© by PSP Volume 30– No. 08/2021 pages 9731-9739 Fresenius Environmental Bulletin

DETERMINATION OF CALORIE VALUE OF COAL BY USING MACHINE LEARNING METHODS

Mehmet Kayakus*

Management Information Systems, Faculty of Social Sciences and Humanities, Akdeniz University, Antalya 07600, Turkey

ABSTRACT

In this study, the calorie value of coal was estimated using three different machine learning methods. For this, sulfur, volatility and ash values of 1665 coal samples were used as independent variables for the determination of calories. Artificial neural networks, support vector regression and multiple linear regression methods were used as estimation methods. At the end of the study, according to the R2 performance measurement, the most successful estimation methods were ANN (0.891), SVR (0.887) and MLR (0.882), respectively. According to RMSE analyzes, SVR (0.017), MLR (0.028) and ANN (0.029) were most successful, respectively. According to the MAE value; ANN (1.3%), SVR (1.7%) and MLR (1.8%) were most successful, respectively. The study can be considered successful when compared with similar studies in the literature.

KEYWORDS:
Coal, calories, machine learning, artificial neural networks, support vector regression, multiple linear regression

INTRODUCTION

Coal is one of the three main energy sources in the world, accounting for one-third of the world's total energy use. Up to 40% of electricity is produced from coal. From the perspective of the scarcity of energy resources, the current reserves of global coal resources are expected to last for about 153 years, much longer than oil reserves of about 50.6 years and natural gas reserves of about 52.5 years [1]. Turkey has a total 15 gigaton lignite reserve and 8.5 gigaton of total lignite reserve is ready to be extracted [2].

Coal is a flammable sedimentary rock of endless variation, composed of layered plant remains consolidated under overlapping layers [3]. Coal is a kind of complex mineral composed of inorganic minerals and organic substances as components [4, 5]. On the other hand, the content and type of these mineral substances affect the transportation and storage, development and use of coal during the coal utilization process [6-8]. Organic ingredients are essential in defining the nature of coal and its value in different usage processes [9]. Depending on the maturation stage of the coaling process, coal is mainly composed of varying proportions of carbon, oxygen, nitrogen and sulfur. Graphite, which is essentially the endpoint of carbonization, contains only carbon [10]. Many industrial products can be produced from coal and are used mainly in thermal power plants, metals and chemicals [11, 12].

By means of coal analysis, both its quality and economic value are determined. Quality coal provides less harmful and environmentally friendly heating. Coals with high calorific value are more preferred and their economic return is higher. In coal analysis studies, the calorie, sulfur, volatility and ash values of coal are measured. Countries subject coals to analysis before selling, both to protect the environment and to prevent the introduction of coals with poor performance.

Calorific value in coal is an important feature that shows the useful energy content of coal and therefore its value as fuel [13]. Technically, the calorific value is defined as the amount of heat generated when the unit weight of coal is completely burned and the combustion products cooled to a standard temperature of 298^{o}K. [14]. The calorific value of charcoal is measured experimentally with sophisticated equipment and a trained chemist [15]. The magnitude of the gross calorific value (GCV) varies significantly depending on the ash, moisture content and the type of coal [14].

Sulfur in coal is found in inorganic and organic forms. Inorganic forms are usually sulfides and sulfates. Analytical methods are available for the determination of these inorganic sulfur components and several commercial processes are available for removing most of the inorganic sulfur from coal [16]. The sulfur content in coals varies considerably, but most commonly the total sulfur content is in the range of 0.5% to 5%. Coal with less than 1% sulfur is classified as low sulfur coal. Coal containing 1% to 3% sulfur is classified as medium sulfur coal, and coal with more than 3% sulfur is classified as high sulfur coal [17].

Volatile Organic Compounds (VOCs) found in coal are molecules containing at least one carbon and one hydrogen atom (i.e. organic compounds) that evaporate easily (i.e. volatile) at room temperature [18]. Most VOCs are toxic and can participate in a

photochemical reaction that leads to smoke generation [19]. Coal has a significant amount of VOCs in its structure [20].

Ash content is one of the most important indexes of coal quality. The standard measurement method of coal ash content is the time-consuming combustion-weighing method, that is, obtaining the impurity percentage according to the weight ratios before and after combustion. In the coal industry, rapid ash measurement is often used to monitor the quality of coal products. However, about 1 hour is still required for manual sampling, sample preparation, and analysis; this leads to delayed information feedback and unqualified coal products [21].

There are many studies in the literature analyzing coal with artificial intelligence techniques. Tan et. al. carried out an analysis of coal using a nonlinear support vector regression (SVR) model. For this, 167 Chinese and 4540 US coal samples were used. As a result of their studies, the average absolute error value was 2.16% for Chinese coal; 2.42% for US coal [22]. Parikh et. al. developed a simple correlation to estimate the calories of coal, biomass and other solid fuels. As a result of their studies, the mean absolute error of their correlations was 3.74% [23]. Patel et al. designed seven nonlinear artificial neural networks model to estimate the caloric value of coal. The average correlation coefficient of the models is greater than 0.984; and the RMSE value was 0.665 [14]. Majumder et. al. developed correlation based on close analysis using 250 coal samples to estimate the calories of coal. The overall mean error between the actual and predicted data was 13.13%, and this value was considered as a high value [15]. Mesroghli et. al. estimated the regression by using 4540 US coal samples and calorie value of coal using the method of artificial neural networks. As a result of their studies, they concluded that it is better to use regression instead of ANN for estimation [24]. Huang et al. estimated the supply characteristics of karstic water in coal mines and its main influencing factors using artificial neural networks. The correlation coefficient in the study was 0.8272 [25].

Here, the calorie value of coal is considered as the input variable for determining machine learning. 1665 coal samples taken from Turkey for the estimations in the study were used as the data set. Artificial neural networks, support vector regression and multiple linear regression methods were used as estimation methods.

MATERIALS AND METHODS

Data Set. This study focuses on estimating the calorie value of coal using machine learning methods. For this, sulfur, volatility and ash values were accepted as independent (input) variables for determination of calories. 1665 coal samples taken from Turkey for the estimations in the study were used as the data set. When acquiring data, SC 144 for sulfur measurements; TGA 701 devices for volatility and ash, and AC 500 devices were used for calories. The sample data set is given in Table 1.

TABLE 1
The sample data set

Sulfur	VOCs	Ash	Calorific value
0,2577	22,15	6,04	7570,82
0,25768	22,2	6,02	7606,58
0,2501	15,58	4,34	7855,07
0,2641	15,72	7,8	7500
0,2632	15,28	7,78	7461,32
0,2729	15,38	6,75	7461,43
0,27361	15,1	6,81	7533,93
0,2711	17,71	9,59	7319,97
0,2774	17,69	9,65	7368,25
0,2595	17,17	7,48	7414,55
0,2609	17,15	7,48	7405,3
0,2576	18,19	5,29	7705,42
0,2549	17,64	5,31	7669,47
0,4208	30,13	17,12	6109,37
0,4292	30,19	17,18	6134,08
0,9159	23,91	8,91	7286,71
0,8831	23,24	8,89	7248,16
0,2535	19,33	6,53	7449,98
0,2525	18,86	6,54	7460,25
0,6279	29,77	14,2	6393,81
0,6265	30,34	14,16	6401,52

In the study, using the cross-validation method, the same data samples were grouped differently to increase the amount of data. Cross validation is a technique widely recognized in the field of statistical data analysis. There are two commonly used methods which are k-bit verification and leave-one-out [26, 27]. The K-bit verification method is divided into k discrete parts in a data set containing a total of n samples, each containing $n\ k$ samples. Each time a different set of data is reserved for testing and the remaining k-1 data set is used for training. The classifier is trained k times by changing the test set each time. Classifier performance is estimated with the average of k errors obtained in this way [28]. According to leave-one-out (LOO) method, each sample in the data set is excluded once and determined as test data, and the remaining samples are modeled as a classifier as training data. In this way, the classification success rate is calculated as much as the number of samples. In the last stage, the arithmetic mean of these success rates is taken and the verified classification success rate with the LOO method is calculated [29].

In the study, the normalization method was used to improve the performance of machine learning methods and increase the accuracy rate. Normalization method includes min-max, z-score, median and sigmoid. In this study, using the min-max normalization method, the original data were normalized with linear conversion in the range of 0-1. In

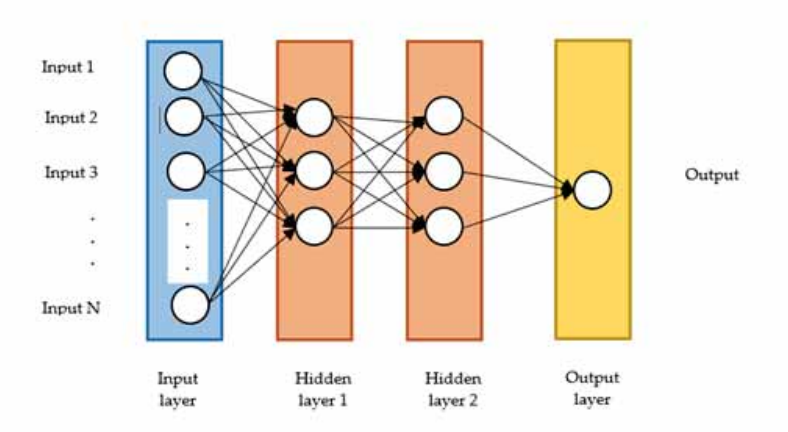


FIGURE 1
The artificial neural network model

Equation 1, the min-max normalization formula is given:

$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

x refers to 'normalized data, x_i refers to the input value; x_{min} refers to the smallest number in the input set, x_{max} refers to the largest number in the input set.

Forecasting models. Artificial Neural Network. Artificial neural networks (ANN) are computer systems developed to simulate the processes of the human brain by mathematically modelling the neuro-physiological structure [30]. ANN is capable of "learning"; that is, it can be trained through supervised or unsupervised learning to improve its performance [31]. Artificial neural networks are categorized according to their architecture (number of layers), topologies (connection model, feed forward or repetitive, etc.) and learning type. Most of the applications use feed forward networking and error back propagation learning.

The feed forward network is generally composed of a hierarchical structure consisting of three layers defined as input, hidden and output layers. Each node in the input tier is linked to all nodes in the hidden tier using weighted links. The weights on the network connections determine the parameters of the data fitting model. The number of nodes in the input layer is equal to the number of inputs in the system to be modelled. The number of nodes and sizes in hidden layers are adjustable parameters that are determined according to the desired prediction accuracy and performance of the ANN mode [14, 32]. Output layer generates information calculated according to input values and selected parameters. Data enters the network through the input layer, travels through the hidden layer, and leaves from the output layer. Each hidden layer and output layer node collects data from the nodes above it (input layer or hidden layer) and applies an activation function. The activation functions can occur in several forms. The type of activation function is indicated by the state of the neuron within the network. In most cases, input layer neurons do not have an activation function because their role is to transfer inputs to the hidden layer [33]. Figure 1. The artificial neural network model is seen.

The back propagation learning algorithm consists of two procedures as forward feed and back-propagation weight training. The input data set is spread over the network to give an estimate of the output. It is used to update the weights systematically according to the error information in the estimate [34]. The network is trained by changing the weights until the error between the training data outputs and the predicted outputs of the network is small enough [24].

In an artificial neural network model, when the information in the input cells is transferred to the intermediate layer cells, it is multiplied by the relevant weights and the NET input of each intermediate layer cell is calculated according to the weighted sum function:

$$NET_p = \sum_i w_{ip} x_i \quad (2)$$

Here, the input into the NET p intermediate layer p. cell represents i. cell of w_{ip} input layer, the weight that connects to intermediate layer p. cell represents, the output of the i. cell in the x_i input layer. The output of the middle layer cells is found by passing through the NET input found from an activation function.

$$F_p = f(NET_p) \quad (3)$$

Here, F_p intermediate layer shows the activation function of the cell p. The activation function to be used must be differentiable.

After this step, the error calculation is made by using the target value and the value in the output cell. Error calculation:

$$\varepsilon_p = (T_p) - (F_p) \quad (4)$$

ε_p shows the error made; T_p targeted output value; F_p p. the calculated value of the cell.

Support Vector Regression (SVR). Support vector regression (SVR) is a universal learning algorithm proposed by Vapnik [35, 36]. SVR is based on

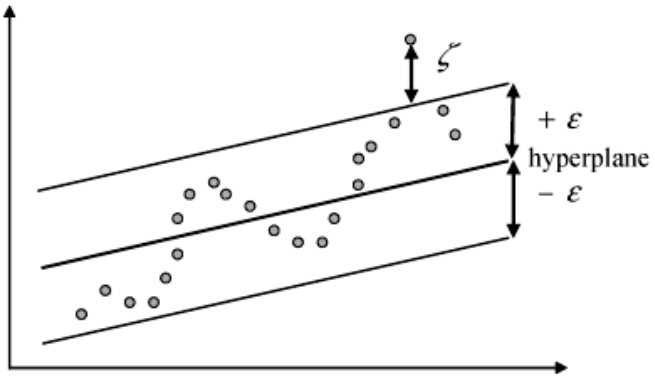


FIGURE 2
The nonlinear support vector regression [42]

a statistical learning theory that allows solving learning problems with a small number of examples [22]. In addition, SVR has a strong ability to handle nonlinear problems and is successfully used in a wide variety of fields [37-40].

Support vector regression began to be used as a powerful approach to predict and solve linear and nonlinear systems with the advent of Vapnik's ε-insensitive loss function [41].

Considering that there is a data set like in Equation 5:

$$D = \{(x^1, y^1), \dots, (x^l, y^l)\}, \quad x \in R^n, y \in R \tag{5}$$

Linear function for support vector regression:

$$f(x) = w \cdot \phi((x) + b \tag{6}$$

Here w and b are the weight vector and bias, respectively, and ϕ (x) is the kernel function.

The nonlinear support vector regression is shown in Figure 2.

The general SVR prediction function is as follows:

$$f(x) = \big(w \cdot \Phi(x)\big) + b \tag{7}$$

Our goal is to find the value of ω and b, so that the x values can be determined by minimizing the risk of regression.

$$R_{reg}(f) = C \sum_{i=0}^{\ell} \Gamma(f(\chi_i) - y_i) + \frac{1}{2}\|w\|^2 \tag{8}$$

Here, Γ(.) is a cost function, C is constant and can be written in terms of vector ω data points as:

$$w = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*)\Phi(x_i) \tag{9}$$

For a SVR, choosing an appropriate core is imperative for the success of the learning process. The regularization parameter, which controls the trade-off between the complexity of an SVR and the number of inseparable points, is also important. SVR kernels include linear kernel, polynomial kernel, index nucleus, dual neural network nucleus, and radial elementary function (RBF) nucleus. In these cores, the RBF core is the most popular core function and has better overall stability [43].

Radial basis function (RBF) kernel:

$$K\big(x_i, x_j\big) = \exp(-\gamma\big|(x_i \cdot x_j\big|^2) \tag{10}$$

The ε-insensitive loss function is the most used cost function. Function:

$$\Gamma(f(\chi_i) - \gamma) = \begin{cases} |f(x) - \gamma| - \varepsilon, & for |f(x) - \gamma| \geq \varepsilon \\ 0, & otherwise \end{cases} \tag{11}$$

By solving the quadratic optimization problem, regression risk and e-insensitive loss function can be minimized.

$$\frac{1}{2}\sum_{i,j=1}^{\ell} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)k(x_i, x_j) - \sum_{i=1}^{\ell} \alpha_i^*(y_i - \varepsilon) - \alpha_i(y_i + \varepsilon)$$

Subject to (12)

$$\sum_{i=1}^{\ell} \alpha_i - \alpha_i^* = 0, \qquad \alpha_i, \alpha_i^* \in [0, C]$$

Lagrange multipliers α_i and α_i^* represent the solutions to the second order problem above as forces pushing the predictions towards the target value of y_i. Only non-zero values of Lagrange multipliers are useful in predicting the regression line and are known as support vectors. The constant C determines the penalties for estimation errors. The value of w is solved according to the Lagrange multipliers. For variable b, it can be calculated by applying Karush – Kuhn – Tucker (KKT) conditions, which means that the product of Lagrange multipliers and constraints must be equal to 0.

$$\alpha_i(\varepsilon + \xi_i - y_i + (w, x_i) + b) = 0$$
$$\alpha_i^*(\varepsilon + \xi_i^* + y_i - (w, x_i) - b) = 0 \tag{13}$$
$$(C - \alpha_i)\xi_i = 0$$
$$(C - \alpha_i^*)\xi_i^* = 0 \tag{14}$$

Where ξ_i and ξ_i^* are slack variables used to measure errors outside the ε tube. Since α_i, $\alpha_i^* = 0$ and $\xi_i^* = 0$ for $\alpha_i^{\alpha}\text{Ɛ}$ (0, C), b can be computed as

$$b = y_i - (w, x_i) - \varepsilon \qquad for\ \alpha_i \in (0, C)$$
$$b = y_i - (w, x_i) + \varepsilon \qquad for\ \alpha_i^* \in (0, C) \tag{15}$$

It is necessary to find the ideal value by trying the values of penalty C and Radius ε, which are core functions.

Multiple Linear Regression (MLR). A simple linear regression shows the relationship between dependent variable y and independent variable x according to the regression equation. According to the

multiple linear regression model, the dependent variable is related to two or more independent variables. The general model for the K variable is as follows [44]:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \quad (16)$$

y dependent variable, x independent variable, β_i, with respect to x_i represents the regression parameter, and ε represents the error term. Error term ε has a mean value equal to zero and constant variance [45].

The difference between the true value of y and the predicted y value tends to an average of 0. Therefore, if the probability distributions of the dependent variable y at various levels of the independent variable are normally distributed (bell-shaped), the error term can be assumed to be the mean or expected value of 0. Therefore, the error term can be omitted in the calculation of parameters.

In the multiple linear regression analysis, the least squares method is used to estimate the regression coefficients. Regression coefficients indicate the irrelevant contributions of each independent variable to predict the dependent variable. Unlike simple linear regression, inferences should be made about the degree of interaction or correlation between each of the independent variables.

The Evaluation of Models. In comparing the success of artificial neural networks, support vector machines and multiple linear regression models, corrected coefficient of determination (R2), mean absolute error (MAE), mean squared error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE) criteria were used.

The equations for these criteria are given below.

$$R^2 = 1 - \frac{\sum(Y_i - Y_i^*)^2}{\sum(Y_i - Y_{ave})^2} \quad (17)$$

$$MAE = \frac{\sum_{i=1}^{n}|Y_i - Y_i^*|}{n} \quad (18)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 \quad (19)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \quad (20)$$

$$MAPE = \frac{\sum_{i=1}^{n}\left|\frac{Y_i - Y_i^*}{Y_i}\right|}{n} \cdot 100 \quad (21)$$

According to these criteria, high R2 and low MAE, MSE, RMSE and MAPE values determine the model that fits well.

CONCLUSION AND DISCUSSION

A total of 1665 coal sample data were used in the study. The Knime program was preferred for the application and analysis of artificial intelligence techniques. In the artificial neural network model, 70% of the data is transferred to the designed network, which is randomly selected for training and 30% for testing purposes. Models with different ANN parameters were tested and the model with the highest performance was decided. Statistical data of success and error of ANN designs are given in Table 2.

A value of R2, which indicates how well the data fits a linear curve, indicates that the test data is provided with a linear curve. According to the ANN method, in the estimation of the calorific value of coal the highest R2 value was 0.891; the lowest 0.871 and the average was 0.880. According to these results, it can be accepted that it is close to the desired values and can be considered successful. In the study, MSE values are ideal because they are very close to 0. The RMSE value is desired to be close to zero; in this study, the average is 0.030 and it is seen that it is very close to the desired value. For MAE and MAPE, values with the smallest value are considered ideal. In the study, The MAE was 0.014; MAPE was 0.012 and appears to provide the desired values.

According to the ANN method, the most successful method was the model consisting of 2 hidden layers and 3 neurons in each layer after 1000 iterations. According to this model, R2 was 0.891, MSE 0.001, RMSE 0.029, MAE 0.013 and MAPE was 0.013.

TABLE 2
Statistical data of success and error of ANN designs

ANN Models			R2	MSE	RMSE	MAE	MAPE
Iteration	Hidden layer	Neuron					
500	2	2	0,871	0,001	0,031	0,016	0,012
	2	3	0,883	0,001	0,03	0,014	0,014
	3	2	0,886	0,001	0,029	0,014	0,013
	3	3	0,880	0,001	0,03	0,014	0,012
	4	2	0,863	0,001	0,032	0,015	0,013
	4	3	0,878	0,001	0,03	0,014	0,011
1000	2	2	0,882	0,001	0,03	0,015	0,015
	2	3	0,891	0,001	0,029	0,013	0,013
	3	2	0,881	0,001	0,03	0,014	0,012
	3	3	0,876	0,001	0,03	0,014	0,011
	4	2	0,885	0,001	0,029	0,014	0,013
	4	3	0,889	0,001	0,029	0,014	0,012
Average			0,880	0,001	0,030	0,014	0,012

TABLE 3
The performances of SVR method for different overlapping penalty

Overlapping penalty	R2	MSE	RMSE	MAE	MAPE
1	0,829	0,001	0,029	0,022	0,031
2	0,827	0,001	0,029	0,021	0,030
5	0,887	0,001	0,024	0,017	0,024
10	0,800	0,001	0,032	0,021	0,030

TABLE 4
The performances of MLR models

Test ID	R2	MSE	RMSE	MAE	MAPE
1	0,889	0,001	0,025	0,018	0,022
2	0,904	0,001	0,024	0,017	0,023
3	0,868	0,001	0,029	0,019	0,022
4	0,879	0,001	0,026	0,019	0,024
5	0,893	0,001	0,026	0,019	0,024
6	0,894	0,001	0,025	0,018	0,022
7	0,895	0,001	0,025	0,018	0,022
8	0,876	0,001	0,026	0,019	0,023
9	0,873	0,001	0,028	0,020	0,025
10	0,851	0,003	0,054	0,021	0,024
Average	0,882	0,001	0,028	0,018	0,023

TABLE 5
The performances of three different machine learning methods

Machine learning methods	R2	MSE	RMSE	MAE	MAPE
Artificial neural networks	0,891	0,001	0,029	0,013	0,012
Support vector regression	0,887	0,001	0,017	0,017	0,024
Multiple linear regression	0,882	0,001	0,028	0,018	0,023

Nonlinear regression method was used as SVR method in the study. In the SVR method, tests were carried out using Polynomial, Hyper Tangent and Radial Basis Function (RBF) techniques, which are core functions, and Radial Basis Function (RBF) has been preferred as the core function since it gives a successful result. The overlapping penalty is a useful parameter when input data is not separable. It determines how many penalties will be given to each wrongly predicted point, thus increasing the accuracy rate. In this study, different overlapping penalty values were tested.

Table 4 shows the performances of the Radial Basis Core Function (RBF) according to the SVR method for different overlapping penalty values.

According to SVR method, R2 value is the highest in the estimation of the calorie value of coal, with the value 0.887; the lowest was 0.800. According to the results, it is seen that the SVR method does not vary much according to different parameters and provides successful results. MSE values were found to be ideal values with 0.001 in the study. The RMSE value is desired to be close to zero. In this study, it was found to be in the range of 0.024-0.032 and very close to the desired value. In the study, the MAE ranged from 0.017 to 0.022; MAPE between 0.024 and 0.031. For MAE and MAPE, values with the smallest value are considered ideal. The study is fairly close to the desired values for both the MAE and the MAPE value.

According to the SVR method, the most successful method is the model where the value of the overlapping penalty parameter is 5. According to this model, R2 was 0.887, MSE 0.001, RMSE 0.017, MAE 0.017 and MAPE was 0.024.

In the MLR method, 3 independent variables and a total of 1665 data were used in the entry. In the MLR method, input samples are randomly selected in the data set. Table 5 shows the coal calorie value estimation performances of the MLR method according to 10 different input data randomly selected.

According to the MLR method, the average R2 value in the estimation of the calorific value of coal was 0.882 the highest and it was observed that it was close to the desired values. The average MSE values in the study are 0.001 and it is ideal. The average RMSE value is 0.028 and it is seen that it is close to the ideal value. The average MAE in the study was 0.018; MAPE is 0.023 and it is seen that values are very close to ideal values.

The ideal and average results of the tests performed according to three different machine learning methods are shown in Table 5.

The graphical representation of the R2 value in Table 5 is shown in Figure 2.

Graphical comparison of MSE, RMSE, MAE and MAPE values in Table 5 is shown in Figure 3.

When the results of the analysis made according to three different artificial intelligence techniques are compared, it is seen that the results are close to each other and successful.

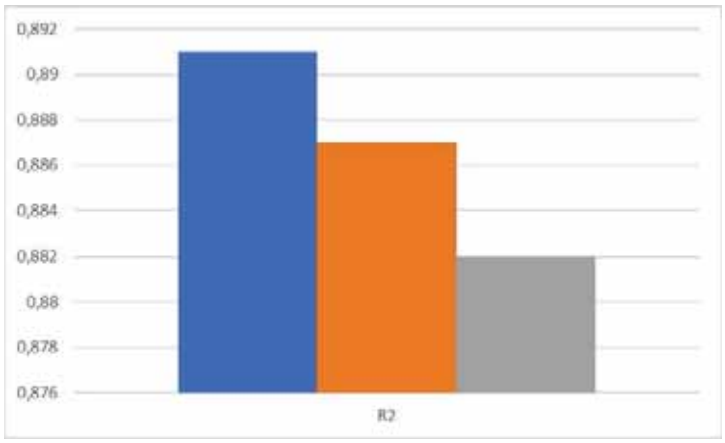


FIGURE 2
Analysis of the R2 value

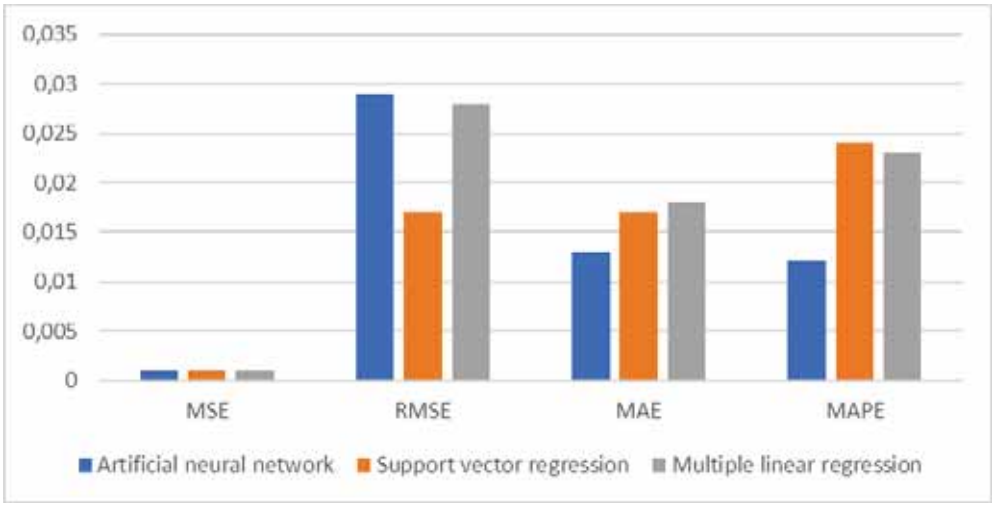


FIGURE 3
The comparison of MSE, RMSE, MAE and MAPE

CONCLUSION

Coal will continue to be an important source of energy throughout the world for many years. Quality coal is economically more valuable and less harmful to the environment. This ensures that coal is preferred when it is of good quality.

Today, expensive measurement systems are needed together with expert staff to determine the quality of coal. With the advances in artificial intelligence technologies, coal analysis and estimates of calorie values can be made via computers.

In this study, the calories of coal were successfully estimated by using three different machine learning methods on 1665 coal samples. According to the R2 performance measurement, the most successful estimation methods were ANN (0.891), SVR (0.887) and MLR (0.882), respectively. According to RMSE analyses, SVR (0.017), MLR (0.028) and ANN (0.029) were most successful, respectively. According to the MAE value, ANN (1.3%), SVR (1.7%) and MLR (1.8%) were most successful, respectively.

The study was considered to be successful when compared with similar studies for estimating coal calories in the literature. In their study, Tan et al., using the nonlinear support vector regression (SVR) model, the average absolute error (MAE) value for Chinese charcoal was 2.16% and 2.42% for US coal. The RMSE value was 0.665 in the artificial neural network model designed by Patel et al. to estimate the calorie value of coal. When compared with these studies, it is seen that the study is successful.

REFERENCES

[1] Song, Y. and Wang, N. (2019) Exploring temporal and spatial evolution of global coal supply-demand and flow structure. Energy. 168, 1073-1080.

[2] Karadirek, S., Ozcelik, O. and Altunsoy, M. (2019) Environmental Aspects and Coal Quality Parameters of Soma, Karapinar and Sahinali Coals in Turkey. Fresen. Environ. Bull. 28, 873-878.

[3] Sivrikaya, O. (2014) Cleaning study of a low-rank lignite with DMS, Reichert spiral and flotation. Fuel. 119, 252-258.
[4] Takagi, H., Maruyama, K., Yoshizawa, N., Yamada, Y., and Sato, Y. (2004) XRD analysis of carbon stacking structure in coal during heat treatment. Fuel. 83 (17-18), 2427-2433.
[5] Oliveira, M.L., Ward C.R., Izquierdo, M., Sampaio, C.H., de Brum, I.A., Kautzmann, R.M., Sabedot S., Querol, X. and Silva L.F.O. (2012) Chemical composition and minerals in pyrite ash of an abandoned sulphuric acid production plant. Science of The Total Environment. 430, 34-47.
[6] Dias C.L., Oliveira, M.L., Hower, J.C., Taffarel, S.R., Kautzmann, R.M. and Silva L.F. (2014) Nanominerals and ultrafine particles from coal fires from Santa Catarina, South Brazil. International Journal of Coal Geology. 122, 50-60.
[7] Garcia K.O., Teixeira, E.C., Agudelo-Castañeda, D.M., Braga, M., Alabarse, P.G., Wiegand, F., Kautzmann R.M. and Silva L.F.O. (2014) Assessment of nitro-polycyclic aromatic hydrocarbons in PM1 near an area of heavy-duty traffic. Science of the Total Environment. 479, 57-65.
[8] Zhang, L., Li, Z., Yang, Y., Zhou, Y., Kong, B., Li, J. and Si L. (2016) Effect of acid treatment on the characteristics and structures of high-sulfur bituminous coal. Fuel. 184, 418-429.
[9] Vassilev, S.V. and Tascón, J.M. (2003) Methods for characterization of inorganic and mineral matter in coal: a critical overview. Energy and Fuels. 17(2), 271-281.
[10] Potgieter-Vermaak, S., Maledi, N., Wagner, N., Van Heerden, J., Van Grieken, R. and Potgieter, J. (2011) Raman spectroscopy for the analysis of coal: A review. Journal of Raman Spectroscopy. 42(2), 123-129.
[11] Carraher Jr., C.E. (2003) Seymour/Carraher's polymer chemistry. CRC Press.
[12] Stanger, R., Tran, Q.A., Browne, M., Lucas, J., Naebe, M., Height, M. and Wall, T. (2018) Evaluating the thermal extrusion behavior of a coking coal for direct carbon fiber production. Energy and Fuels. 32(4), 4528-4537.
[13] Hower, J. and Eble, C. (1996) Coal quality and coal utilization. Energy Miner. Div. Hourglass. 30(7), 1-8.
[14] Patel, S.U., Kumar, B.J., Badhe, Y.P., Sharma, B.K., Saha, S., Biswas, S., Chaudhury, A., Tambe, S.S., Kulkarni, D. (2007) Estimation of gross calorific value of coals using artificial neural networks. Fuel. 86(3), 334-344.
[15] Majumder, A.K., Jain, R., Banerjee, P. and Barnwal, J. (2008) Development of a new proximate analysis based correlation to predict calorific value of coal. Fuel. 87(13-14), 3077-3081.
[16] Calkins, W.H. (1994) The chemical forms of sulfur in coal: A review. Fuel. 73(4), 475-484.
[17] Chou, C. (1997) Geologic factors affecting the abundance, distribution, and speciation of sulfur in coals. 30th Int'l Geol. Congr. 18, 47-57.
[18] Ni, J.-Q., Robarge, W.P., Xiao, C. and Heber, A.J. (2012) Volatile organic compounds at swine facilities: A critical review. Chemosphere. 89(7), 769-788.
[19] Hanks, J.C. and Louglin, S.O. (2011) Volatile organic compounds. Nova Science Publishers New York.
[20] Cheng, J., Zhang, Y., Wang, T., Xu, H., Norris, P. and Pan, W.P. (2018) Emission of volatile organic compounds (VOCs) during coal combustion at different heating rates. Fuel. 225, 554-562.
[21] Zhang, Z., Yang, J., Wang, Y., Dou, D. and Xia, W. (2014) Ash content prediction of coarse coal by image analysis and GA-SVM. Powder Technology. 268, 429-435.
[22] Tan, P., Zhang, C., Xia, J., Fang, Q.-Y. and Chen, G. (2015) Estimation of higher heating value of coal based on proximate analysis using support vector regression. Fuel Processing Technology.138, 298-304.
[23] Parikh, J., Channiwala, S. and Ghosal, G. (2005) A correlation for calculating HHV from proximate analysis of solid fuels. Fuel. 84(5), 487-494.
[24] Mesroghli, S., Jorjani, E. and Chelgani, S.C. (2009) Estimation of gross calorific value based on coal analysis using regression and artificial neural networks. International Journal of Coal Geology, 79(1-2), 49-54.
[25] Huang, P., Wang, X. and Han, S. (2019) Artificial Neural Network Prediction Model of Karst Water in Coal Mines. Fresen. Environ. Bull. 28(1), 452-458.
[26] Devijver, P.A. and Kittler, J. (1982). Pattern recognition: A statistical approach. Prentice Hall.
[27] Geisser, S. (1975) The predictive sample reuse method with applications. Journal of the American statistical Association. 70(350), 320-328.
[28] Isler, Y., Narin, A. and Ozer, M. (2015) Comparison of the effects of cross-validation methods on determining performances of classifiers used in diagnosing congestive heart failure. Measurement Science Review. 15(4), 196-201.
[29] Gümüşçü, A., Karadağ, K., Tenekecı, M.E. and Aydılek, İ.B. (2017) Genetic algorithm based feature selection on diagnosis of Parkinson disease via vocal analysis. 25th Signal Processing and Communications Applications Conference (SIU). 1-4.
[30] Zhang, G.P. (2001) An investigation of neural networks for linear time-series forecasting. Computers and Operations Research. 28(12), 1183-1202.

[31] Chang, T.C. and Chao, R.J. (2006) Application of back-propagation networks in debris flow prediction. Engineering Geology. 85(3-4), 270-280.
[32] Kayakuş M. and Terzioğlu, M. (2021) The Prediction of Pension Fund Net Asset Values Using Artificial Neural Networks and Multiple Linear Regression Methods. The Journal of Information Technologies. 14(1), 95-103.
[33] Khashei, M., Hamadani, A.Z. and Bijari, M. (2012) A novel hybrid classification model of artificial neural networks and multiple linear regression models. Expert Systems with Applications. 39(3), 2606-2620.
[34] Hancke, G.P. and Malan, R. (1998) A modal analysis technique for the on-line particle size measurement of pneumatically conveyed pulverized coal. IEEE Transactions on Instrumentation and Measurement. 47(1), 114-122.
[35] Cortes, C. and Vapnik, V. (1995) Support-vector networks. Machine learning, 20(3), 273-297.
[36] Vapnik, V. (2013) The nature of statistical learning theory. Springer Science and Business Media.
[37] Azamathulla, H.M., Ghani, A.A., Chang, C.K., Hasan, Z.A. and Zakaria, N.A. (2010) Machine learning approach to predict sediment load–a case study. CLEAN–Soil, Air, Water. 38(10), 969-976.
[38] Lu, C.J., Lee, T.S. and Chiu, C.C. (2009) Financial time series forecasting using independent component analysis and support vector regression. Decision Support Systems. 47(2), 115-125.
[39] Rao, B.V., Gopalakrishna, S. (2009) Hardgrove grindability index prediction using support vector regression. International Journal of Mineral Processing. 91(1-2), 55-59.
[40] Kayakuş, M., Yetiz, F. and Terzioğlu, M. (2021) The Impact of Internet Banking Usage on Bank Profitability: A Study on Deposit Banks Operating in Turkey. Turkish Studies-Economics-Finance-Politics. 16(1), 321-334.
[41] Lijuan, W. and Guohua C. (2016) Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow. Knowledge-Based Systems. 110, 157-166.
[42] Wu, C.H., Ho, J.M. and Lee, D.T. (2004) Travel-time prediction with support vector regression. IEEE Transactions on Intelligent Transportation Systems. 5(4), 276-281.
[43] Hsu, C.W. and Lin, C.J. (2002) A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks. 13(2), 415-425.
[44] Brown, S.H. (2009) Multiple linear regression analysis: a matrix approach with MATLAB. Alabama Journal of Mathematics. 34, 1-3.
[45] Amral, N., Ozveren, C. and King, D. (2007) Short term load forecasting using multiple linear regression. 42nd International Universities Power Engineering Conference. 1192-1198.

Received: **18.12.2020**
Accepted: **25.04.2021**

CORRESPONDING AUTHOR

Mehmet Kayakus
Management Information Systems
Faculty of Social Sciences and Humanities
Akdeniz University
Antalya 07600 – Turkey

e-mail: mehmetkayakus@akdeniz.edu.tr