Çılgın, C., & Özdemir, M.O. (2023). Time Series Forecasting of Covid-19 Confirmed Cases in Turkey with Stacking Ensemble Models. *Bingöl University Journal of Social Sciences Institute, 26*, 504-520. https://doi.org/10.29029/busbed.1299248

Article Type : Research Article
Date Received : 18.05.2023
Date Accepted : 24.10.2023

https://doi.org/10.29029/busbed.1299248

TIME SERIES FORECASTING OF COVID-19 CONFIRMED CASES IN TURKEY WITH STACKING ENSEMBLE MODELS

Cihan ÇILGIN[1], Mehmet Ozan ÖZDEMİR[2]

ABSTRACT

Since COVID-19 has spread almost across any country and is a serious threat to mankind, it was declared to be a pandemic by WHO. Forecasting the results of a pandemic is a quite important and difficult task for policy makers and decision makers. The aim of this study is to forecast the daily case numbers in Turkey by using various time series modeling approaches. In this context, positive case numbers between March 11, 2020, and December 24, 2021, were taken into account in this study. This study, with the number of observations it covers, differentiates from other studies which have been conducted with few number of observations. In this study, all the waves during the COVID 19 pandemic were included in the analysis by studying a more extensive time period. Moreover, in our study, along with a comparison of machine learning algorithms by making case forecasting with these algorithms, increasing the forecasting performance was aimed by combining the predictions of all models used with the stacking approach under a single model. By taking all the related studies analyzed into account, our study, as far as we know, is the first one to assess this many model performances together and make a stacking model on COVID-19 case numbers. The findings obtained from the study prove that forecasting of the cases validated via the developed stacking model were made with high accuracy, and all ensemble learning approaches produce better results than individual methods.

Keywords: COVID-19, time-series, forecasting, machine learning, coronavirus, generalized stacking models

[1] *Dr. Res. Asst., Bolu Abant İzzet Baysal University, Faculty of Applied Sciences, cihancilgin@ibu.edu.tr, https://orcid.org/0000-0002-8983-118X*

[2] *Ass. Prof. Dr., Dokuz Eylül University, Faculty of Economics and Administrative Sciences, ozan.ozdemir@deu.edu.tr, https://orcid.org/0000-0002-4224-1190*

Çılgın, C., & Özdemir, M.O. (2023). Türkiye'de Doğrulanmış Covid-19 Vakalarının İstifleme Topluluk Modeller İle Zaman Serisi Tahmini. *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 26*, 504-520. https://doi.org/10.29029/busbed.1299248

Makalenin Türü : Araştırma Makalesi
Geliş Tarihi : 18.05.2023
Kabul Tarihi : 24.10.2023

https://doi.org/10.29029/busbed.1299248

TÜRKİYE'DE DOĞRULANMIŞ COVİD-19 VAKALARININ İSTİFLEME TOPLULUK MODELLER İLE ZAMAN SERİSİ TAHMİNİ

Cihan ÇILGIN[1], Mehmet Ozan ÖZDEMİR[2]

ÖZ

COVID-19 hemen hemen her ülkeye yayıldığı ve insanlık için ciddi bir tehdit oluşturduğu için DSÖ tarafından pandemi olarak ilan edilmiştir. Bir pandeminin sonuçlarını tahmin etmek, politika yapıcılar ve karar vericiler için oldukça önemli ve zor bir görevdir. Bu çalışmanın amacı, çeşitli zaman serisi modelleme yaklaşımlarını kullanarak Türkiye'deki günlük vaka sayılarını tahmin etmektir. Bu kapsamda 11 Mart 2020 ile 24 Aralık 2021 tarihleri arasındaki pozitif vaka sayıları bu çalışmada dikkate alınmıştır. Bu çalışma, kapsadığı gözlem sayısı ile daha az gözlem sayısı ile yapılmış diğer çalışmalardan ayrılmaktadır. Bu çalışmada COVID 19 pandemisi sırasındaki tüm dalgalar daha geniş bir zaman diliminde incelenerek analize dâhil edilmiştir. Ayrıca çalışmamızda makine öğrenmesi algoritmalarının bu algoritmalar ile durum tahmini yapılarak karşılaştırılması yanında, yığınlama yaklaşımı ile kullanılan tüm modellerin tahminleri tek bir model altında birleştirilerek tahmin performansının artırılması hedeflenmiştir. Çalışmamız, incelenen ilgili tüm çalışmaları dikkate alarak, bildiğimiz kadarıyla, bu kadar çok model performansını bir arada değerlendiren ve COVID-19 vaka sayıları üzerinde bir yığınlama modeli oluşturan ilk çalışmadır. Çalışmadan elde edilen bulgular, geliştirilen istifleme modeli ile doğrulanan durum tahminlerinin yüksek doğrulukta yapıldığını ve tüm toplu öğrenme yaklaşımlarının bireysel yöntemlere göre daha iyi sonuçlar verdiğini kanıtlamaktadır.

Anahtar Kelimeler: COVID-19, zaman serisi, tahmin, makine öğrenimi, koronavirüs, genelleştirilmiş yığınlama modelleri

[1] *Dr. Araş. Gör., Bolu Abant İzzet Baysal Üniversitesi, Uygulamalı Bilimler Fakültesi, cihancilgin@ibu.edu.tr, https://orcid.org/0000-0002-8983-118X*

[2] *Dr. Öğr. Üyesi, Dokuz Eylül Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, ozan.ozdemir@deu.edu.tr, https://orcid.org/0000-0002-4224-1190*

1. INTRODUCTION

Although it has been almost 2 years since it has started, the COVID-19 pandemic, which has affected the whole world and caused permanent damages both on economic and social areas, is still acknowledged as a pandemic by the World Health Organization (WHO) and maintains its course with multifarious variants. The COVID-19 pandemic affecting more than 200 countries, infecting more than 330 million people and causing more than 5,5 million deaths since its beginning had quite heavy effects also on Turkey and brought along many problems in many areas. Modeling the extensity and modeling, predicting and forecasting the epidemiological properties of the virus are important topics in providing the necessary equipment to deal with possible outcomes (Maleki et al., 2020). Notwithstanding, it was hard to forecast about the COVID-19 pandemic in its early stages, in todays's data and information age the ever-increasing data on the spread of the pandemic makes it easier to forecast the course of the status for the researchers. In the struggle against this pandemic, which has taken a hold of the whole world, threatened public health as much as the health of individuals and tested the health systems and economies, it is vital to take early precautions and planning a course of action in this context. When patient capacity, test reserve and the presence of protective equipment particularly during the pandemic period are taken into consideration, forecasting the future with current cases is critical for logistics, planning the hospital personnel and equipment and even vaccination process (Özen et al., 2021). When taken into consideration precautionary actions like full lockdown, which have immense negative outcomes, preventing COVID-19 from spreading and being able to forecast its spread is important for managing this process with less adversary precautions. Mathematical model applications, artificial intelligence, methodologies like big data and so on carries great potential in the prediction of the amount of possible necessary additional equipment and resources (Ceylan, 2020), along with the extents of the spread and efficiency of restricting strategies to prevent the spread of this epidemic disease (Abdulmajeed et al., 2020). Time series models may take an important part in predicting diseases, and particularly in pandemics like COVID-19, and incidence data can be used to forecast the future course of the disease (Kane et al., 2014).

The effects of the pandemic have been felt also in Turkey parallel to other countries, and maybe even more destructive, and many courses of action were planned and applied to hold the pandemic under control. In this context, this study aims to create time series prediction models with machine learning algorithms and econometric models by taking daily case numbers in Turkey, where more than 12 million people have been infected in total, and more than 90 thousand people have died since the first case was recorded, into consideration. Moreover, new COVID-19 variants are emerging day by day and the variable effects of each new variant on the number of cases increases the need for new models with different observation numbers and developed in different time periods. In this study, as well as comparing prediction methods by taking these models into account that were developed on time series in the context of already existing related studies (Khan et al., 2021), enhancing the performance of prediction by using the stacking approach, which is an ensemble learning approach with all the models developed is aimed. Although many factors such as the interventions of country governments to cope with the pandemic, the lack of epidemiological information (Pontoh et al., 2020), epidemiological facts, randomness, stay-at-home compliance and curfews (Abdulmajeed et al., 2020) make it difficult to develop prediction models on COVID-19 data, this study focuses on the performance of prediction models even in these challenging situations. More importantly, the results obtained within the scope of this study are thought to provide an academic impetus to guide policy makers in terms of developing real-time health policies by governments and health institutions, along with making basic public information about future conditions available to individuals.

In the continuing sections of the study, studies on COVID-19 data in different countries and Turkey were analyzed and presented in a systematic way in the second part. While the third section contains the explanations of the data set, models and the proposed method used in the study, the application steps carried out within the framework of the proposed method are explained in the following section. In the last part, the study is concluded with the findings obtained and the conclusion part formed in the light of these findings.

2. RELATED WORK AND BACKGROUND

Like the bird flu H5N1 (Kane et al., 2014; Biswas et al., 2014), SARS (Earnest et al., 2005, Lai, 2005), hepatitis (Guan et al., 2004; Sumi et al., 2013) in the past, in many diseases and epidemics, time series models have been used by researchers to forecast. Considering the effects of the pandemic and the fact that Covid-19 is the biggest pandemic in the last century, it has triggered many researchers to be productive on this issue. Many researchers have conducted mathematical models to predict how many people are infected in order to provide accurate information to their governments (Pontoh et al., 2020) and to develop real-time health policies by health institutions with the right decisions and action plans in the light of this accurate information (Koçak, 2020). In particular, the pandemic's close impact on many vital issues such as the economy, health system, security and food access, especially public health, and the fact that the course of the pandemic is the focus of all humanity has led to an increase in both the scope and the number of researches on the Covid-19 pandemic. Within the scope of this study, many studies have been carried out in order to obtain predictions about the course of the pandemic through various time series models. In accordance with this purpose Turkey (Karcıoğlu et al., 2021; Akay & Akay, 2021; Fidan & Yuksel, 2022), Greece (Katris, 2021), Portugal (De Oliveira, et al., 2021), Italia (Ding et al., 2020; Ceylan, 2020; Dehesh et al., 2020), Worldwide (Sevli & Gülsoy; 2020, Maleki et al., 2020, Petropoulos et al., 2022), USA (De Oliveira, et al., 2021; Özen et al., 2021; Shastri et al., 2020; Zeroual et al., 2020), Malaysia (Purwandari et al., 2022), Spain (Ceylan, 2020; Zeroual et al., 2020), Brazil (De Oliveira, et al., 2021; Dairi, 2021), İndia (Shastri et al., 2020; Arora et al., 2020; Tandon et al., 2020), Nigeria (Abdulmajeed et al., 2020), France (Zeroual et al., 2020; Ceylan, 2020; Dairi, 2021), Canada (Chimmula & Zhang, 2020), South Korea (Pontoh et al., 2020; Dehesh et al., 2020), China (Dehesh et al., 2020; Zeroual et al., 2020), Iran (Dehesh et al., 2020; Talkhi et al., 2021), Austuralia (Zeroual et al., 2020), Russia (Dairi, 2021), Pakistan (Ali et al., 2020) etc. many studies with COVID-19 time series were conducted for many countries. While the scope of these studies mostly consisted of the daily number of positive cases (Kumar & Susan, 2020; Papastefanopoulos et al., 2020; Zeroual et al., 2020; Chimmula & Zhang, 2020; Arora et al., 2020; Ding et al., 2020; Karcıoğlu et al., 2021; Dairi, 2021; Sevli & Gülsoy, 2020; Özen et al., 2021; Tandon et al., 2020; Purwandari et al., 2022; Pontoh et al., 2020; De Oliveira, et al., 2021; Shastri et al., 2020; Talkhi et al., 2021, Petropoulos et al., 2022), which is also the subject of this study, the authors also took into account the daily number of deaths (Karcıoğlu et al., 2021; Sevli & Gülsoy, 2020; Purwandari et al., 2022; Petropoulos et al., 2022, Pontoh et al., 2020; Shastri et al., 2020; Talkhi et al., 2021) and the daily recoveries (Karcıoğlu et al., 2021; Sevli & Gülsoy, 2020; Purwandari et al., 2022; Purwandari et al., 2022; Pontoh et al., 2020; Zeroual et al., 2020; Dairi, 2021). In addition, many statistical and machine learning based time series models such as ARIMA (Ali et al., 2020; Tandon et al., 2020; Karcıoğlu et al., 2021; Ding et al., 2020; Özen et al., 2021; Abdulmajeed et al., 2020; Ceylan, 2020; Dehesh et al., 2020; Talkhi et al., 2021; Katris, 2021), Prophet (Sevli & Gülsoy, 2020; Özen et al., 2021; Abdulmajeed et al., 2020; Talkhi et al., 2021), Holt-Winter (Abdulmajeed et al., 2020; Talkhi et al., 2021), Random Forest (Özen et al., 2021), Linear Regression (Özen et al., 2021), Artificial Neural Networks (Pontoh et al., 2020; Purwandari et al., 2022; De Oliveira, et al., 2021; Katris, 2021), Extreme Learning Machine (Pontoh et al., 2020; Talkhi et al., 2021; Purwandari et al., 2022), Support Vector Machine (Dairi, 2021), LSTM (Dairi, 2021; Chimmula and Zhang, 2020; Karcıoğlu et al., 2021) and RNN (Shastri et al., 2020; Arora et al., 2020; Zeroual et al., 2020) were used to estimate the daily positive cases, the number of deaths and the number of recoveries.

Existing studies in the literature have generally focused on the development and comparative analysis of various methods on one or more time series. Due to the data-based learning and stochastic nature of machine learning approaches, it is inevitable that different results will be obtained on different data sets, and different model advantages are observed in studies on COVID-19 data. Özen et al. (2021), in parallel with the study on Nigeria by Abdulmajeed et al. (2020), conducted prediction studies on the number of cases in the United States with various methods such as Prophet, Random Forest, and ARIMA, and they revealed that the Polynomial Regression method among the existing methods is more reliable than other methods. Ceylan (2020), unlike other authors, conducted research on various model orders with the Autoregressive Integrated Moving Average (ARIMA), whose effectiveness has been tested in various previous epidemics and diseases (Guan et al., 2004; Earnest et al., 2005), to predict the epidemiological trend of the prevalence of COVID-19 in Italy, Spain and France, the most affected countries in Europe and reported the effectiveness of the Autoregressive Integrated Moving Average on COVID-19 data. A similar study was carried out with ARIMA by Ali et al. (2020) in the Pakistan sample, and it was emphasized by the authors that ARIMA (1, 0, 4) is the best model configuration on the available data. Shastri et al. (2020) designed a methodology with variants of deep learning models, such as long short term memory networks (LSTM) and recurrent neural network (RNN) based, using confirmed cases and death cases for both the United States and India. They concluded that for all four datasets of both countries, the Convolutional LSTM predicted Covid-19 cases with higher accuracy and much less error than other model variations. In a similar approach, Arora et al. shared their findings with long short-term memory models on the next day and one week prediction of COVID-19 cases, reporting the COVID-19 prediction for India's 32 states and union territories, with an error of 3%. Another study using long short term memory models was carried out on case data approved by Chimmula and Zhang (2020) in Canada. In the study, they emphasized that they obtained satisfactory results with LSTM, which supports the studies carried out with similar methods. In addition to these studies, Maleki et al. (2020) used econometric models that require more model specification and an improved autoregressive time series model based on two-part scaled mixture normal (TP-SMN) distributions. In the study, they developed a new efficient prediction model to forecast confirmed and recovered COVID-19 cases in the world using past and current data. The results reveal that the proposed method performs well in worldwide forecasting confirmed and recovered COVID-19 cases.

Ahmar and del Val (2020) used the SutteARIMA and ARIMA methods to forecast short-term confirmed cases of COVID-19 in Spain. The data on Covid-19 was obtained from Worldometer. To evaluate the forecasting methods, mean absolute percentage error (MAPE) was used as an evaluation metric. Based on the findings obtained from the ARIMA and SutteARIMA methods, they concluded that the SutteARIMA method is superior to the ARIMA method to forecast the daily confirmed cases of COVID-19. Riberio et al. (2020) applied autoregressive integrated moving average (ARIMA), cubist regression (CUBIST), random forest (RF), ridge regression (RIDGE), support vector regression (SVR), and stacking-ensemble learning methods to forecast the COVID-19 cumulative confirmed cases in ten Brazilian states. The CUBIST regression, RF, RIDGE, and SVR models are used as base-learners and the Gaussian process (GP) as meta-learner in the stacking-ensemble learning approach. Mean absolute error and symmetric mean absolute percentage error criteria were used to evaluate. In most cases, the SVR and stacking-ensemble learning perform better regarding adopted criteria than compared models. Oliveira et al. (2021) applied the ANN model to predict the number of COVID-19 confirmed cases and deaths and to forecast the future seven days for the series of USA, Brazil and Portugal. The results obtained from the simulations showed that the prediction of confirmed cases and deaths made by ANN is successful. Chandu (2020) conducted research on a number of confirmed cases in India and Thailand with the ARIMA model. It was emphasized that ARIMA (2,1,1) model is the best model based on the available data set, and ARIMA (2,1,1) modeling was performed to forecast the confirmed cases in both countries. Singh et al. (2020) proposed a hybrid model to forecast the daily COVID-19 death cases from the five countries, namely USA, the UK, France, Italy and Spain. The proposed method involves the application of wavelet decomposition to split the data set into component series and then applying ARIMA models to each component series to forecast the death cases. The hybrid model is compared with ARIMA model. The results obtained from the hybrid model showed better performance as compared with ARIMA model.

Although there are many studies in the literature on modeling time series data for the Covid-19 pandemic, studies conducted in Turkey are generally based on a single model and ignore alternative model validity. Regarding the number of observations it covers, this study differs from other existing studies, which generally have a small number of observations made in the early stages of the pandemic. In addition, considering the related studies examined, our study is the first study to evaluate the performance of so many models both for Turkey data set and on behalf of other country data sets and to perform a stacking model on the number of COVID-19 cases.

3. MATERIAL AND METHOD

3.1. Data

The data used within the scope of the study were obtained from the data provided by Our World In Data organization. The data obtained to be used in the study consists of the number of positive cases for a total of 656 days between 11.03.2020 and 24.12.2021, starting from the first case in Turkey. As seen in Figure 1, although the number of cases is relatively low in the first days of the pandemic compared to the following days, a significant increase is observed in the number of cases in each different pandemic wave, as defined by field experts. Although there are pandemic waves occurring in certain periods, as can be seen from Figure 1, the number of cases does not have a clear seasonality and does not have a sharp observable trend.

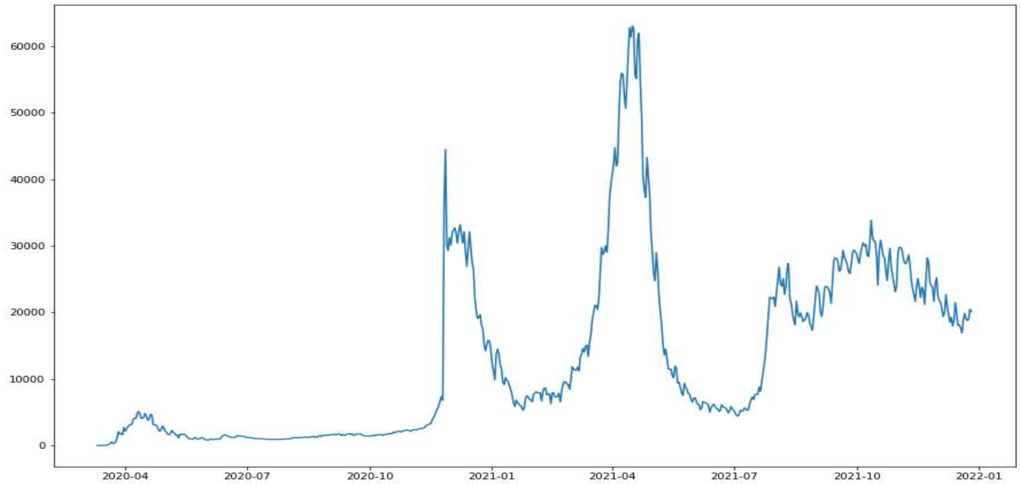

Figure 1. Daily positive case numbers.

3.2. ARIMA

The ARIMA(p,d,q) model is a combination of the Autoregressive (AR) model and the Moving Average (MA) model and the "I" stands for integration. Here, p is the degree of autoregression, d is the degree of difference; q is the degree of moving average. AR(p) model

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + a_t \tag{1}$$

In which: $Z_{t-1}, Z_{t-2}, Z_{t-p}$ are the lags; $\phi_1, \phi_1, \ldots \phi_p$ are the coefficients of the lags and α_t is the white noise series. MA(q) model

$$Z_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} \tag{2}$$

where: $\theta_1, \theta_2, \ldots, \theta_q$ are the lags and $\alpha_t, \alpha_{t-1}, \alpha_{t-2}, \ldots, \alpha_{t-q}$ are the white noise error terms. ARIMA(p,d,q) model

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + a_t + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} \tag{3}$$

The ARIMA model can be estimated if the series is stationary, that is, if its mean and variance are constant. The difference parameter d in the model is the degree of transformation made to make the series stationary (Box et al., 2016: 88).

3.3. Random Forest

Random forest (RF), developed by Breiman (2001) as an ensemble classification and regression approach, provides a unique and very high combination of predictive success and model-interpretability among popular machine learning methods (Qi, 2012: p.307). RF is an ensemble learning technique that consists of many decision trees and results in a decrease in variance compared to a standard single decision tree (Couronné et al., 2018). The RF algorithm, which has proven its validity with many studies, can show very successful results in classification and regression tasks. RF independently generates K number of regression trees for $h_k(x), (k = 1, \ldots\ , K)$ using an x input vector, and the model prediction is obtained as the average of the prediction from each tree in the generated forest (Seo et al., 2017). The equation of the RF regression is presented in Equation (4).

$$Random\ Forest\ Prediction = \frac{1}{K}\sum_{k=1}^{K} h_k(x) \tag{4}$$

As can be seen from Equation (4), the sample variance decreases because the model averages the predicted values from individual trees.

3.4. Support Vector Regression

Support Vector Machines (SVM) was introduced by Vapnik (1995) and gained popularity in the application field due to its many attractive features and promising application performance (Gunn, 1998). Although it is mostly used in classification tasks, it also provides successful results in regression tasks. SVMs can be easily applied to regression problems (Smola, 1996) by introducing an alternative loss function, including a distance measure (Gunn, 1998), and successful results are obtained. Support Vector Regression (SVR) uses the principle of inherent risk minimization, which tries to minimize an upper bound of the generalization error rather than minimizing the prediction error in the training set. This mechanism offers SVR a greater advantage in generalizing the input-output relationship learned during the training phase to generate better predictive values with new input data (Chen & Wang, 2007).

3.5. K-Nearest Neighbors

Due to its simplicity and intuitiveness, the k-nearest neighbors (KNN) algorithm has become one of the widely used machine learning techniques for classification and regression. The KNN algorithm is a method used to classify objects based on the closest training observations within the sample set. The same method can be used for regression by assigning a property value to the object to be the average of the values of its nearest K neighbor. Generally, it can be beneficial on performance to weight the contributions of neighboring observations to obtain better results. Thus, close neighbors will contribute more to the mean than distant neighbors (Imandoust & Bolandraftar, 2013). KNN regression is basically a sample-based learning algorithm, which is non-parametric and does not require any assumptions about the distribution of the data. It is a very useful method in terms of learning the complex target function quickly without losing information. As can be seen in Equation (5), for a certain x input of the training data, K observations close to x_i are taken into account, and the mean of these K independent observations constitutes the $\hat{y}$ value (Goyal et al., 2014).

$$\hat{y}(x) = \frac{1}{k}\sum_{x_i \in N_{k(x)}} y_i \tag{5}$$

The $N_{k(x)}$ in Equation (5) represents the nearest point to its neighbor K for the x_i observation. Although various distance measures are used to measure the proximity between points, the Euclidean distance is commonly used.

3.6. Extreme Gradient Boosting

Developed by Chen and Guestrin (2016), Extreme Gradient Boosting, or more commonly XGBoost, is an efficient and scalable implementation of the gradient boosting framework (Friedman, 2001). Unlike traditional tree learning, XGBoost not only extracts information from its predecessors, but also aggregates the scores on the corresponding leaves to reduce the errors of the previous tree and get more accurate results at the end. XGBoost has many other features such as parallel and distributed computing that can accelerate learning and predict high accuracy (Zhao et al., 2019). The most obvious difference between XGBoost and other gradient boosting methods is that XGboost uses a new editing technique to control over fitting. In this way, it can provide faster and more reliable results during model creation (Al Daoud, 2019). The editing technique is done by adding a new term to the loss function as follows:

$$L(f) = \sum_{i=1}^{n} L(\hat{y}_i, y_i) + \sum_{m=1}^{M} \Omega(\delta_m) \tag{6}$$

$$\Omega(\delta_m) = a|\delta| + 0.5\beta\|w\|^2 \tag{7}$$

where: $|\delta|$ is the number of branches, w is the value of each leaf, and Ω is the regularization function. In addition, XGBoost uses a new gain function presented in Equation (8), different from the split criteria of standard decision trees.

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L+\beta} + \frac{G_R^2}{H_R+\beta} + \frac{(G_L+G_R)^2}{H_L+H_R+\beta}\right] - a \tag{8}$$

$$\mathrm{G}_j = \sum_{i \epsilon I_j} g_i \; ve \; \mathrm{H}_j = \sum_{i \epsilon I_j} h_i \tag{9}$$

where: G represents the score of the child on the right, H is the score of the child on the left and $Gain$ the score if there is no new child (Zhang and Haghani, 2015).

3.7. CatBoost

CatBoost, is a new gradient boosting algorithm proposed by Prokhorenkova et al. (2018) and it is a method that works successfully with categorical features with the lowest information loss and can be successful in regression tasks, although it is mainly used in classification tasks (Jabeur et al., 2021). This machine learning method is based on an advanced gradient boosting decision tree that can solve problems with noisy data, heterogeneous features and complex dependencies. (Shahriar et al., 2021). This algorithm differs from traditional gradient boosting decision tree algorithms in the following aspects.

First, it deals with categorical features during the training period rather than the preprocessing time. CatBoost allows the use of the entire dataset for training. Second, it can identify all categorical features as a new feature. CatBoost isn't particularly stingy at factoring in combinations for categorical data. Finally, it can cope with the deviation that occurs when converting categorical variables to numerical values with the TS method (Huang et al., 2019).

3.8. AdaBoost

Freund and Schapire's (1997) AdaBoost algorithm is one of the first practical boosting algorithms (Schapire, 2013: p.38). In other words, AdaBoost can transform a less successful learning algorithm into an arbitrarily more accurate and powerful learning algorithm with slightly better accuracy than random guessing. AdaBoost brings a new method and a new design idea to the design of the learning algorithm (Ying et al., 2013). The method corrects incorrect predictions made by weak learning algorithms and is less sensitive to overfitting than most learning algorithms (Hu et al., 2008). Given a set of training examples, AdaBoost first preserves the current probability distribution W, and then AdaBoost calls the weak learner in a loop. Training samples are obtained with the W_t distribution at each loop step (t stands for loop step). Then the weak learner h_t is trained. The W_t distribution is updated after each cycle according to the prediction results in the training samples. "Easy" examples correctly classified by the weak learner receive low weight, and "difficult" examples that are misclassified are assigned a higher weight, allowing AdaBoost to focus on "difficult" examples in the next step (Li et al., 2005).

3.9. Multi-Layer Perceptron

Artificial Neural Networks (ANN) are structures similar to the large neuron network in the brain, inspired by the functioning of the human brain. ANNs are basically semi-parametric regression predictors, and although they are successful on linear relationships, they are also very successful structures in simulating the behavior of complex nonlinear relationships (McCluskey et al., 2013). These Networks can predict model functions and process linear-non-linear functions by learning from data relationships and generalizing invisible states. One of the popular Neural Networks is Multi-Layer Perceptron (MLP) (Taud & Mas, 2018). Although the architecture of a multilayer perceptron is variable depending on the data and the task to be applied, it generally consists of several neuron layers, as shown in Figure 2 (Gardner & Dorling, 1998).

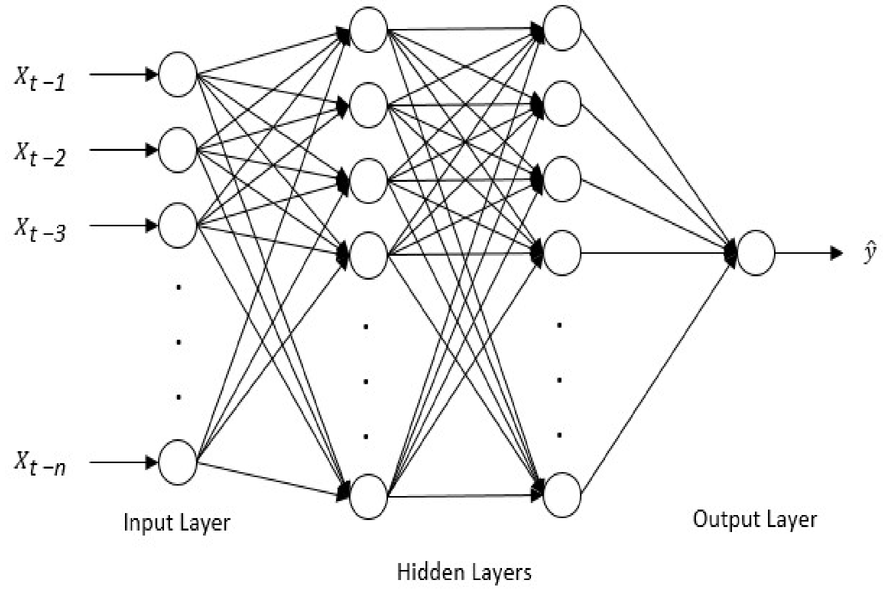

Figure 2. Multi-Layer perceptron architecture

Although there are various methods and network topologies for the development and implementation of the neural network prediction model, the most common usage is the feedforward neural network topology with backpropagation learning algorithm used in this study (Ali et al., 2017). In an MLP, the model layers are completely interconnected, that is, each neuron in one layer is connected to each neuron in the next layer. The learning process in artificial neural networks is realized by calculating the weights in each connection so that the expected output can be calculated correctly in response to the observations given during the training phase. In an ANN, the learning algorithm is basically responsible for the task of calculating the weights of the neural links of the network we have given. This learning algorithm is used to train the entire neural network with gradient search to minimize the square of errors between the output and the expected.

3.10. Stacked Generalization

Stacked generalization, which is an ensemble learning method, or stacking with its simpler use, is an approach that treats the results of the predictions in the validation set as input regressions for next-level models (Pavlyshenko, 2018), and it is a very effective application in increasing the performance of individual models (Wolpert, 1992). Its main purpose is to enable researchers to combine different prediction algorithms on the same task into an individual algorithm (Naimi & Balzer, 2018). Stacking is about combining learners created using different machine learning algorithms $L_1, \ldots, L_N$, in a single dataset consisting of $s_i = (x_i, y_i)$ samples, that is, pairs of feature vectors (x_i) (Džeroski and Ženko, 2004). Predictions created by different learners in the first stage are used as an input for second-level learning algorithms, often called meta-learners, in the second stage.

3.11. Propesed Approach

The proposed approach aims to effectively combine the forecasts of different modeling techniques, especially machine learning models used in time series modeling, along with different model orders of the ARIMA process. For this purpose, a meta-learner is developed by using the prediction results obtained from more than one individual model. Thus, it is aimed to increase the forecast performance with the stacking approach, which is a sub-branch of the ensemble learning methodology.

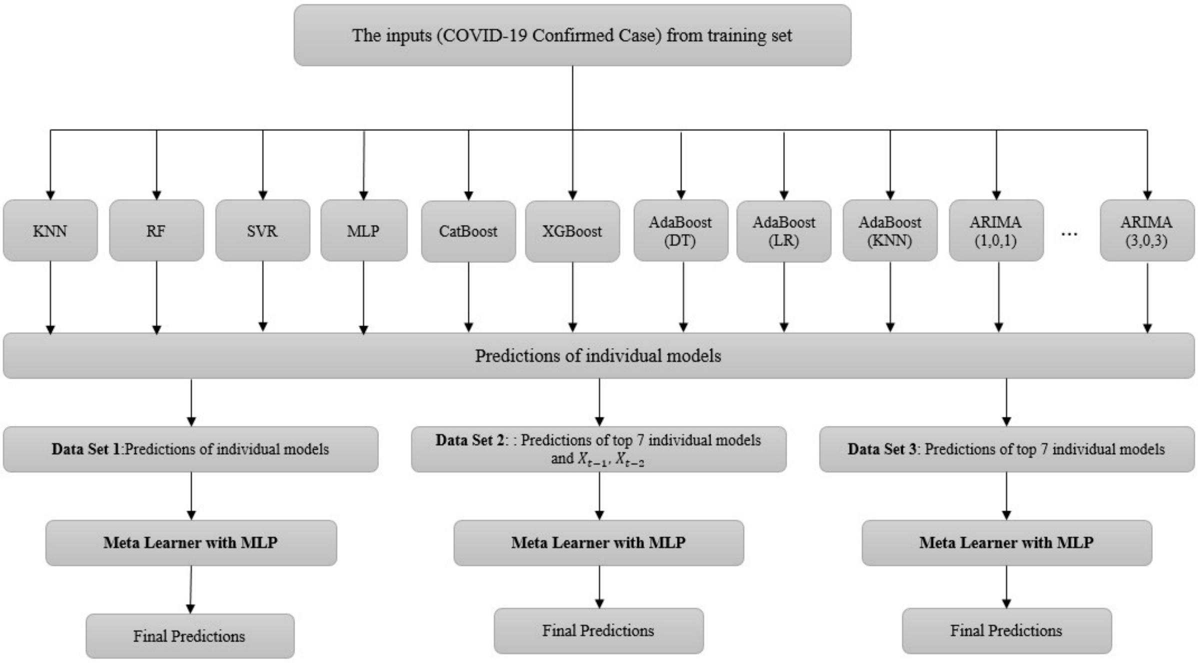

Figure 3. The diagram of the proposed stacking ensemble.

As seen in Figure 3, daily positive case numbers of COVID-19 were used as input dataset for all models. Then, parameter adjustments are made that will allow each model to forecast with the highest success rate on this data set. After suitable model designs are found, each model is separately designed iteratively for 60 observations that do not exist in the data set, with walk-forward validation, and forecast values for 60 days are obtained. These forecast values obtained within the scope of each model form the input data set for the multi-layer perceptron, which is the meta-learner to be used in the next step. An important point here is to present the prediction data of 60 observations used as the test data set to the meta-learner. Thus, the risk of both overfitting and bias of the meta-learner is eliminated. For this reason, it would not be wrong to classify the stacking method created within the scope of the study as a super learner. In this context, it is aimed to compare the forecasting performances of three different meta-learners created by experimental design by deriving different feature sets from these forecasted values. As can be seen in Figure 3, the first data set created with the forecasting values obtained from the individual models contains the forecasting values of only 60 observations created by 16 different models. The second data set, on the other hand, has nine different independent variables in total with the forecasting values of the seven models with the best performance value and the lags $x_{(t-1)}$, $x_{(t-2)}$. The third data set is created with the forecasting values of only the seven models with the best performance value in order to better understand the effect of individually successful individual models on the forecasting. In this context, the selected models and their hyperparameter estimates are presented in detail in the next section.

3.12. Model Evaluation

Within the scope of the study, there is a two-stage approach in evaluating model performances. First, walk-forward validation was used in the performance evaluation and comparison of the individual models. Walk-forward validation is an approach in which the model makes one-by-one forecasting for each observation in the test dataset (Kaastra & Boyd, 1996). After each forecasting is made for a time step in the test dataset, the actual observation for the forecasted value is added to the test dataset and presented to the model. For this purpose, 10 percent of the total data set was used as test data. In the study, the data dated 24-10-2021 constitutes the first test data. Within the scope of each model, the data between 11-03-2020 and 23-10-2021 constitute the first training data, and after the models are created with this training data, an forecast is made for the observation dated 24-10-2021. Then, the observation dated 24-10-2021 is added to the training dataset, and the model is trained again and makes a forecast for the next day. Thus, this step is repeated 60 times to obtain the forecast values one by one for the last 60 days. In the second stage, a separation of 20 percent by 80 percent was used for the data set consisting of 60 predictive values of different models in order to evaluate the performance of meta-learners. Thus, while 48 forecasting values were presented to meta-learners for training, performance metrics were calculated using data from the last 12 days for testing purposes. In this study, it is aimed to evaluate the findings together with the existing studies by preferring the frequently used performance metrics in the time series studies carried out for the COVID-19 pandemic within the scope of the literature. For this purpose, Mean Absolute Error (MAE) in Equation 10 (Ceylan, 2020; Pontoh et al., 2020; Özen et al., 2021; Purwandari et al., 2022; Talkhi et al., 2021), Mean Absolute Percentage Error (MAPE) in Equation 11 (Sevli & Gülsoy, 2020; Karcıoğlu et al., 2021; Özen et al., 2021; Shastri et al., 2020; Arora et al., 2020; Talkhi et al., 2021) and Root Mean Square Error (RMSE) in Equation 12 (Ceylan, 2020; Pontoh et al., 2020; Özen et al., 2021; Talkhi et al., 2021; Purwandari et al., 2022) were used to compare model performances as evaluation metrics. Models with minimum values are selected in the most model selection according to the criteria below.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{10}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} * 100\% \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{12}$$

4. APPLICATION AND FINDING

All models in the study were carried out with the Python 3.7 programming language. As mentioned before, in the first stage of the study, individual models should be created and forecast values should be obtained. For this reason, first of all, the hyper-parameters of the machine learning algorithms, which are the individual models, must be determined and appropriate orders must be selected for the ARIMA models. For this purpose, parameter optimization is the first step of the experimental design. In the study, the parameters given in Table 1 were used in the creation of the models in order to obtain the highest success rate by using the Grid Search approach in parameter selection on behalf of machine learning algorithms. Grid Search is a traditional hyperparameter optimization method, and it is a search algorithm that performs a full search on a certain subset of the hyperparameter space of the training algorithm (Liashchynskyi & Liashchynskyi, 2019).

Table 1.
Hyper-Parameters Used In Individual Models

Model	Hyper-parameters	Selected Value
MLP	Batch Size	256
	Hidden Neurons	400
	Epochs	2000
	Lag	8
CatBoost	Learning_Rate	0.03
	Iterations	3000
	Depth	10
	Bagging_Temperature	0.2
	Lag	8
KNN	K	3
	Lag	8
Random Forest	N_Estimators	100
	Max_Depth	4
	Lag	8
XGBoost	N_Estimators	0,053
	Max_Depth	0,063
	Lag	12

Table 1. continuation

Hyper-Parameters Used In Individual Models

Model	Hyper-parameters	Selected Value
SVR	C	0.1
	Gamma	0.0000001
	Kernel	linear
	Lag	12
AdaBoost - LR	Learning_Rate	0.00001
	N_Estimator	2000
	Lag	17
AdaBoost - DT	Learning_Rate	0.05
	N_Estimators	100
	Max_Depth	6
	Lag	10
AdaBoost - KNN	Learning_Rate	3000
	N_Estimators	0.01
	N_Neighbors	3
	Lag	17

In addition, Table 1 also presents the lag values used for each model, except for the hyper-parameters used in the model configuration. Although eight lag data are considered for most models, it has been observed in empirical studies that different lag values for some models result in superior (meta) model performance. ARIMA, another model used in the study, has a different modeling process than machine learning algorithms.

After the parameter adjustments of all models are completed, walk-forward validation is performed to forecast the observation data for the last 60 days. In Table 2, the forecast performance of each individual model is summarized according to three different metrics, and as can be seen, the multi-layer perceptron exhibited the best forecasting performance among the individual models with an MAE improvement of approximately 1,8 percent compared to its closest competitor, AdaBoost with the linear regression basic model. After the two models, SVR and ARIMA(3,0,3) models come with MAE values of 976,60 and 1023,51, respectively. Although the model that achieved the best forecast value was MLP, the models indicated in bold in Table 2 showed an acceptable forecasting performance close to the MAE value obtained by MLP. Among ARIMA models, ARIMA(3,0,3) is clearly seen as the most successful model, but the performances of ARIMA(3,0,2), ARIMA(2,0,3) and ARIMA(2,0,2) models are also at an acceptable level.

Table 2.

Forecast Performance of Individual Models Over 60 Observations

Model	MAE	MAPE	RMSE
ARIMA(1,0,1)	1167,33	5,2	1534,30
ARIMA(1,0,2)	1140,22	5,0	1460,95
ARIMA(2,0,1)	1149,93	5,1	1495,12
ARIMA(2,0,2)	**1135,15**	**5,0**	**1455,72**
ARIMA(2,0,3)	**1129.33**	**5,0**	**1451.71**
ARIMA(3,0,2)	**1115,25**	**4,9**	**1440,30**
ARIMA(3,0,3)	**1023,51**	**4,5**	**1329,49**
MLP	**925,87**	**4,0**	**1181,55**
Random Forest	1467,19	6,5	1776,08
KNN	1474,55	6,5	1761,23
SVR	**976,60**	**4,2**	**1271,74**
CatBoost	1201,77	5,2	1441,01
XGBoost	1215,15	5,3	1469,91
AdaBoost – Decion Tree	1426,38	6,3	1712,93
AdaBoost – Linear Regression	**942,52**	**4,1**	**1192,53**
AdaBoost – KNN	1374,24	5,9	1740,89

In the next step, a new data set is created with the forecast values of 60 observations obtained from the individual models and given to the meta-learner as input. As mentioned before, 3 different feature sets were created from the forecasted values obtained. Thus, a basic feature engineering approach was carried out in order to achieve higher forecasting performance. While the first dataset consists of only model forecast values, the second dataset includes the forecast values and x_{t-1}, x_{t-2} lags of the seven models with the best performance values. The third data set was created using only the forecasted values of the seven models with the best performance values indicated in bold in Table 2. In the study, multi-layer perceptron is used as a meta-learner in order to combine the forecasts of the individual models. In addition, although many methods were tested as meta-learners within the scope of the study, the results of meta-learners created with other models were not reported due to the superior performance of MLPs as a result of almost all trials. Three different model configurations are needed for three different data sets. As a result of the results obtained by applying hyper-parameter optimization in the meta learner, the parameter values used in the models are given in Table 3.

Table 3.
Forecast Performance of Individual Models Over 60 Observations

Data Set	Hyper-parameters	Selected Value
Dataset 1	Batch Size	128
	Hidden Neurons-1	32
	Hidden Neurons-2	12
	Epochs	2000
Dataset 2	Batch Size	256
	Hidden Neurons	36
	Hidden Neurons-2	10
	Epochs	1000
Dataset 3	Batch Size	300
	Hidden Neurons-1	21
	Hidden Neurons-2	14
	Epochs	1500

Table 4 contains metrics for the performance of meta-learners and some high-performing individual models. In addition, the results of simple ensemble learning approaches, which are formed with the mean and median values of the forecast values of all individual models, are presented in order to form a basis for performance comparison. It should be noted that the performance metrics presented here belong to the test dataset, which accounts for 20 percent of the last 60 observations. In Table 4, it is seen that all of the ensemble learning approaches created give better results than individual methods. In addition, among the meta-learners, it is seen that the feature set including the best seven models and the variables x_{t-1}, x_{t-2} produced very successful results with an MAE value of 559,97 and a MAPE value of 2,9 percent. Considering the individual model forecast, it is seen that the most successful model for the test data set used is Adaboost based on linear regression. It should also be emphasized that, due to the stochastic nature of machine learning approaches, the forecast values obtained from each model run change, and therefore performance metrics also change. For this reason, the metrics presented in Table 4 are calculated by averaging the forecast results obtained as a result of 30 repeated model runs, especially for MLP, SVR and stacking models. Similarly, the forecast values presented in Table 5 reflect the averages obtained as a result of repeated model tests.

Table 4.
Forecast Performance of Individual Models And Meta Learners Over Last 12 Observations

Model	MAE	MAPE	RMSE
Simple Mean	659,53	3,5	855,92
Simple Median	650,78	3,4	871,71
Metalearner with Dataset 1	**592,69**	**3,1**	**833,08**
Metalearner with Dataset 2	**559,97**	**2,9**	**799,82**
Metalearner with Dataset 3	**575,13**	**3,0**	**834,93**
ARIMA(3,0,2)	715,48	3,7	971,55
ARIMA(3,0,3)	689,88	3,6	878,18
MLP	773,87	4,1	1010,21
SVR	628,74	3,3	837,38
AdaBoost – Linear Regression	594,13	3,1	798,46

As many existing studies with ensemble learning methods demonstrate, the results show that a collective approach combining the forecasts of individual methods produces much better results than individual methods. Although it uses different period intervals and country data, considering similar studies in the literature, MAPE values obtained from meta-learners created using stacking show much better results than many existing studies. In order to better understand the forecast results obtained, the meta-learner forecast values and the observations of the expected values are presented in Table 5. As can be seen here, there are very few forecasting errors between the forecast values and the expected value.

Table 5.
Forecast values of meta-learners over test set

Date	Confirmed Positive Case	Metalearner (Dataset 1) Forecast	Metalearner (Dataset 2) Forecast	Metalearner (Dataset 3) Forecast
15.12.2021	19872	20815,33	20963	21126,02
16.12.2021	18100	18462,38	18683,33	18419,38
17.12.2021	18141	17961,69	17991,33	18219,63
18.12.2021	17644	18071,18	18024,86	18121,82
19.12.2021	16910	16751,55	16755,15	16689,23
20.12.2021	18762	17775,09	17550,85	17314,39
21.12.2021	19859	19743,81	19782,57	19714,21
22.12.2021	19095	18861,13	19065,68	18962,33
23.12.2021	18771	17828,8	17967,35	18164,91
24.12.2021	18910	18593,52	18747,38	18868,26
25.12.2021	20470	18235,66	18522,7	18505,65
26.12.2021	20138	20350,7	20267,48	20351,67

5. CONCLUSION

Throughout history, humanity frequently encounters natural disasters such as earthquakes, floods, avalanches and hurricanes that cause significant human and economic suffering. Therefore, experience is gained to deal with such natural disasters and to minimize their damage. However, individuals and governments seem to be caught off guard as pandemics rarely occur (Petropoulos et al., 2022). While the number of epidemics worldwide has nearly quadrupled in the last 60 years, their annual number has more than tripled since 1980 (Walsh, 2017). Epidemiologists state that due to globalization and increasing social interaction among people around the world, pandemics will occur more frequently and therefore it is necessary to be prepared for unexpected pandemics (Petropoulos et al., 2022).

Since COVID-19 has spread almost across any country and is a serious threat to mankind, it was declared to be a pandemic by WHO. Forecasting the results of a pandemic is a quite important and difficult task for policy makers and decision makers. Predicting the prevalence of the disease is critical for health departments to strengthen surveillance systems and reallocate resources for the entire health system during the pandemic (Ceylan, 2020).

In our study, in addition to comparing these algorithms by making case forecasts with machine learning algorithms, the forecasts of all models used with the stacking approach were combined under a single forecast. The forecasting of validated cases through the developed stacking model has been carried out with high accuracy. As many existing studies with ensemble learning methods have shown, the results exhibit that a collective approach that combines the forecasts of individual methods produces much better results than individual methods. However, it is seen that all of the developed ensemble learning approaches give better results than individual methods. In addition, among the meta-learners, it is seen that the best seven models and the feature set with the variables $x_{(t-1)}$, $x_{(t-2)}$ performed very successful results with the MAE value and the MAPE value of 2,9 percent. Although we use different time periods and country data, considering similar studies in the literature, MAPE values obtained from meta-learners created using stacking show much better results than many existing studies. This study provides evidence that the number of cases in Turkey can be forecasted by taking into account only the historical data, without considering the ensemble model proposed within the scope of the study and other factors affecting the current number of cases. At the same time, this study shows promising results in creating action plans using advanced time series forecasting models for both current pandemic conditions and possible pandemic situations.

REFERENCES

Abdulmajeed, K., Adeleke, M., & Popoola, L. (2020). Online forecasting of COVID-19 cases in Nigeria using limited data. *Data in Brief*, *30*. https://doi.org/10.1016/j.dib.2020.105683

Ahmar, A. S., & del Val, E. B. (2020). SutteARIMA: Short-term forecasting method, a case: Covid-19 and stock market in Spain. *Science of the Total Environment*, *729*. https://doi.org/10.1016/j.scitotenv.2020.138883

Akay, S., & Akay, H. (2021). Time series model for forecasting the number of COVID-19 cases in Turkey. *Turkish Journal of Public Health*, *19(2*), 140-145. https://doi.org/10.20518/tjph.809201.

Al Daoud, E. (2019). Comparison between XGBoost, LightGBM and CatBoost using a home credit dataset. *International Journal of Computer and Information Engineering*, *13*(1), 6-10. https://doi.org/10.5281/zenodo.3607805

Ali, M., Khan, D. M., Aamir, M., Khalil, U., & Khan, Z. (2020). Forecasting COVID-19 in Pakistan. *PLoS One*, *15*(11). https://doi.org/10.1371/journal.pone.0242762.

Ali, Z., Hussain, I., Faisal, M., Nazir, H. M., Hussain, T., Shad, M. Y., ... & Hussain Gani, S. (2017). Forecasting drought using multilayer perceptron artificial neural network model. *Advances in Meteorology, 2017*. https://doi.org/10.1155/2017/5681308.

Arora, P., Kumar, H., & Panigrahi, B. K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, Solitons & Fractals*, *139*. https://doi.org/10.1016/j.chaos.2020.110017.

Biswas, P. K., Islam, M. Z., Debnath, N. C., & Yamage, M. (2014). Modeling and roles of meteorological factors in outbreaks of highly pathogenic avian influenza H5N1. *PloS One*, *9*(6). https://doi.org/10.1371/journal.pone.0098471.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. John Wiley & Sons.

Breiman, L. (2001) Random forests. *Machine Learning*, *45*, 5-32. https://doi.org/10.1023/A:1010933404324.

Ceylan, Z. (2020). Estimation of COVID-19 prevalence in Italy, Spain, and France. *Science of The Total Environment*, *729*. https://doi.org/10.1016/j.scitotenv.2020.138817.

Chandu, V. C. (2020). Time series forecasting of COVID-19 confirmed cases with ARIMA model in the South East Asian countries of India and Thailand: A comparative case study. *medRxiv*, 2020-05.

Chen, K. Y., & Wang, C. H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism management*, *28*(1), 215-226. https://doi.org/10.1016/j.tourman.2005.12.018.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785-794, arXiv:1603.02754.

Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, *135*. https://doi.org/10.1016/j.chaos.2020.109864

Couronné, R., Probst, P., & Boulesteix, A. L. (2018). Random forest versus logistic regression: A large-scale benchmark experiment. *BMC bioinformatics*, *19*(1), 1-14. https://doi.org/10.1186/s12859-018-2264-5.

Dairi, A., Harrou, F., Zeroual, A., Hittawe, M. M., & Sun, Y. (2021). Comparative study of machine learning methods for COVID-19 transmission forecasting. *Journal of Biomedical Informatics*, *118.* https://doi.org/10.1016/j.jbi.2021.103791

De Oliveira, L. S., Gruetzmacher, S. B., & Teixeira, J. P. (2021). COVID-19 time series prediction. *Procedia Computer Science, 181,* 973-980. https://doi.org/10.1016/j.procs.2021.01.254.

Dehesh, T., Mardani-Fard, H. A., & Dehesh, P. (2020). Forecasting of covid-19 confirmed cases in different countries with arima models. *MedRxiv.* https://doi.org/10.1101/2020.03.13.20035345.

Ding, G., Li, X., Jiao, F., & Shen, Y. (2020). Brief Analysis of the ARIMA model on the COVID-19 in Italy. *medRxiv*. https://doi.org/10.1101/2020.04.08.20058636.

Džeroski, S., & Ženko, B. (2004). Is combining classifiers with stacking better than selecting the best one?. *Machine learning*, *54(3*), 255-273, https://doi.org/10.1023/B:MACH.0000015881.36452.6e .

Earnest, A., Chen, M. I., Ng, D., & Sin, L. Y. (2005). Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore. *BMC Health Services Research*, *5*(1), 1-8. https://doi.org/10.1186/1472-6963-5-36.

Fidan, H., & Yuksel, M. E. (2022). A comparative study for determining Covid-19 risk levels by unsupervised machine learning methods. *Expert Systems with Applications*, *190.* https://doi.org/10.1016/j.eswa.2021.116243.

Freund, Y., & Schapire, R.E. (1997). A decision-theoretic generalization of on-line learning and anapplication to boosting. *J. Comput. Syst. Sci. 55*(1), 119–139. https://doi.org/10.1006/jcss.1997.1504.

Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451.

Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*, 32(14-15), 2627-2636. https://doi.org/10.1016/S1352-2310(97)00447-0

Goyal, R., Chandra, P., & Singh, Y. (2014). Suitability of KNN regression in the development of interaction based software fault prediction models. *Ieri Procedia*, *6*, 15-21. https://doi.org/10.1016/J.IERI.2014.03.004.

Guan, P., Huang, D. S., & Zhou, B. S. (2004). Forecasting model for the incidence of hepatitis A based on artificial neural network. *World journal of gastroenterology: WJG*, *10*(24), 3579-3582. https://doi.org/10.3748/wjg.v10.i24.3579.

Gunn, S. R. (1998). Support vector machines for classification and regression. *ISIS technical report, 14*(1), 5-16. https://see.xidian.edu.cn/faculty/chzheng/bishe/indexfiles/new_folder/svm.pdf.

Hu, W., Hu, W., & Maybank, S. (2008). Adaboost-based algorithm for network intrusion detection. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *38*(2), 577-583. https://doi.org/10.1109/TSMCB.2007.914695.

Huang, G., Wu, L., Ma, X., Zhang, W., Fan, J., Yu, X., ... & Zhou, H. (2019). Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *Journal of Hydrology*, *574*, 1029-1041. https://doi.org/10.1016/j.jhydrol.2019.04.085.

Imandoust, S. B., & Bolandraftar, M. (2013). Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background. *International Journal of Engineering Research and Applications*, *3(5)*, 605-610. https://www.scopus.com/record/display.uri?eid=2-s2.0-84934906871&origin=inward&txGid=194a430bb8abb0e83d36a946950d48e1&featureToggles=FEATURE_NEW_DOC_DETAILS_EXPORT:1

Jabeur, S. B., Gharib, C., Mefteh-Wali, S., & Arfi, W. B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, *166.* https://doi.org/10.1016/j.techfore.2021.120658.

Kaastra, I., & Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, *10*(3), 215-236. https://doi.org/10.1016/0925-2312(95)00039-9.

Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P. (2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC bioinformatics*, *15*(1), 1-9. https://doi.org/10.1016/j.jbi.2021.103791

Karcıoğlu, A. A., Tanışman, S., & Bulut, H. (2021). Time series forecasting of COVID-19 transmission in Turkey using ARIMA model and LSTM Network. *Avrupa Bilim ve Teknoloji Dergisi*, *(32)*, 288-297. https://doi.org/10.31590/ejosat.1039394. (In Turkish)

Katris, C. (2021). A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. *Expert systems with applications*, *166*. https://doi.org/10.1016/j.eswa.2020.114077.

Khan, M., Mehran, M. T., Haq, Z. U., Ullah, Z., Naqvi, S. R., Ihsan, M., & Abbass, H. (2021). Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review. *Expert systems with applications*, *185*. https://doi.org/10.1016/j.eswa.2021.115695.

Koçak, M. (2020). A comparison of time-series models in predicting COVID-19 cases. *Türkiye Klinikleri Biyoistatistik*, *12*(1), 89-96. https://doi.org/10.5336/biostatic.2020-75402.

Kumar, N. & Susan, S. (2020). COVID-19 pandemic prediction using time series forecasting models. In: *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. IEEE, 1-7, https://doi.org/10.1109/ICCCNT49239.2020.9225319

Lai, D. (2005). Monitoring the SARS epidemic in China: a time series analysis. *J Data Sci*, *3(3*), 279-93. https://doi.org/10.6339/JDS.2005.03(3).229.

Li, X., Wang, L., & Sung, E. (2005). A study of AdaBoost with SVM based weak learners. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks,* IEEE, 1, 196-201. https://doi.org/10.1109/IJCNN.2005.1555829

Liashchynskyi, P., & Liashchynskyi, P. (2019). Grid search, random search, genetic algorithm: A big comparison for NAS. *arXiv.* https://doi.org/10.48550/arXiv.1912.06059%20Focus%20to%20learn%20more

Maleki, M., Mahmoudi, M. R., Wraith, D., & Pho, K. H. (2020). Time series modelling to forecast the confirmed and recovered cases of COVID-19. *Travel medicine and infectious disease*, 37. https://doi.org/10.1016/j.tmaid.2020.101742.

McCluskey, W. J., McCord, M., Davis, P. T., Haran, M., & McIlhatton, D. (2013). Prediction accuracy in mass appraisal: a comparison of modern approaches. *Journal of Property Research*, *30(4)*, 239-265. https://doi.org/10.1080/09599916.2013.781204.

Naimi, A. I., & Balzer, L. B. (2018). Stacked generalization: an introduction to super learning. *European journal of epidemiology*, *33(5)*, 459-464. https://doi.org/10.1007/s10654-018-0390-z.

Özen, N. S., Saraç, S., & Koyuncu, M. (2021). Prediction of COVID-19 Cases in the United States of America with Machine Learning Algorithms. *Avrupa Bilim ve Teknoloji Dergisi, (22),* 134-139. https://doi.org/10.31590/ejosat.855113Abstract. (In Turkish)

Pavlyshenko, B. (2018). Using stacking approaches for machine learning models. In *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP),* IEEE, 255-258. https://doi.org/10.1109/DSMP.2018.8478522.

Papastefanopoulos V., Linardatos P., & Kotsiantis S. (2020). COVID-19: A comparison of time series methods to forecast percentage of active cases per population. *Applied Sciences*, *10*(11), 3880. https://doi.org/10.3390/app10113880.

Petropoulos, F., Makridakis, S., & Stylianou, N. (2022). COVID-19: Forecasting confirmed cases and deaths with a simple time series model. *International Journal of Forecasting*, *38*, 439-452. https://doi.org/10.1016/j.ijforecast.2020.11.010

Pontoh, R. S., Zahroh, S., Hidayat, Y., Aldella, R., Jiwani, N. M., & Firman, S. (2020). Covid-19 modelling in South Korea using a time series approach. *Int. J. Adv. Sci. Technol*, *29*(7), 1620-1632. http://sersc.org/journals/index.php/IJAST/article/view/16246.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural İnformation Processing Systems*, *31*. https://doi.org/10.48550/arXiv.1706.09516%20Focus%20to%20learn%20more

Purwandari, T., Zahroh, S., Hidayat, Y., Sukonob, S., Mamat, M., & Saputra, J. (2022). Forecasting model of COVID-19 pandemic in Malaysia: An application of time series approach using neural network. *Decision Science Letters*, *11*(1), 35-42. https://doi.org/10.5267/j.dsl.2021.10.001

Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C., & dos Santos Coelho, L. (2020). Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos, Solitons & Fractals, 135,* 109853. https://doi.org/10.1016/j.chaos.2020.109853

Qi, Y. (2012). Random forest for bioinformatics. In *Ensemble machine learning*. Springer, MA, 307-323, https://doi.org/10.1007/978-1-4419-9326-7_11

Schapire, R. E. (2013). Explaining adaboost. *In Empirical inference*. Springer, Berlin, Heidelberg, 37-52. https://doi.org/10.1007/978-3-642-41136-6_5

Seo, D. K., Kim, Y. H., Eo, Y. D., Park, W. Y., & Park, H. C. (2017). Generation of radiometric, phenological normalized image based on random forest regression for change detection. *Remote Sensing*, *9*(11). https://doi.org/10.3390/rs9111163.

Sevli, O., & Gülsoy, V. G. B. (2020). Machine learning based case estimation using prophet model with time series data for covid-19 outbreak. *Avrupa Bilim ve Teknoloji Dergisi*, (19), 827-835, https://doi.org/10.31590/ejosat.766623. (In Turkish)

Shahriar, S. A., Kayes, I., Hasan, K., Hasan, M., Islam, R., Awang, N. R., ... & Salam, M. A. (2021). Potential of ARIMA-ANN, ARIMA-SVM, DT and CatBoost for atmospheric PM2. 5 forecasting in Bangladesh. *Atmosphere*, *12*(1), 100. https://doi.org/10.3390/atmos12010100.

Shastri, S., Singh, K., Kumar, S., Kour, P., & Mansotra, V. (2020). Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. *Chaos, Solitons & Fractals*, 140, https://doi.org/10.1016/j.chaos.2020.110227

Singh, S., Parmar, K. S., Kumar, J., & Makkhan, S. J. S. (2020). Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. *Chaos, Solitons & Fractals*, 135, 109866. https://doi.org/10.1016/j.chaos.2020.109866.

Smola, A. J., & B. Schölkopf, (1998). On a kernel–based method for pattern recognition, regression, approximation and operator ınversion. *Algorithmica* 22, 211– 231. *Technical Report 1064.* https://doi.org/10.1007/PL00013831.

Sumi, A., Luo, T., Zhou, D., Yu, B., Kong, D., & Kobayashi, N. (2013). Time-series analysis of hepatitis A, B, C and E infections in a large Chinese city: Application to prediction analysis. *Epidemiology & Infection*, *141*(5), 905-915. https://doi.org/10.1017/S095026881200146X.

Talkhi, N., Fatemi, N. A., Ataei, Z., & Nooghabi, M. J. (2021). Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A comparison of time series forecasting methods. *Biomedical Signal Processing and Control*, 66. https://doi.org/10.1016/j.bspc.2021.102494.

Tandon, H., Ranjan, P., Chakraborty, T., & Suhag, V. (2020). Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future. *arXiv preprint arXiv:2004.07859*. https://doi.org/10.48550/arXiv.2004.07859.

Taud, H., & Mas, J. F. (2018). Multilayer perceptron (MLP). In *Geomatic approaches for modeling land change scenarios*. Springer, 451-455. https://doi.org/10.1007/978-3-319-60801-3_27.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer https://books.google.com.tr/books?hl=tr&lr=&id=sna9BaxVbj8C&oi=fnd&pg=PR7&dq=).+The+Nature+of+Statistical+Learning+Theory&ots=oqL9H_jrc6&sig=a5xo-MNgbvj_3GrI92a2kcUHDPE&redir_esc=y#v=onepage&q=).%20The%20Nature%20of%20Statistical%20Learning%20Theory&f=false.

Walsh, B. (2017). *The world is not ready for the next pandemic*. https://time.com/4766624/next-global-security/
Wolpert, D. H. (1992). Stacked generalization. *Neural networks, 5*(2), 241-259. https://doi.org/10.1016/S0893-6080(05)80023-1
Ying, C., Qi-Guang, M., Jia-Chen, L., & Lin, G. (2013). Advance and prospects of AdaBoost algorithm. *Acta Automatica Sinica, 39*(6), 745-758. https://doi.org/10.1016/S1874-1029(13)60052-X.
Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons & Fractals*, 140, https://doi.org/10.1016/j.chaos.2020.110121.
Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies, 58*, 308–324. https://doi.org/10.1016/j.trc.2015.02.019.
Zhao, Y., Chetty, G., & Tran, D. (2019). *Deep Learning with XGBoost for Real Estate Appraisal*, In 2019 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, pp. 1396-1401, Xiamen- China, December, https://doi.org/10.1109/SSCI44817.2019.9002790.

ÇALIŞMANIN ETİK İZNİ

Yapılan bu çalışmada "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbiri gerçekleştirilmemiştir. Ayrıca çalışma kapsamında "Etik İzin" gerektiren bir durum bulunmamaktadır.

ARAŞTIRMACILARIN KATKI ORANI

1.yazarın araştırmaya katkı oranı %50, 2. yazarın araştırmaya katkı oranı %50'dir.

Yazar 1: Metodoloji, Yazılım, Araştırma, Veri düzenleme, Yazma – Orijinal taslak, Görselleştirme.

Yazar 2: Metodoloji, Yazma – Orijinal taslak, İnceleme ve Düzenleme, Kavramsallaştırma, Proje yönetimi.