neural networks are an information processing system inspired by biological neural networks. Artificial neural networks consist of several interconnections of artificial nerve cells and generally arranged in layers. Neural networks consist of two or three layers. These layers are input, hidden and output layers. Neural networks consisting of two layers have not hidden layers [14]. The structure of artificial neural networks is shown in Figure 11.

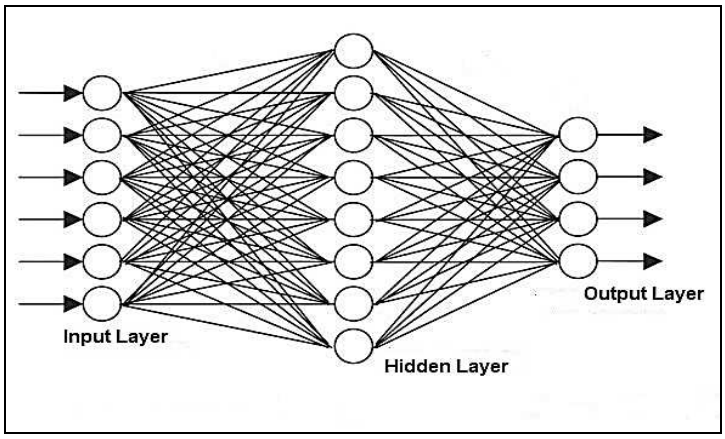

Figure 11. Structure of Artificial Neural Network

In this section, Multilayer Perceptron (MLP) and Radial Basic Function (Radial Based Function Network-RBFN) methods are used. Multilayer Perceptron method is a feed forward learning algorithm based on error. It performs the learning and decision making stages, which have two basic functions, by weighting, activation function and bias . Weight is the coefficient that each input is multiplied before going to the next stage. All inputs are collected by multiplying with their owns weights. The response that occurs as a result of sending this value to the activation function becomes the decision of the system.

Bias, on the other hand, is a user-added parameter that can vary from user to user, the way the mechanism works or its purpose [15]. Artificial neural network structures created in MLP method, which used when the estimate for fuel types of year 2019, are given in Figures 12, 13 and 14. In these shapes, 2015-2018 shows the years given as input in the network, the neurons hidden layers and 2019 shows the target column.

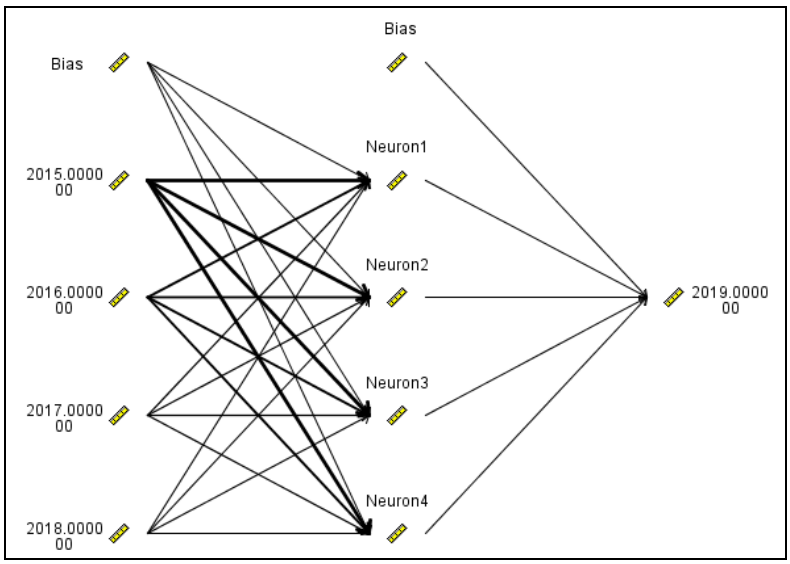

Figure 12. ANN Structure for 2019 Diesel Sales Forecast

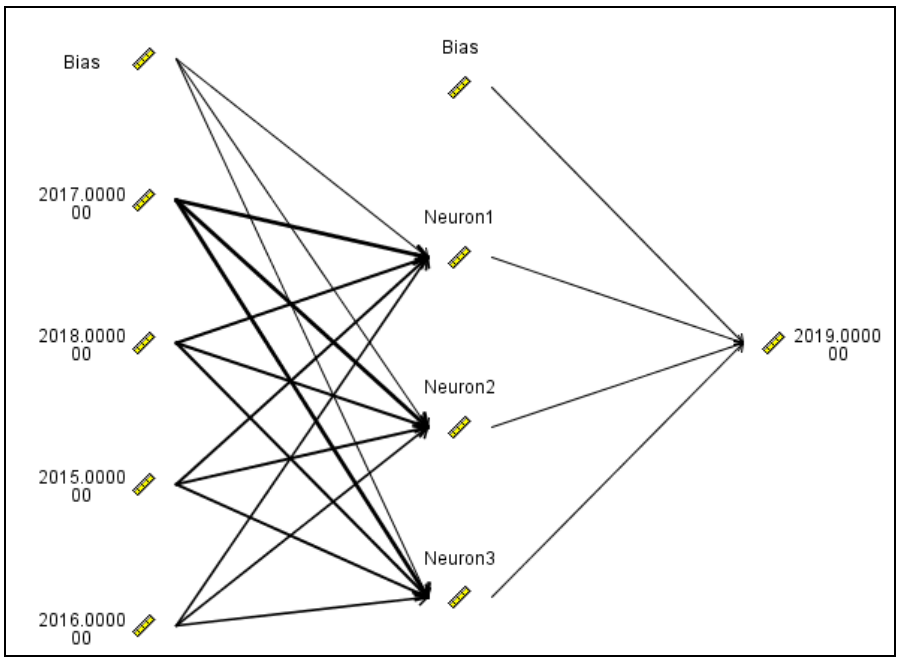

Figure 13. Ann Structure for 2019 Auto Gas Sales Forecast

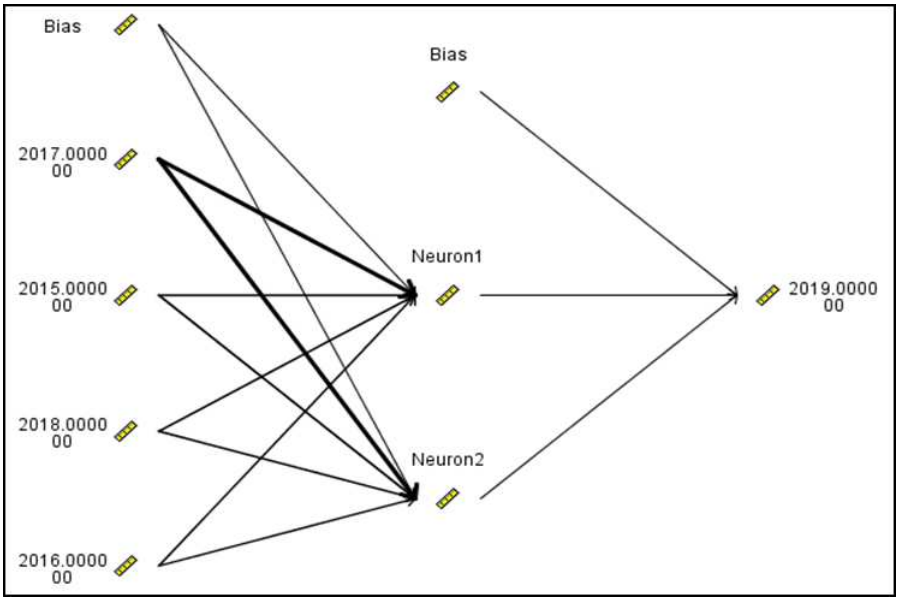

Figure 14. ANN Structure for 2019 Unleaded Gasoline Sales Forecast

When the given neural network structure is examined, the number of neurons used in the hidden layer in the estimation of each species is different. A neuron usually receives many inputs simultaneously. Each entry has its own weight. The weights indicated in bold in the artificial neural networks given in the figure indicate that the inputs are more important than the other input values in the neurons to which they are connected. For this reason, they are more effective in generating value at the neuron. In addition, weights are adaptive coefficients that determine the strength of the input signal. That is, it is a measure of the connecting power of the input. The network with the smallest structure in neural network structures is the model created for the estimation of

unleaded gasoline consisting of 2 neurons. As a result of the network structure created, the values which have been estimated for the sale of fuel types are given in Table 2.

Table 2. Fuel Sales Values Estimated by MLP

2019 (Months)	Realized Fuel Sales			Predicted Values with ANN Multiplayer Perceptron		
	Diesel	Unleaded	Autogas	Diesel	Unleaded	Autogas
January	88212	15638	80831	89123,43	15228,36	80803,07
February	105957	16348	71306	102207,7	16466,72	71293,02
March	103705	18815	81264	101761,3	18732,89	81466,35
April	90698,4	19987	77399	92848,63	20251,71	77598,87
May	105315	22509	65882	95068,62	20951,84	65127,58
June	89905	24520	60551	93896,42	22667,70	54045,15
July	93807,39	26059,98	55156,3	96289,93	24828,57	50842,81
August	89671	25426	50572	100855,4	24930,64	49371,31
September	92349	21132	44411	96906,84	22773,59	46850,64
October	85277	21225	49662	78235,11	22306,36	50721,26
November	70872	19772	49460	73861,42	22537,85	53750,39
December	77969	19009	47633	81198,87	19663,63	52011,7

The values predicted by the MLP algorithm and the actual values were examined on the line chart, and as a result of the comparison, it was observed that the algorithm made close predictions especially for January, February, March and April. The graph of the comparison of these values is given in Figure 15.

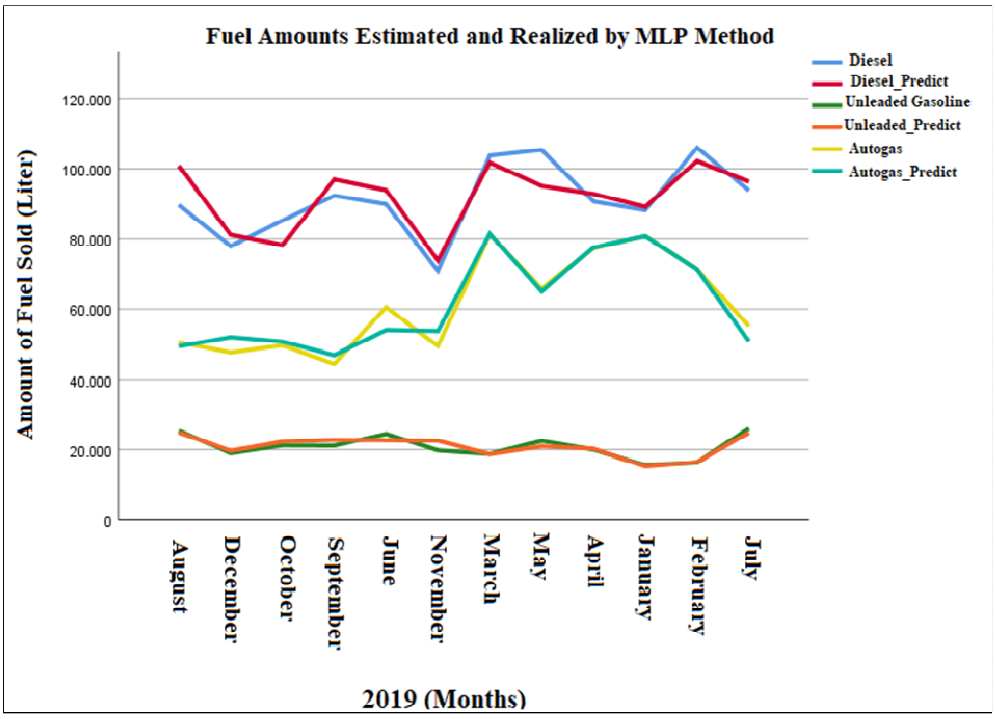

Figure 15. Fuel Amounts Estimated and Realized by MLP Method

The RBFN (Radial Basic Function Network) method used as another prediction algorithm is a curve fitting approach depending on the values of the target variable in multi-dimensional space. Training of the model requires less time and more data is needed to get good results [16]. The number of neurons in the hidden layer of artificial neural networks created for the prediction process in the RBFN algorithm differs as in the MLP algorithm. The number of neurons in the hidden layer is 6 in the network created for unleaded gasoline and 8 for diesel and auto gas. As a result of the created structures, the values estimated by the network regarding the sales of fuel oil types are given in Table 3.

Table 3. Estimated Values with RBFN

2019 (Months)	Realized Fuel Sales			Predicted Values with ANN Radial Basic Function		
	Diesel	Unleaded	Autogas	Diesel	Unleaded	Autogas
January	88212,00	15638,00	80831,00	88212,00	15992,86	80831,00
February	105957,00	16348,00	71306,00	105957,00	15993,11	71306,00
March	103705,00	18815,00	81264,00	103705,00	18911,43	81264,00
April	90698,40	19987,00	77399,00	90698,40	19987,09	77399,00
May	105315,00	22509,00	65882,00	105315,00	22509,00	65882,00
June	89905,00	24520,00	60551,00	94405,37	22519,80	58055,79
July	93807,39	26059,98	55156,30	93807,39	26059,98	55156,30
August	89671,00	25426,00	50572,00	98165,59	26236,93	51742,59
September	92349,00	21132,00	44411,00	92349,00	21132,00	44411,00
October	85277,00	21225,00	49662,00	89522,26	22520,82	54380,78
November	70872,00	19772,00	49460,00	73465,38	22036,46	53131,30
December	77969,00	19009,00	47633,00	77969,00	18912,51	47633,00

When the actual sales values and the estimated values were compared, it was observed that the algorithm found the same values especially in January, February, March, April, May and September.

2) Sales Forecasting with Multiple Linear Regression: In the multiple linear regression model, the aim is to explain the total change in independent variables (explanatory variables) and in the dependent variable (response variable). In multiple linear regression analysis, the contribution of some of the model-forming independent variables to the model may be insignificant. Therefore, it is necessary to determine the independent variables that will explain the dependent variable in the most appropriate way and to removal the insignificant variables from the model. This process is called "variable selection" [17]. Different methods have been developed for the selection of variables. These methods are examined in two groups as classical methods and step methods according to calculation techniques.

At this stage, while performing the estimate of the sales, the forward, backward and stepwise methods were used. In the forward selection method, it wants to find the most suitable regression model by adding one independent variable at a time. In the backward selection method, all variables are included in the model in the first stage. In the next steps, the process is continued by throwing the independent variable, which has the lowest partial F value. The contribution of the discarded variable is tested each time. If the contribution of the discarded variable is statistically significant, the throw is not performed and the process is stopped there [18]. In the stepwise selection method, both the forward selection method and the reverse selection method are used simultaneously [19]. The stepwise selection method consists of editing the forward selection method. The independent variable previously added to the model is re-evaluated with partial F statistics. An argument previously added to the model can be removed from the model in later steps [20]. In this method the purpose is, to determine what the independent variables that may affect Y dependent variable are theoretically, and to

choose among them the variables that do not have interrelations and affect the dependent variable the most. The most important benefit of the stepwise selection method is that it solves the problem of multiple linear connection [21]. The estimated of values as a result of the methods used are given in Table 4.

Table 4. Estimated Values by Multiple Linear Regression Methods

2019 (Months)	Realized Fuel Sales			Values Estimated by Forward Regression			Values Estimated by Backward Regression			Values Estimated by Stepwise Regression		
	Diesel	Unleaded Gasoline	Autogas	Diesel	Unleaded Gasoline	Autogas	Diesel	Unleaded Gasoline	Autogas	Diesel	Unleaded Gasoline	Autogas
January	88212,00	15638,00	80831,00	96505,08	15820,27	69759,807	96505,08	16115,43	70421,34	93421,41	15820,27	70710,66
February	105957,00	16348,00	71306,00	10032,42	16291,84	82540,384	100392,4	16048,42	78399,15	100300,6	16291,84	77989,78
March	103705,00	18815,00	81264,00	99579,05	17811,47	73380,077	99579,05	17789,66	80136,61	99054,94	17811,47	80029,25
April	90698,40	19987,00	77399,00	92894,64	20792,27	60210,418	92894,64	21590,9	75337,39	93022,15	20792,27	75806,43
May	105315,00	22509,00	65882,00	91767,21	20526,10	63582,079	91767,21	20941,49	67813,21	93177,95	20526,1	67210,57
June	89905,00	24520,00	60551,00	93700,51	22591,09	58874,588	93700,51	22571,87	56507,57	93544,59	22591,09	56477,82
July	93807,39	26059,98	55156,30	95019,53	24931,87	54295,876	95019,53	24570,35	47646,7	94597,71	24931,87	47392,57
August	89671,00	25426,00	50572,00	95017,22	25594,08	46009,174	95017,22	25455,13	51880,93	96970,79	25594,08	53241,5
September	92349,00	21132,00	44411,00	94989,81	22694,49	55856,765	94989,81	23106,18	48807,7	95021,45	22694,49	48131,14
October	85277,00	21225,00	49662,00	81050,63	21965,77	54661,673	81050,63	21503,26	47616,31	81574,27	21965,77	47716,74
November	70872,00	19772,00	49460,00	71214,16	21577,36	54685,903	71214,16	21506,81	57508,01	69046,83	21577,36	58343,71
December	77969,00	19009,00	47633,00	81607,53	19844,37	60270,556	81607,53	19241,48	52052,39	84005,13	19844,37	51077,15

3) Comparison of the Results of Artificial Neural Networks and Multiple Linear Regression Methods: Artificial neural network methods and multiple linear regression methods, the compared among themselves in order to estimate the 12-month fuel sales of 2019 year and the best predictive models were determined. The results of artificial neural network methods are given in Table 5.

Table 5. Artificial Neural Networks Methods and Results for Estimation of Fuel Types

	Autogas		Unleaded Gasoline		Diesel	
	Multiplayer Perceptron	RBFN	Multiplayer Perceptron	RBFN	Multiplayer Perceptron	RBFN
Minimum Error	-4378,704	-4718,783	-2765,856	-2264,461	-11184,44	-8494,589
Maximum Error	6505,853	2495,213	1852,291	2000,205	10246,384	0
Mean Error	20,43	-588,788	-74,912	-197,584	-709,654	-1652,8
Mean Absolute Error	2115,466	1004,657	1012,903	606,181	4539,877	-1652,8
Standard Deviation	3128,671	1887,317	1339,996	1012,08	5714,032	2769,612
Linear Correlation	0,974	0,991	0,915	0,953	0,841	0,965

When the results in Table 5 are analyzed, it is seen that the most adaptive method for 2019 fuel types is RBFN. The values realized with the values estimated of the RBFN method are given in the histogram graph (Figure 16).

When the histogram graph in Figure 16 is examined, it is observed that the intake of diesel oil generally increases in winter and spring and decreases in autumn. It is observed that purchases for unleaded gasoline increase partially in winter and spring, decrease in summer and autumn. It is observed that purchases for autogas, another type of fuel, increase in winter and spring, decrease in summer and autumn.

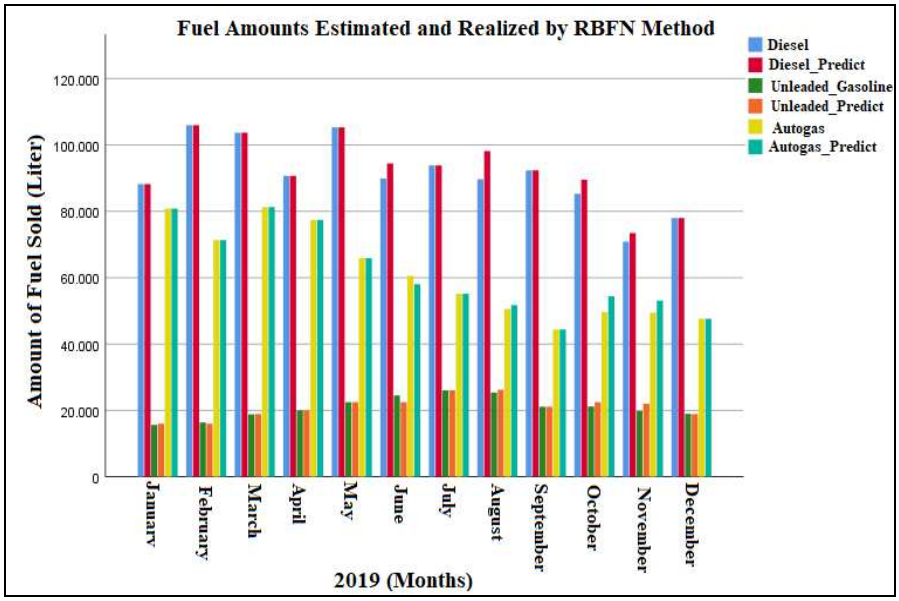

Figure 16. Fuel Amounts Estimated and Realized by RBFN Method

In order to find out which of the multiple linear regression methods produce values closer to the real values, the R^2 values of the methods were examined. R^2 is a commonly known measure and is a measure of the fit rate of the linear model. It is often called the determination coefficient. It ranges from 0 to 1, and small values indicate that the model is not accord [16]. The R^2 values of the methods used in each type of fuel are given in Table 6.

Table 6. Multiple Linear Regression Estimation Methods R^2 Values

	Multiple Linear Regression Estimation Methods R^2 Values		
	Forward	Backward	Stepwise
Autogas	0,525	0,832	0,832
Unleaded	0,855	0,84	0,855
Diesel	0,682	0,682	0,703

When R^2 values are analyzed, it is observed that backward and stepwise linear regression for autogas, forward and stepwise linear regression for unleaded gasoline and stepwise linear regression method it appears to estimate values close to real values.

III. RESULTS

In this case study, an automation software for a fuel-oil station has been created; the data taken from this automation software has been transferred to a database and then analyzed by data mining techniques. 48-month (2015-2018) data were used for the analysis, and it was aimed to estimate the sales expected to occur in 2019 based on the data. In the analysis, it is aimed to predict fuel sales and determine the model that makes the

best estimation by using two different data mining techniques and their different models. Due it is a web based project, a synchronized access to the system through different stations was made aimed. The dispersion graphics of the data, success rates of the used algorithms, and the outputs produced by the algorithms have been obtained by using data mining technics. Thanks to the line graphic which is occurred by analyzing the data, it is inferred that the most sold fuel-oil type is the diesel. It is seen that the LPG and unleaded types of fuel-oil at the station is the sold fuel in the second order.

The project developed in the study makes the people who want to make future planning and forecasting in the fuel-oil sector, to make the accurate predictions easy by data mining. Based on the data obtained from the analysis, it has been reached to the result that the station should prefer to have the diesel fuel mostly; then unleaded and LPG type ones respectively.

To future works, it is thought that the interface part of the software could be developed in order to take attentions of users. To more precise predictions by data mining on future, it is proposed to use monthly sales data of other stations of the company. Thus, subtractions on how much the station should be supplied regarding to the types of fuel-oil by examining the increase and decrease at sales could be made.

The stations to be closed can be determined by looking at the sales. It can be benefited that the more accurate subtractions (results) could be obtained on prediction processes by different algorithms.

ACKNOWLEDGEMENTS

The author(s) of the paper thank to Technology Faculty of Muğla Sıtkı Koçman University for its valuable supports during progress of the study.

REFERENCES

[1] Eroglu, H. “Stock Control of Fuel Products and Evaluation of Wastes in Fuel Stations”, İstanbul, Turkey: Beykent University Graduate School of Social Sciences, 2019.

[2] Skeet, J., “C# in Depth, Manning Publications Company”, 4th ed., ISBN-13: 978-1617294532, 497 pages, 2019.

[3] Esposito, D., “Programming ASP.NET Core (Developer Reference)”, ISBN: 978-509304417, 416 pages, Publisher: Microsoft Press; 1st Edition, USA, May 19, 2018.

[4] Pala, Z. “Step by Step Web Applications by ASP.NET”, Turkmen Book-House, ISBN-13: 978-9756392485, 359 pages, 2006.

[5] Sonmez, E., Kacar, S., *Web Laboratory Design Based Matlab Builder Ne and Asp.Net for Control Systems Course*, Sakarya University Journal of Science, vol. 20, iss. 2, p.p. 155-165, 2016.

[6] [Online]. Available: https://jamshidhashimi.com/net-ve-c/ 26.08.2019.

[7] Yildiz, M., Seker, S. E., *Data Mining Tools*, YBS Encyclopedia, vol. 3, no. 4, p.p. 10–19, 2016.

[8] Silahtaroglu, G., “Data Mining Concepts and Algorithms”, ISBN: 9756797815, 3rd Publishment, 304 pages, Turkey, 2016.

[9] Kocaman, A.E., “Sentiment Analysis on Twitter by Data Mining Technics”, B.Sc. Thesis, Technology Faculty of Mugla Sitki Kocman University, Mugla, 2018.

[10] Atan, S., *Data, Big Data and Business Administration*, Balikesir University The Journal of Social Sciences Institute, vol. 19, no. 35, p.p. 137–153, 2016.

[11] Zhong, N., Zhou, L., "Methodologies for Knowledge Discovery and Data Mining", in Proc. Third Pacific-Asia Conference, Pakdd-99, p. 26-28, China, 1999.

[12] Akpinar, H., *Knowledge Discovery and Data Mining in Databases*, İstanbul Univ. Journal of Business Administration, vol. 29, no. 1, p.p. 2000.

[13] [Online]. Available: https://www.ibm.com/tr-tr/products/spss-modeler 27.02.2020

[14] Silahtaroğlu, G., Data Mining, İstanbul:Papatya Publishing, 2016

[15] Bulut, F . (2016). *The Right Professional Preference with Multilayer Perceptron*. Anadolu University Journal of Science and Technology A - Applied Sciences and Engineering , 17 (1) , 97-109 . DOI: 10.18038/btda.45787

[16] Irmak, S., Köksal, C. D., Asilkan, Ö., (2012) *Predicting Future Patient Volumes of the Hospitals by Using Data Mining Methods*. International Journal of Alanya Faculty of Business, 4 (1),101-114.

[17] Alpar, R. (2003) *Introduction to Applied Multivariate Statistical Methods* 1. 2nd Edition, Nobel Book-House, Ankara.

[18] Efe, E., Bek, Y., Şahin, M. (2000) *Statistical Methods with Solutions in SPSS II*, Kahramanmaraş Sütçü İmam University Rectorate Publication No: 10, Kahramanmaraş.

[19] Çakır Zeytinoğlu, F. (2007) *Effects of Current Assets of Businesses on Sales: Selection of the Best Regression Equation and Sectoral Comparison*. Marmara University İ.İ.B.F. Journal, 23 (2): 331-349.

[20] Montgomery, D.C., Peck, E.A., Vining, G.G. (2001) Introduction to Linear Regression Analysis, 3rd Edition, John Wiley & Sons, New York

[21] Işık, A.(2006) Applied Statistics-II. Beta Edition, İstanbul.