ISSN: 0974-5823 Vol. 7 No. 1 January, 2022

International Journal of Mechanical Engineering

Performance comparison of regression learning methods: COVID-19 case prediction for turkey

Ipek Atik

Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, Gaziantep Islam Science and Technology University, Gaziantep, Turkey

Abstract - The World Health Organization has listed Covid-19 as a pandemic disease in the medical literature. The Covid-19 as a causative agent of respiratory diseases has spread rapidly and affected the whole world. Besides the rapid spread of the disease, alpha, delta, beta, gamma, and lastly the omicron variants of the virus have emerged. During the disease progression, countries experienced serious economic, educational and social problems. Everyday life was seriously restricted. Hundreds of people around the world are still losing their lives every day in the fight against the disease that threatens humanity. The study estimates the number of Covid 19 cases for Turkey using Deep Learning. For the analyses, the number of cases, deaths, vaccinated people, test positive_rates between February 12, 2021 and December 01, 2021 for Turkey was used. The input data are decomposed into their subcomponents in the study using the Empirical Mode Decomposition (EMD) signal decomposition method. In the prediction study, deep learning regression learner algorithms in MATLAB®2020a were used. Support Vector Machine (SVM), Linear regression (LR), Bagged Tree (BT), Fine Tree (FT) algorithms estimation results performance metrics were obtained. In the estimation studies, the best performance was obtained with the SVM algorithm. R values for SVM, LR, BT and FT were obtained R values as 99%, 94%, 90% and 88%; RMSE values as 0.017, 0.034, 0.052 and 0.109; MAE values as 0.014, 0.036, 0.040 and 0.087 respectively. The accurate assessment of the number of cases will help countries to take action.

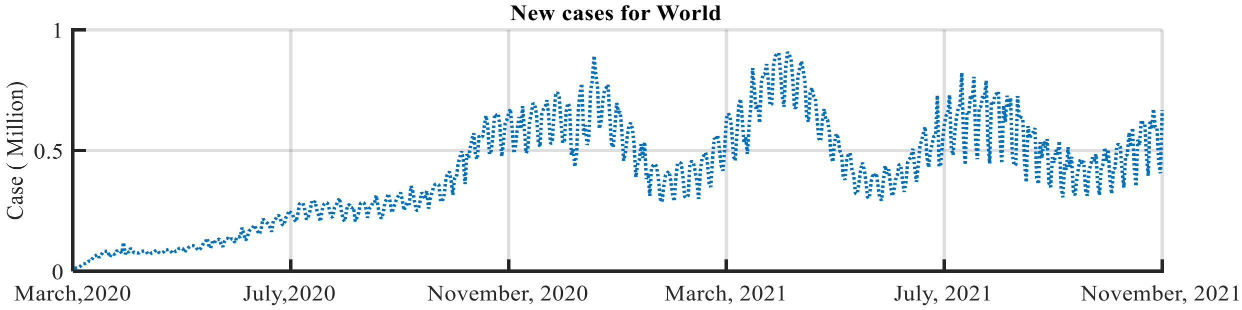
Index Terms - coronavirus, Covid-19, deep learning, forecasting, Turkey.

INTRODUCTION

Covid-19 is a deadly disease that emerged in Wuhan, China in 2019 and spread very quickly in a short time. It was entered into the literature as a pandemic disease by the World Health Organization in March 2020 [1].

The disease has deadly effects. For this reason, countries have taken important radical measures regarding social and economic life. Curfew, remote teaching, and quarantine are some of these measures. The disease has affected countries in terms of education, social and economic, especially health. To ensure the continuation of life, transportation, education, and social life were tried to be continued with certain precautions. Social spaces were closed by switching to the distance education system in education. Curfews have been implemented to reduce the contact areas of people [2].

The first case in Turkey was seen on March 11, 2020. However, as in the whole world, strict measures have been applied regarding the epidemic in Turkey as well [3]. Social distance rules, remote education, closure of social places, taking necessary measures in the health sector are some of them. Despite all the precautions, because the transmission factor of the disease is high, it has spread all over the country in a short time and has become difficult to control [4]. As in the whole world, there has been a decrease in the production sector and economic measures have also been implemented. The change in the number of new cases of covid-19 between 01 March 2020 and 30 November 2021 in the whole world and Turkey is as seen in Figure 1[5].



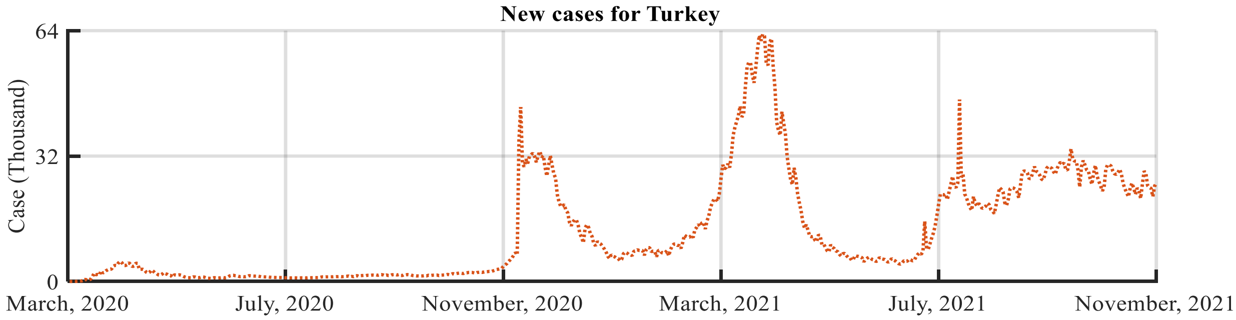


Figure 1. Change in the number of new cases of covid-19 in the world and Turkey between 01.03.2020-30.11.2021[5].

When the total number of covid-19 cases worldwide is examined, it is seen that there are fluctuations in the number of new cases between March 2020 and the date of the study, 30 November 2021. According to the official data announced, the number of new cases in the world was 2.378 on 01 March 2020; 95,705 on 01 June 2020; 266,043 on 01 September 2020; 566,625 in January 2021; 713657 in April 2021; It was identified as 440986 in July 2021 and 706863 in December 2021. In Turkey, on March 11, 2020; 827 on 01 June 2020; 1,572 on 01 September 2020; 12,203 in January 2021; 40,806 in April 2021; It was determined as 5,288 in July 2021 and 22,556 in December 2021[5].

The most important development in the fight against the disease has been the discovery of the vaccine. Vaccination has been the most effective method in the treatment of the disease. The first COVID-19 vaccine in the world was administered in December 2020. However, it is aimed to control the disease by increasing the number of people vaccinated all over the world [6]–[8].

In the study, the estimation of the number of covid-19 cases in Turkey was estimated by the deep learning regression analysis method. Data on the number of people vaccinated, the number of people who tested positive, the number of cases, and the number of deaths were used as input data. The data set subcomponents were obtained by decomposing the data set into its components using the EMD method. In the data set, the data between the date of the first vaccination in Turkey 12 February 2021, and the date of the study 01 December 2021 were taken. An accurate estimation of the number of cases will guide the measures to be taken. Especially the measures to be taken in the health sector are very important for countries. In addition, the relationship between the vaccine and the number of corona cases will be seen in the study [9]–[11].

RELATED WORKS

Since the covid-19 disease affected the whole world in a short time, many scientists have worked in this field. Studies on covid-19 include research in various aspects. Since our study is a machine learning prediction study, the studies in this field are summarized. Kırbaş et al. [4], ARIMA used the Nonlinear Autoregression Neural Network (NARNN) and Long-Short Term Memory (LSTM) method. He conducted a cumulative confirmed case and total increase rate estimation and comparison study of eight different European countries. In the study, it was estimated 14 days ahead.

Arora et al. [12], daily and weekly estimations were made with Deep LSTM/Stacked LSTM, Convolutional LSTM, and Bidirectional LSTM method. In the study, an increased positive rate estimation was made by using the data of Indian countries. Zeroual et al. [13], used RNN (Recurrent Neural Network), LSTM, Bi-LSTM (Bi-directional), VAE (Variational Auto Encoder) methods. Estimated the number of New COVID-19 cases and recovered cases 17 days ahead for six different countries. Shahid et al. [14], made predictions for 48 days ahead by using ARIMA, support vector regression (SVR), long short-term memory (LSTM), Bi-LSTM methods. The number of confirmed cases, deaths, and recovered cases was estimated in the study.

Chimmula and Zhang [7], estimated the number of confirmed cases for Canada and Italy using the LSTM method. It is estimated 14 days ahead in the forecast. Alzahrani et al. [15], ARIMA made a cumulative daily cases estimation study for 1 month ahead with the Autoregressive Moving Average (ARMA) method. The data of Saudi Arabia was used as the data set. Ogundokun et al. [16], performed a confirmed case estimation study for 8 days ahead with linear regression. Nigeria data was used as the data set.
Ribeiro et al. [17], Cumulative confirmed cases prediction study for 1,3 and 6 days ahead with ARIMA, cubist regression (CUBIST), random forest (RF), ridge regression (RIDGE), support vector regression (SVR) and stacking-ensemble learning he did. The data of the country of Brazil was used as the data set. Shastri et al. [18], made predictions for 30 days ahead by using LSTM, Stacked LSTM, Bi-directional LSTM, and Convolutional LSTM methods. The USA confirmed case data was used as the data set. Papastefanopoulos et al. [19], made predictions for 7 days ahead with ARIMA, the Holt-Winters additive model (HWAAS), TBAT, Facebook's Prophet, Deep AR models. He estimated the number of COVID-19 confirmed, death, and recovered cases.

MATERIAL AND METHODS

The estimation of the number of new cases of Covid-19 for Turkey was carried out with the regression analysis method, which is one of the deep learning methods. Among the regression analysis methods, Linear regression, Support Vector Machine (SVM), Bagged Tree, Fine Tree algorithms were used. In this part of the study, information about the data set is given. Regression methods are explained. Since the data set is decomposed by the EMD method, information is given about the EMD method. Performance metrics are given for comparison of model performance metrics.

A. Dataset Description

Turkey's daily new cases, number of deaths, number of people vaccinated, and positive test results were used as input data. Input data is separated into its components by the EMD method and decomposed into sub-data. To get a highly accurate estimate from the regression analysis, different regression parameters were tried to find the highest success rate. Data between 12.02.2021, the first vaccination date for Turkey, and 01.12.2021, the date of the study, were used in the study. The data set was randomly allocated as 80% training data and 20% test data. The data set time-series to change is as given in Figure 2 [5].

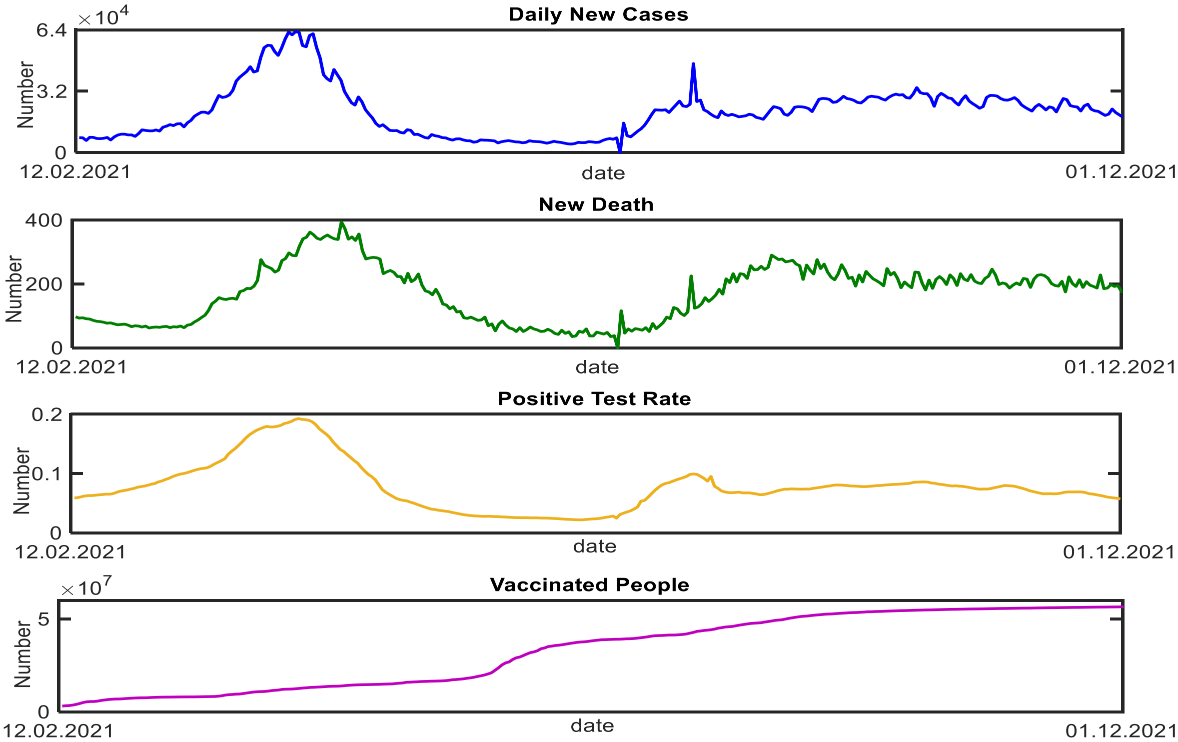


FIGURE 2. Turkey's coronavirus data [5].

B. Empirical Mode Decomposition (Emd)

The Empirical Mode Decomposition (EMD) method, which is called the flexible analysis method, is used for data belonging to stationary and nonlinear processes. The most important feature that distinguishes this method from the others is the eigenmode functions (IMF), each of which has different oscillations and is created by the local endpoints of each oscillation, which is symmetrical concerning their local mean. These oscillations are produced separately from the signal, assuming that a randomly determined signal consists of self-mode oscillations with different frequencies that are thought to be it's own [20].

To calculate the instantaneous frequency over the IMFs, the difference between the number called zero intercepts and the number of local endpoints must be a maximum of 1 and the local average value must be zero. To calculate this local mean value, a local time scale must be determined. In the EMD algorithm, the local minimum and maximum points of the signal are used for this purpose [21]. First, two signs are obtained by the cubic curve interpolation method of the sign and its local maximum and minimum points. The local average is obtained from the point average of the two obtained signals. This process repeats until the local mean value becomes zero and when the desired value is obtained, the current IMF value is accepted and subtracted from the original sign. If the sign is no longer monotonous, the IMF index i is increased sequentially and recorded as $c_1[n]$ by subtracting the first IMF from $[n]$ by elimination. The resulting IMF continues until its values become monotonous with the $[n]$ sign. When the monotony condition is met, the remaining $[n]$ is called a residual sign or trend. The final selection of IMFs is sorted from low to high. In the calculation given in Eq.(1), the IMF represents the eigenmode function, n represents the number of steps in the algorithm, and the r(t) value represents the residual signal [20].

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \qquad (1)$$

The five components and residuals obtained are given in Figure 3, with each component representing different properties of the data.

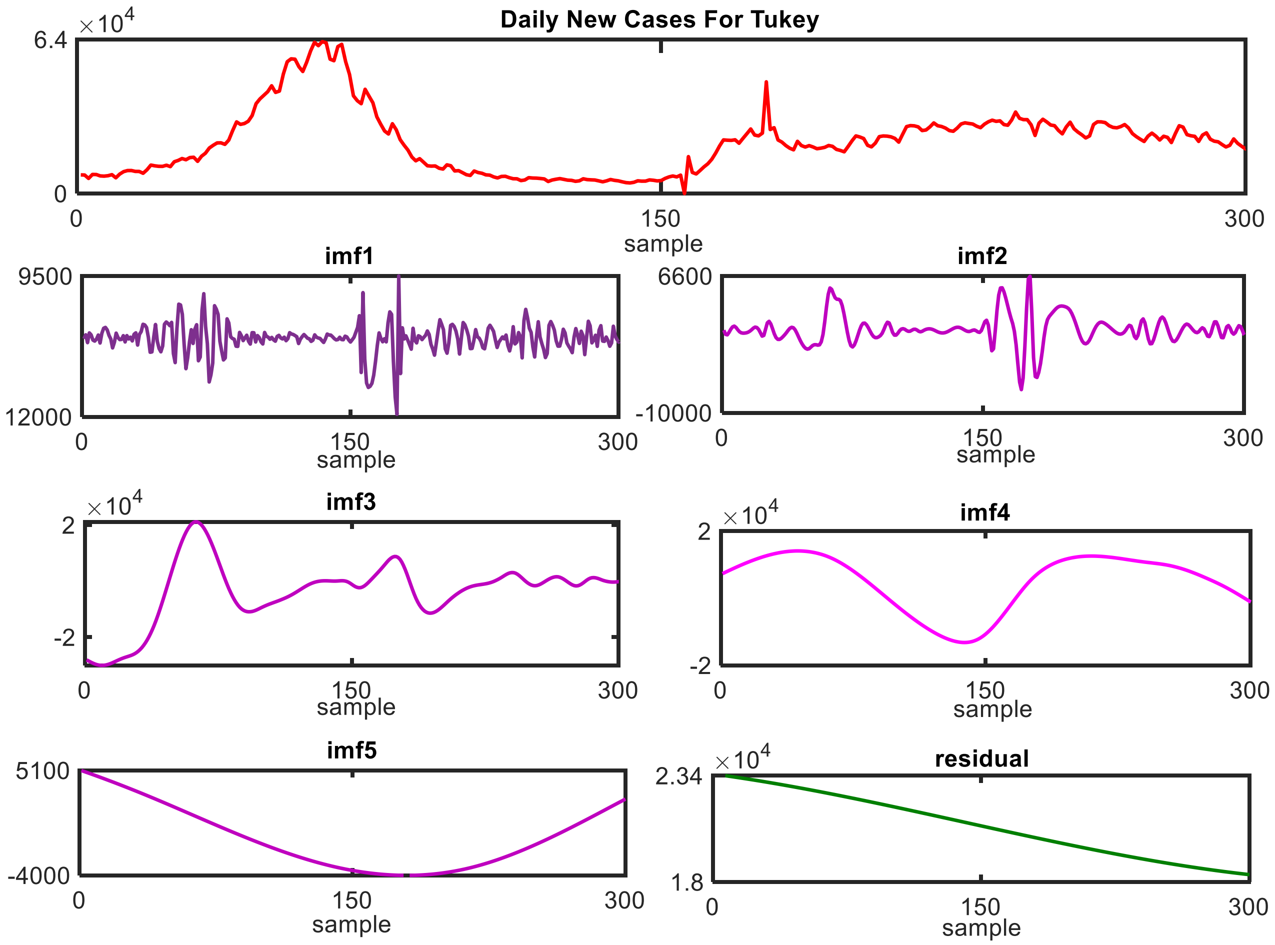


FIGURE 3. Components of Confirmed new cases in EMD.

C. Regression Analysis

1) LINEAR REGRESSION

The regression analysis method aims to estimate the approximate relationship between dependent and independent variables. If the number of variables is single, it is called univariate, and if there is more than one variable, it is called multivariate regression. In the method, one of the variables in the data set is dependent, while the others are independent variables that affect this variable at different rates. The simple linear regression method is expressed as in Eq.(2) [22];

$$Y_i = \theta_i + \theta_2 x_i + \varepsilon_i \qquad i = 1,2,\dots,n \qquad (2)$$

Here, i. dependent random variable x_i , observable i. independent variable (estimation variable), ε_i random error term, n represents the number of observations in the sample, i=1,2,….,n. θ_1 in the model; The point where the regression line intersects the y-axis is the slope of the θ_2 regression line, its representation is given in Figure 4 [22], [23].

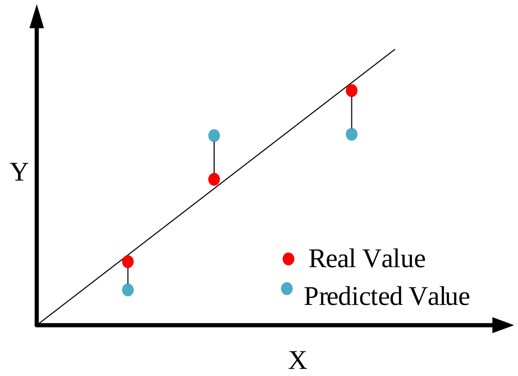


FIGURE 4. Linear regression for a data series.

2) DECISION SUPPORT MACHINE REGRESSION (SVR)

Support vector machines were first proposed by Vapnik. Support vector machines method is one of the learning statistical algorithms used for classification and regression analysis. It has also popularized the use of high performance in nonlinear problem solving with SVR. Although it was originally designed for regression analysis, it has also been used in regression analysis. The support vector machines method used in regression analysis was developed as Support Vector Regression (SVR). The estimation performance of the model varies according to the parameters forming the model. The SVR method aims to reach a function that approximates the training dataset to increase the prediction performance. For this reason, the three main parameters, Epsilon value (ε), capacitance value (C), and kernel function type, and if any, the parameter of the kernel function should be carefully determined. Commonly used kernel functions; linear, polynomial, radial basis function (RBF), and sigmoid. Consider a set with training points $\{(x_1, y_1), \ldots, (x_l, y_l)\}$, the future vector $x_i \in R^n$, and the target output $y_i \in$ R. Here, the relationship between nonlinear input and output is expressed as in Eq.(3) [24], [25].

$$f(x) = w^T \Phi(x) + b \qquad (3)$$

Here f(x) is the predicted values, Φ; are the nonlinear mapping function and w ($w \in R^n$) and b ($b \in R$) are the coefficients. C>0 and ε>0 for SVR;

$$Min_{w,b,\xi,\xi^*} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i + \xi_i^* \qquad (4)$$

constraints;

$$w^T \Phi(x_i) + b - y_i \leq \epsilon + \xi_i$$
$$y_i - w^T \Phi(x_i) - b \leq \varepsilon + \xi^*$$
$$\xi_i, \xi^* \geq 0; i = 1,2, \ldots, l \qquad (5)$$

Here ξ_i^* and ξ_i denote the training errors above and below ε, respectively. The other parameter vector w is as in Eq.(6);

$$w = \sum_{i=1}^{l} (\lambda_i^* - \lambda_i) \Phi(x_i) \qquad (6)$$

Here λ_i^* ve λ_i are Lagrange multipliers. The SVR equation is obtained as in Eq.(7);

$$f(w) = \sum_{i=1}^{l} (\lambda_i^* - \lambda_i) K(x_i, x_j) + b \qquad (7)$$

Here, the equation $K(x_i, x_j) = \exp\left(-y\|x_i - x_j\|^2\right)$ represents the radial basis kernel function [24], [26].

3) BAGGED TREE

In the bagged tree method, estimators are applied to the bootstrap samples from the original dataset and an ensemble is created. The purpose of the pre-loading application is to generate the subsamples by making a random selection with a return. The number of subsamples is the same as the original data set. While some observations are not included in the samples, some may take place several times. While merging estimates is averaged in regression trees, results in classification trees are determined by voting. The bagged tree method gives more successful results than single trees. The bagged tree flowchart is as given in Figure 5 [25], [27];

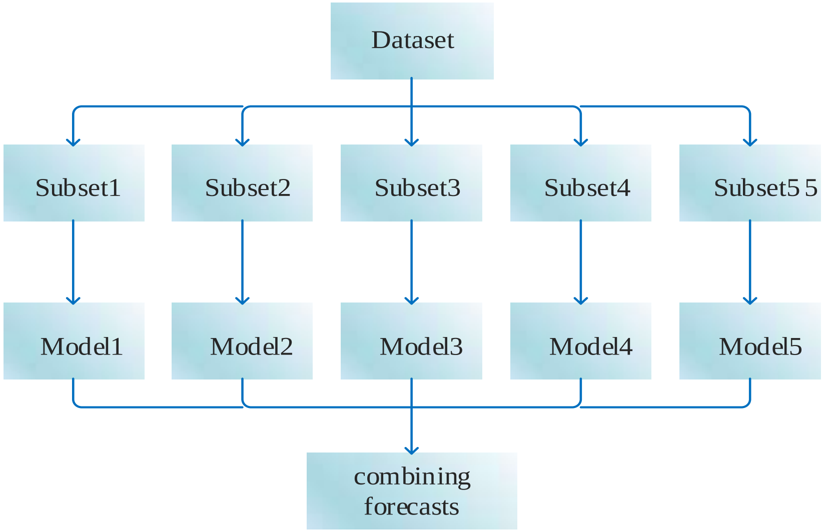


FIGURE 5. Bagged tree structure

In the analysis with the bagged tree, multiple subsets of the data set are created in the first step. A base model is created in each subgroup. The models work independently of each other and finally, the prediction results of all models are combined to determine the final prediction [28], [29].

4) DECISION TREE REGRESSION

The decision tree regression structure consists of nodes and leaves. The dataset is divided into subsets, and the decision nodes are divided into two or more sub-branches. There are decision and leaf nodes in the decision tree. While performing the node calculations, the standard deviation is used as given in Eq.(8) while performing the regression process. First, the standard deviation of the target cluster is calculated. Then, as given in Equation 8, binary standard deviation values are calculated between other clusters and target clusters. As given in Eq.(9), each result is subtracted from the standard deviation value of the target cluster. The set with the smallest standard deviation is the root set. By repeating these operations for each node, the tree structure is formed, expressed as given in Eq. (10) [30];

$$S = \sqrt{\frac{\sum(x-\mu)^2}{n}} \tag{8}$$

$$S(T,X) = \sum c\epsilon x\, P(c)\, S(c) \tag{9}$$

$$SDR(T,X) = S(T) - S(T,X) \tag{10}$$

Here, μ is the mean of the feature, n is the number of data in the feature, and c is the different values that the feature can take [30].

D. Performance Comparison Metrics

The expressions of the correlation coefficient (R), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) metrics used in the performance comparisons of the methods are given in Eq.(11)-(13) [24].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - X_i)^2} \tag{11}$$

$$R = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y}_i)(Y_i - \bar{X}_i)}{\sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y}_i)^2 \sum_{i=1}^{N}(Y_i - \bar{X}_i)^2}} \tag{12}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Y_i - X_i| \tag{13}$$

Here, the input value, the predictive value, the mean of the values determined by the analysis, and the number of samples in the test subset are represented by Y, Y', $\bar{Y}$ and n, respectively. In the analysis, 10-fold cross-validation was applied to increase system stability. In cross-validation, the original dataset is divided into k subsets, a single subset from each subset is used as the test set and the k-1 subset is used as the training set. When this process is repeated k times, all data is used in both the training and test set.

E. Experimental Studies And Results

In the study, the number of new cases of covid-19 in Turkey was estimated. For this, the data of Turkey's number of new cases, number of deaths, number of positive tests, and number of people vaccinated between 12.02.2021 and 01.12.2021 were used. In the data set, the data on the number of new cases were divided into components using the EMD method and used as input data. 80% of the data set was used as training data and 20% as test data.

Linear regression, decision support vector regression, bagged tree and decision tree methods were used in the estimation study. All experiments were performed with Intel in MATLAB R 2020 (a) environment (R) with i7-10750 H CPU @2.60 GHz, NVIDIA Quadro P620 GPU, and 16 GB RAM. The parameters, comparison metrics, and regression analysis results of the regression models used in the latter part of this section are given. The working flow diagram is as given in Figure 6 and model parameters are summarizen in Table 1. As a result of the analyzes made, the exact number of new cases of covid-19 in Turkey and the model results are as shown in Figure 7 and model performances are summarized in Table 2.

TABLE I.
PARAMETERS OF THE METHODS USED IN MODEL ANALYSIS

Algorithm	Parameters
Support Decision Regression (SVR)	Kernel Fonkisyonu: Gaussian
	Kernel scale:0.62
Linear Regression (LR)	Term: Interaction
	Robust option: Off
Bagged Tree Regression (BG)	Minimum Leaf size: 16
	Number of learner:32
Decision Tree Regression_ Fine Tree (FT)	Minimum Leaf size: 12
	Decision Split: Off

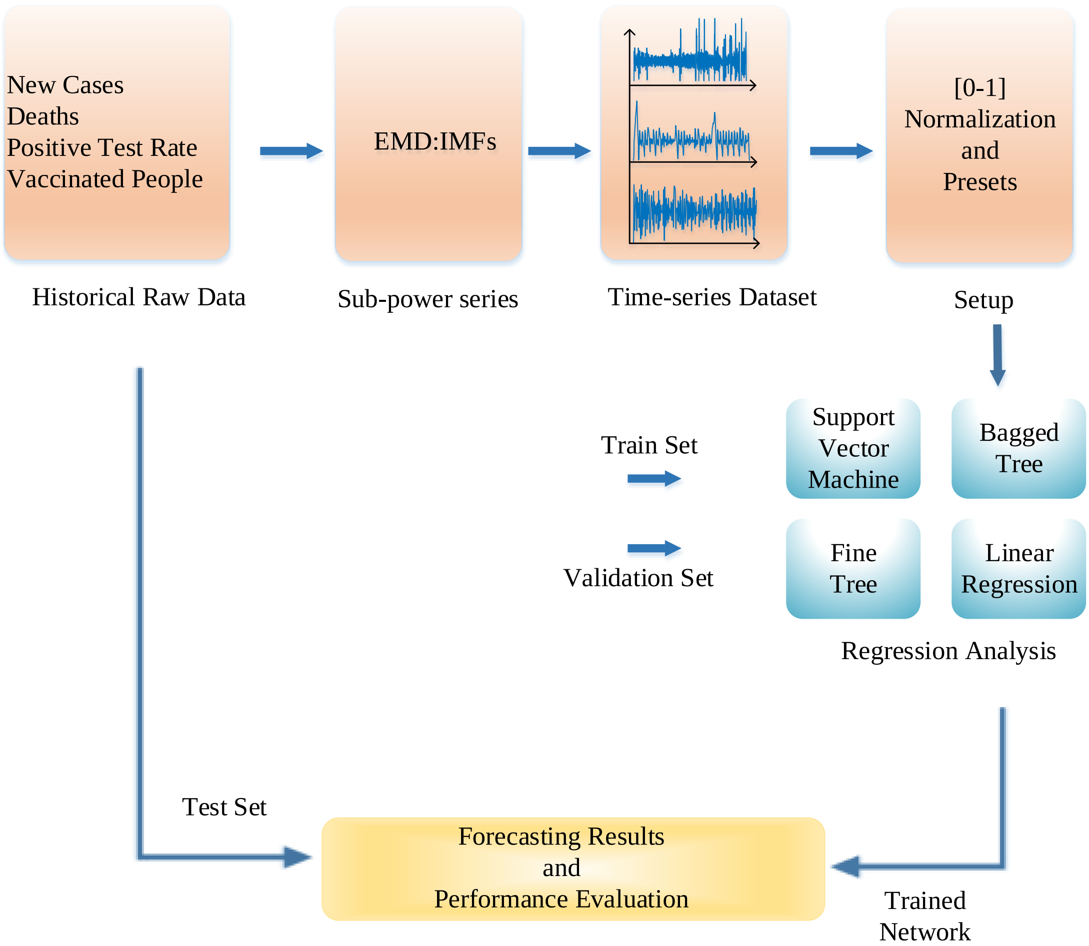


FIGURE 6. Operation flow diagram

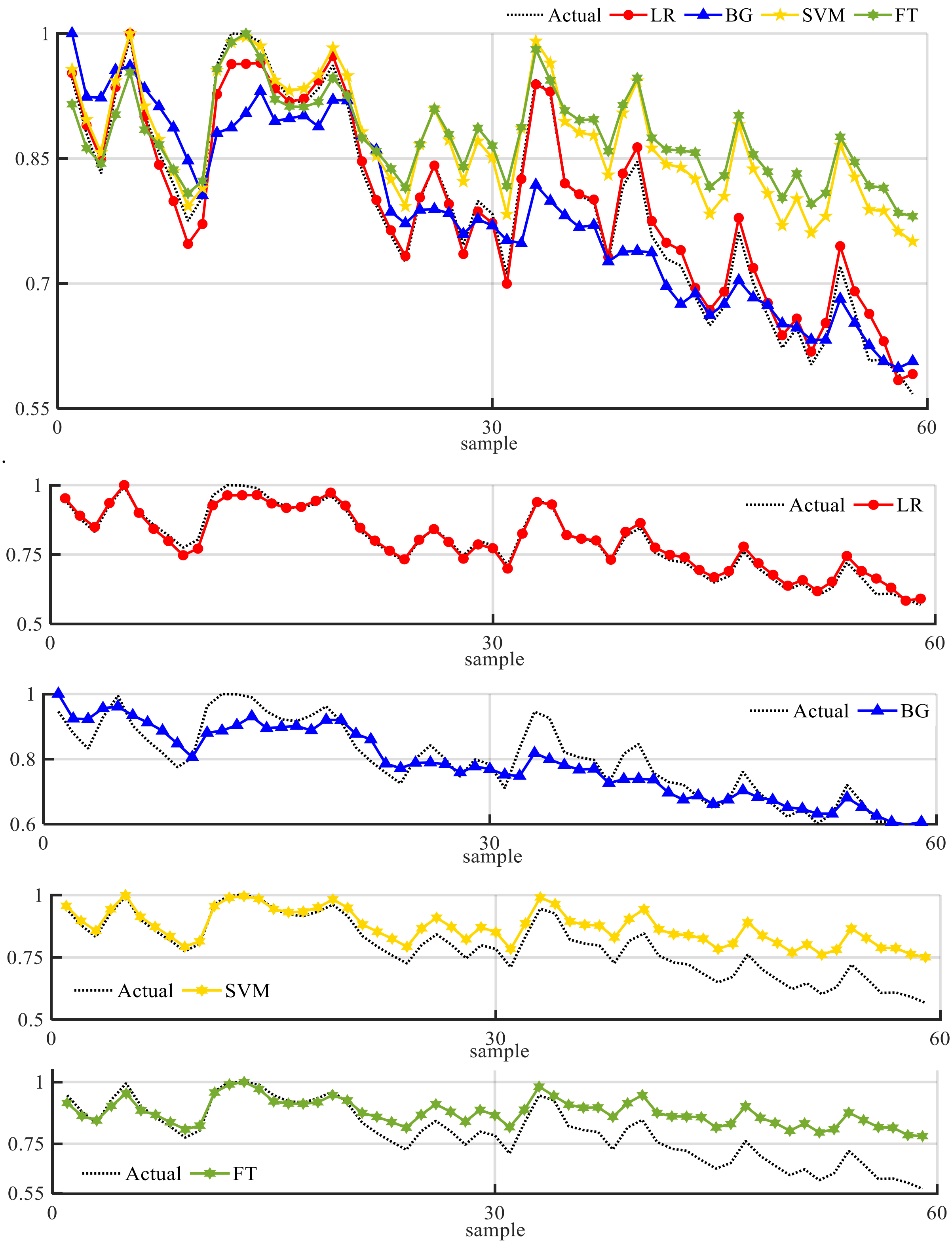


FIGURE 7. Prediction results with SVR, LR, BG, and FT models.

TABLE II

MODEL PERFORMANCE METRICS

Metrics	LR	SVR	BG	FT
R	0.991	0.944	0.906	0.881
RMSE	0.017	0.034	0.052	0.109
MAE	0.014	0.036	0.040	0.087

When the performance values of the prediction models are compared, the highest R-value was obtained in the LR model with 0.991, while the lowest R-value belonged to FT with 0.881. As can be seen in Figure 7, while the values in the LR analysis are close to the real values, the error is higher in FT. When model performances were evaluated, RMSE values of LR, SVR, BG, and FT models were obtained as 0.017, 0.034, 0.052, and 0.109, respectively. When MAE values are compared, the lowest MAE value belongs to LR, while the model with the highest MAE value is FT with 0.087.

CONCLUSIONS

In this study, the number of new cases of covid-19 for Turkey was estimated by the machine learning regression method. In the analysis, daily data of new cases, deaths, positive tests, and vaccinated people between 12.02.2021 and 01.12.2021 were taken. According to the results obtained in the estimation studies, the LR model showed the highest performance and the FT method showed the worst performance. R, RMSE, and MAE values of the LR method are respectively; 0.991, 0.017, and 0.014 were obtained. Performance values of the SVM method; 0.944, 0.034 and 0.036. According to these values, the models with the best performance were obtained as LR, SVR, BG, and FT, respectively. Accurate estimation of new cases of Covid-19 is very important in terms of health, economic and social measures taken by countries. Thus, the planning will be done more accurately and the fight against the disease will be taken under control.

REFERENCES

[1] H. Özbay and A. Dalcalı, "Effects of COVID-19 on electric energy consumption in Turkey and ANN-based short-term forecasting," Turk. J. Electr. Eng. Comput. Sci., vol. 29, pp. 78–97, Jan. 2021.

[2] K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. M. Hyman, P. Yan, and G. Chowell, "Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020," Infect. Dis. Model., vol. 5, pp. 256–263, 2020.

[3] J. Stübinger and L. Schneider, "Epidemiology of Coronavirus COVID-19: Forecasting the Future Incidence in Different Countries," Healthcare, vol. 8, no. 2, 2020.

[4] İ. Kırbaş, A. Sözen, A. D. Tuncer, and F. Ş. Kazancıoğlu, "Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches," Chaos Solitons Fractals, vol. 138, p. 110015, 2020.

[5] "Johns Hopkins Coronavirus Resource Center," Coronavirus Research Center. [Online]. Available: https://coronavirus.jhu.edu/map.html. [Accessed: 11-Dec-2021].

[6] N. Ghazaly, M. Abdel-Fattah, and A. E.-A. Ahmed, "Novel Coronavirus Forecasting Model using Nonlinear Autoregressive Artificial Neural Network," J. Adv. Sci., vol. 29, pp. 1831–1849, Apr. 2020.

[7] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," Chaos Solitons Fractals, vol. 135, p. 109864, Jun. 2020.

[8] F. Demir, "DeepCoroNet: A deep LSTM approach for automated detection of COVID-19 cases from chest X-ray images," Appl. Soft Comput., vol. 103, p. 107160, 2021.

[9] D. A. Pustokhin, I. V. Pustokhina, P. N. Dinh, S. V. Phan, G. N. Nguyen, G. P. Joshi, and S. K., "An effective deep residual network based class attention layer with bidirectional LSTM for diagnosis and classification of COVID-19," J. Appl. Stat., pp. 1–18, Nov. 2020.

[10] Y. Eroğlu, "FORECASTING MODELS FOR COVID-19 CASES OF TURKEY USING ARTIFICIAL NEURAL NETWORKS AND DEEP LEARNING," End. Mühendisliği, vol. 31, no. 3, pp. 353–372, Dec. 2020.

[11] M. Hawas, "Generated time-series prediction data of COVID-19′s daily infections in Brazil by using recurrent neural networks," Data Brief, vol. 32, p. 106175, Oct. 2020.

[12] P. Arora, H. Kumar, and B. K. Panigrahi, "Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India," Chaos Solitons Fractals, vol. 139, p. 110017, Oct. 2020.

[13] A. Zeroual, F. Harrou, A. Dairi, and Y. Sun, "Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study," Chaos Solitons Fractals, vol. 140, p. 110121, Nov. 2020.

[14] F. Shahid, A. Zameer, and M. Muneeb, "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM," Chaos Solitons Fractals, vol. 140, p. 110212, 2020.

[15] S. I. Alzahrani, I. A. Aljamaan, and E. A. Al-Fakih, “Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions,” J. Infect. Public Health, vol. 13, no. 7, pp. 914–919, Jul. 2020.

[16] R. O. Ogundokun, A. F. Lukman, G. B. M. Kibria, J. B. Awotunde, and B. B. Aladeitan, “Predictive modelling of COVID-19 confirmed cases in Nigeria,” Infect. Dis. Model., vol. 5, pp. 543–548, Jan. 2020.

[17] M. H. D. M. Ribeiro, R. G. da Silva, V. C. Mariani, and L. dos S. Coelho, “Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil,” Chaos Solitons Fractals, vol. 135, p. 109853, Jun. 2020.

[18] S. Shastri, K. Singh, S. Kumar, P. Kour, and V. Mansotra, “Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study,” Chaos Solitons Fractals, vol. 140, p. 110227, Nov. 2020.

[19] V. Papastefanopoulos, P. Linardatos, and S. Kotsiantis, “COVID-19: A Comparison of Time Series Methods to Forecast Percentage of Active Cases per Population,” Appl. Sci., vol. 10, no. 11, 2020.

[20] A. Altan and S. Karasu, “Ayrıştırma Yöntemlerinin Derin Öğrenme Algoritması ile Tanımlanan Rüzgâr Hızı Tahmin Modeli Başarımına Etkisinin İncelenmesi,” Avrupa Bilim Ve Teknol. Derg., no. 20, pp. 844–853, Dec. 2020.

[21] H. Zheng, J. Yuan, and L. Chen, “Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation,” Energies, vol. 10, no. 8, 2017.

[22] X. Yan and S. Xiaogang, Linear regression analysis: theory and computing. World Scientific, 2009.

[23] B. Bianco, O. Manca, and S. Nardini, “Electricity consumption forecasting in Italy using linear regression models,” Energy, vol. 34, pp. 1413–21, 2009.

[24] V. O. Akande,, T. O. Owolabi,, S. Twaha,, and S. O. Olatunji,, “Performance Comparison of SVM and ANN in Predicting Compressive Strength of Concrete,” IOSR J. Comput. Eng., no. 16, 2014.

[25] M. Božić and M. Stojanović, “Application of SVM Methods for Mid-Term Load Forecasting,” Serbian J. Electr. Eng., vol. 8, no. 1, pp. 73–83, 2011.

[26] A. Ghanbari, A. Naghavi, S. F. Ghaderi, and M. Sabaghian, “Artificial Neural Networks and Regression Approaches Comparison for Forecasting Iran’s Annual Electricity Load,” presented at the International Conference on Power Engineering, Energy and Electrical Drives, 2009, pp. 675–679.

[27] K. K. Al-Barazanchi, A. Q. Al-Neami, and A. H. Al-Timemy, “Ensemble of bagged tree classifier for the diagnosis of neuromuscular disorders,” presented at the 2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME),, 2014, pp. 1–4.

[28] T. G. Dietterich, “An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization.,” vol. 40, pp. 139–157, 2000.

[29] T. K. Ho, “Random decision forests,” presented at the In: Proceedings of 3rd İnternational Conference on Document Analysis and Recognition. IEEE, Montreal, QC, Canada, pp. 278–282.