SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
VOL. 5, NO. 1, APRIL 2022
DOI: 10.35377/saucis.05.01.999373

Research/Review Article

Augmented Artificial Neural Network Model for the COVID-19 Mortality Prediction: Preliminary Analysis of Vaccination in Turkey

Sena Kır [1], Elif Elçin Günay [2]
[1] Department of Industrial Engineering, Sakarya University, Sakarya, 54050, Turkey; senas@sakarya.edu.tr; (090)264295-5965
[2] Department of Industrial Engineering, Sakarya University, Sakarya, 54050, Turkey; ekabeloglu@sakarya.edu.tr;

Received 22 September 2021; Accepted 26 January 2022; Published online 30 April 2022

Abstract

The spread and severity of coronavirus disease 2019 (COVID-19) have a severe impact on our lives, such that over 5.8 million lives have been lost since it has been first emerged. Although prediction of the COVID-19 mortality may be inevitably accompanied by uncertainty, it is helpful for health politicians and public health decision-makers to take proper precautions to diminish the pandemic's severity. Therefore, this study proposed a mortality prediction model for the deaths that occur on-day, lag 1 day, lag 7 day, and lag 14 day in Turkey, considering 16 variables under four categories as follows: (i) severity of the disease, (ii) vaccination policy as a preventive strategy, (iii) exposure duration in society, (iv) time series impact. The developed Augmented-Artificial Neural Network (ANN) model took advantage of Auto-Regressive Integrated Moving Average (ARIMA) and ANN models to capture the linear and nonlinear components of the mortality. The proposed model was able to predict mortality with the lowest error compared to ARIMA and ANN models. A set of experiments was designed to reveal the impact of each responsible category on mortality. In the experimental study, it was observed that the impact of four categories from highest to the lowest importance on prediction performance were exposure duration in society, vaccination policy, severity of disease, and time series, respectively. According to these results, new virus-fighting policies can be developed, and the existing model can be used as a simulation tool with the new data to be obtained.

Keywords: ANN, ARIMA, coronavirus, vaccination, estimation

1. Introduction

Starting in December 2019, the coronavirus 2019 (COVID-19) pandemic has been a global threat all around the world. According to the World Health Organization (WHO), there have been over 400 million confirmed COVID-19 cases, including nearly 5.8 million deaths [1]. These numbers have been increasing, which continue to affect our lives, education, and economies severely day by day. To diminish the burden of the pandemic, researchers and scientists put a tremendous amount of effort into how to prevent and treating the pandemic [2]. From this aspect, it is evident that predicting the spread of COVID-19 and its impacts contributes to understanding the adversity of the pandemic and developing effective preventive public health emergency strategies for authorities to limit the disease spread promptly [3]. Also, the medical systems can be designed to be more robust to patient overflows, illness, and deaths considering the predictions of the spread of COVID-19 [4].

Estimating the future trend of the COVID-19 and its impact (e.g., hospitalization, mortality, and demand for ICU beds) has been the focus of recent studies given in Table 1. Due to the dependency of epidemic diseases on many different factors and uncertainties [5], the time series modeling approach is quite favorable. Time-series methods help establish a valid prediction model when there is insufficient data to explain the relationship between the dependent and independent variables [6]. Especially, ARIMA [5] and Artificial Neural Network (ANN) models [7] took overwhelming attention because of their capability to work with noisy, complex, and incomplete data.

Table 1 Approaches employed to predict the impact of COVID-19

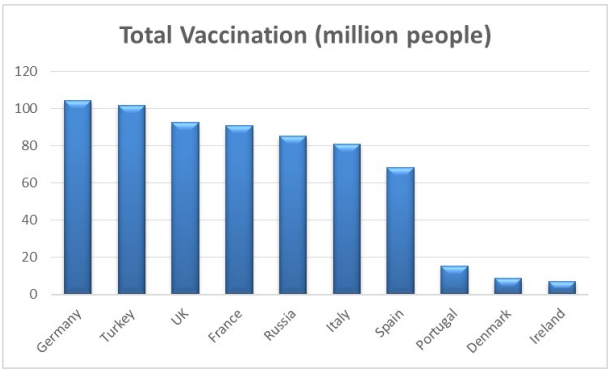
Source	Prediction method/approach	Countries	Dependent variables	Independent variable
[14]	SEIR	China	The inflection point, possible ending time, and total infected cases	Confirmed infected cases, recovered cases latent time, quarantine time, and effective reproduction number
[15]	ARIMA and NARNN	India	Infected cases	Confirmed cases
[16]	Several machine learning models	Italy, the US, France, and the UK	Infected cases, the spread of the virus (growth rate), and the mortality rate	Weather variables and census features
[17]	Logistic function, Hill function, Gompertz function, and LSTM	China and Netherlands	Infected cases	Population size, infected case, static traffic network, and dynamic traffic network
[18]	PNN, RBFNN, and GRNN	India	Deaths (mortality rate)	Deaths and confirmed cases
[19]	MLPICA-ANFIS	Hungary	Infected cases and the mortality rate	Odd days and even days' cases, and mortality rates
[20]	ARIMA, and ANN	India	Infected cases	Confirmed cases
[21]	Gompertz model, Logistic model, and ANN	Mexico	Infected cases	Confirmed cases
[22]	RNN-LSTM variants	India	Daily and weekly infected cases	Confirmed cases
[23]	Machine learning techniques (EN, PCR, PLSR, KNN, RT, RF, GBM, ANN)	US metropolitan counties	Daily and cumulative infected cases	Daily and cumulative confirmed cases, county-level demographic, environmental, mobility, and time-series data
[24]	ANN	Pakistan	Deaths, recovered, and infected cases	Date (day number)
[25]	ANFIS improved by FPA-SSA	China	Daily infected cases	Confirmed cases
[26]	ARIMA (with Hannan, Rissanen algorithm)	India, US, Brazil, Russia, and Spain	Total infected cases	Total confirmed cases
[27]	ARIMA	Italy, Spain, and Turkey	Deaths and infected cases	Confirmed cases
[28]	ARIMA, NARNN, and LSTM	Denmark, Belgium, Germany, France, United Kingdom, Finland, Switzerland, and Turkey	Cumulative infected cases	Cumulative confirmed case
[29]	Recursive based model, Boltzmann function-based model, and Beesham's model	Iran and Turkey	Cumulative deaths and infected cases	Cumulative deaths and confirmed cases
[30]	SEIR	Turkey	Infected cases, intensive care needs, hospitalizations, and deaths	Average and age-specific Infection Fatality Ratio (IFR), infection rate, census data, age-specific hospitalization, and intensive care ratios,
[31]	Gene expression programming	India	Deaths and infected cases	Deaths and confirmed cases
[32]	ARIMA, and NARNN	Egypt	Cumulative infected cases	Cumulative confirmed cases
[4]	Bayesian optimization-based LSTM	USA, Switzerland, India, Slovakia, Russia, Uruguay, Greenland, Malta, Denmark, Brazil, Zimbabwe, Japan, Mexico, Germany, and Norway	Deaths, recovered, infected cases, and country risk classification	Time data, country, confirmed cases, recovered cases, deaths, and weather data

(ANFIS: adaptive network-based fuzzy inference system, ANN: artificial neural network, ARIMA: autoregressive integrated moving average, EN: elastic net model, FPA: flower pollination algorithm, GBM: gradient boosted tree models, GRNN: generalized regression neural network, KNN: k-nearest neighbors regression model, LSTM: long-short term memory, MLPICA: multi-layered perceptron-imperialist competitive algorithm, NARNN: nonlinear autoregressive neural network, PCR: principal components regression model, PLSR: partial least squares regression model, PNN: probabilistic neural network, RBFNN: radial basis function neural network, RF: random forests model, RNN: recurrent neural network, RT: regression tree model, SEIR: susceptible exposed infectious recovered model, SSA: salp swarm algorithm,)

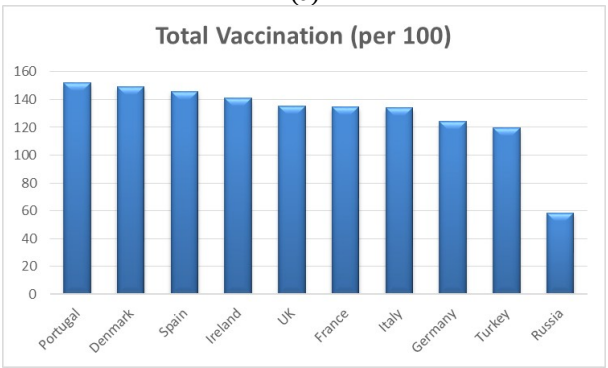
The literature review papers [8-13] are helpful to have more information about the approaches employed to detect and predict the spread of COVID-19. In one of the recent reviews, [33] criticizes state-of-the-art prediction approaches providing generic tools and neglecting major factors, such as social distancing, test rate, and vulnerability issues related to chronic diseases, which significantly affect the prediction accuracy. Ignoring the impact of those factors can deviate the predictions dramatically. In this manner, vaccination is a highly influential factor that affects the spread of the COVID-19 because it is considered a practical way to develop herd immunity [34]. Otherwise, establishing herd immunity via infections may cause a vast number of deaths, as experienced in Sweden [35]. As more people are vaccinated, fewer individuals are expected to be infected, and the severity of the disease is expected to diminish [34]. Predictions before vaccination started are insufficient to accurately predict the spread and impact of outbreaks. Therefore, there is a need to accommodate vaccination impact in the mortality prediction models.

Considering the impact of vaccination, we propose an augmented-ANN approach for COVID-19 mortality prediction for the deaths that occur on-day, lag 1 day, lag 7 day, and lag 14 day. The augmented-ANN utilizes the two most preferred techniques, ANN and ARIMA. Due to the complex and random structure of the epidemic disease, mortality data may include both linear and nonlinear components. Therefore, the augmented-ANN provides a better fit to capture the linear part by ARIMA and the nonlinear part by ANN. Taking into account the suggestion of the previous studies [33, 23], we accommodate the following significant factors that impact the prediction accuracy: (i) severity of the disease, (ii) vaccination policy as a preventive strategy, (iii) time series, (iv) exposure duration in society.

Since the middle of January 2021, Turkey has followed serious vaccination policies successfully, which lead the country to take second place (see Figure 1a) in terms of total vaccination and ninth place in total vaccination per hundred (see Figure 1b) in Europe. Besides, to the best of our knowledge, there is no study on Turkey's mortality prediction after the vaccination has been initiated; instead, mortality prediction research has focused on countries in Europe and China.



(a)



(b)

Figure 1 Vaccination rates and numbers of top 10 European countries [36]

Our study contributes to the literature in the following aspects: i) integration of ARIMA and ANN to capture both linear and nonlinear components of the data, ii) accommodation of the vaccination rates and numbers to better explain the mortality, iii) prediction of mortality in Turkey where the nationwide vaccination program against COVID-19 has been present. The remaining sections are organized as follows. Section 2 presents the precautions taken to fight the epidemic in Turkey and discusses the responsible factors for COVID-19 mortality. Section 3 presents the ARIMA, ANN, and augmented-ANN models. Section 4 demonstrates the results and the advantage of the proposed model and examines the impact of responsible factors on mortality. Section 5 concludes the study by discussing the limitations and suggesting possible future extensions.

2. Data Pretreatment: Turkey Case

In addition to various control measures, such as the general use of face mask and travel bans, stay-at-home policies has been quite strictly applied in Turkey to mitigate the spread of COVID-19. Especially when the rate of increase in COVID-19 cases is high, citizens who violate measures have been subjected to penalties or fines. Various precautions have been taken considering the spread of the virus since March 13, 2020, when the first case was seen in Turkey. These precautions include partial lockdowns (for weekends or specific dates and times), lockdown over 65 and under 18, transition to distance education, full lockdown, etc. By lockdowns, the social distance has been maintained; therefore, the increase in COVID-19 cases and mortality has been reduced.

Policies, such as weekend lockdowns and lockdowns for over the age of 65 and under 20, have restricted community movements. The reflection of these policies can be observed by the Google community mobility report. The report includes the trends on community movements over time by location across six categories: retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential. Google mobility data shows the visitors' relative change in the categorized locations compared to baseline days, representing a typical day. The baseline day is given as the median value over the five weeks from January 3 to February 6, 2020 [37]. The community mobility graph for Turkey between January 14, 2021, and July 03, 2021, as discussed in this study, is shown in Figure 2.

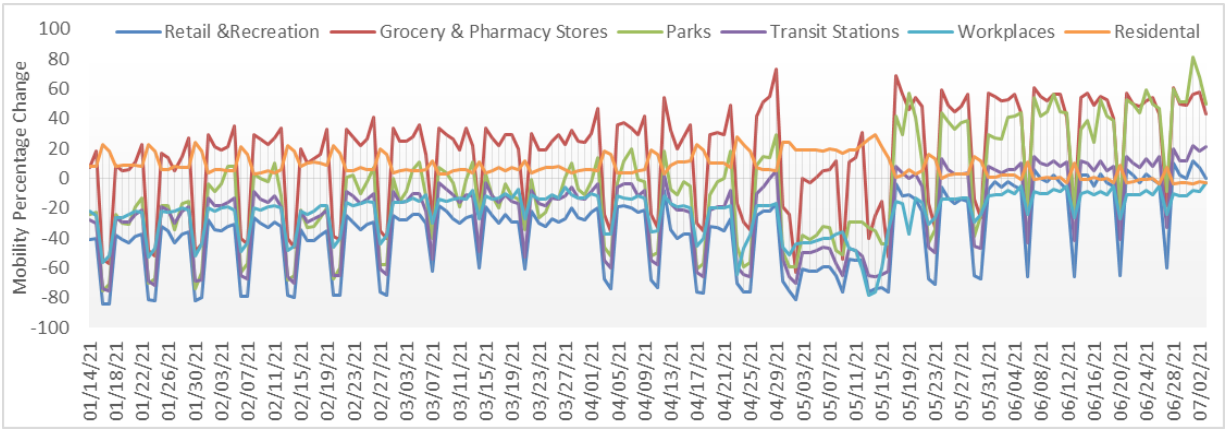


Figure 2 Mobility percentage change in Turkey between January 14 and February 19, 2021

In Figure 2, for example, the weekend lockdowns strictly encourage people to stay at home; thus, generate a decreasing pattern for the weekends followed by an increasing pattern for the weekdays in all the mobility categories except for residential, as expected. Similarly, the effect of full lockdown between April 29 and May 17 on mobility changes can also be seen from the graph in Figure 2.

Considering the COVID-19 data in Turkey and relevant studies from literature, factors affecting mortality can be summed up following four categories: (i) severity of the disease [14, 4], (ii)

vaccination policy as a preventive strategy [34], (iii) exposure duration in society [23], (iv) time series [23]. First, the severity of the disease is associated with the number of severe patients since the number of deaths will be affected by the critical health status of the individuals. The rate of COVID-19 pneumonia is another factor that represents the severity of the disease because pneumonia fills air cells in the lungs with fluid and is deadly [38]. Moreover, the number of patients recovering is incorporated; as more people recover, fewer deaths occur. Additionally, the severity of the disease is also related to late diagnosis. Therefore, the number of COVID-19 PCR tests daily is considered as a factor affecting mortality. We integrated vaccination data as a preventive strategy to the model as a second category. Starting January 14, 2021, to maximize the vaccination's efficiency, Turkey has set a vaccination plan that prioritizes individuals under high risks, such as healthcare workers, older people (above 65 years old), handicapped, and the staff of the nursing homes. Then, gradually the vaccination plan is extended to include all the individuals 18 years old and above. Sinovac and Biontech (since April 12) vaccines are used, and each individual receives two doses approximately 3-4 weeks apart. Three months after the second dose, the third (booster) dose is recommended. In terms of vaccination data, total vaccinations, people fully vaccinated (received two doses), the ratio of total vaccinations to population, and the ratio of people fully vaccinated to population are determined as input variables. Third, we also integrated exposure duration in society to the prediction model by community mobility report retrieved by Google. Finally, because of weekend lockdown policies, mobility has different structures during weekends and weekdays. Thus, the status of the day as weekday or weekend is incorporated as an indicator of the movement trends. All 15 variables considered affecting the mortality in Turkey, the sources, and the types of data were presented in Table 2.

Table 2 Input variables, data sources, periods, and types in ANNs

Sym.	Variable	Source	Data Period	Data Type
X_1	Pneumonia rate in patients	TR Ministry of Health[a]	01/14/2021-07/03/2021	Continuous
X_2	# of severe patients	TR Ministry of Health[a]		Discrete
X_3	# of recoveries	TR Ministry of Health[a]		Discrete
X_4	# of PCR test	TR Ministry of Health[a]		Discrete
X_5	Total vaccinations	Our World in Data[b]		Discrete
X_6	People fully vaccinated	Our World in Data[b]		Discrete
X_7	The ratio of total vaccinations to population	Our World in Data[b]		Continuous
X_8	The ratio of people fully vaccinated to population	Our World in Data[b]		Continuous
X_9	Retail and recreation percent change	Google[c]		Continuous
X_{10}	Grocery and pharmacy percent change	Google[c]		Continuous
X_{11}	Parks percent change	Google[c]		Continuous
X_{12}	Transit stations percent change	Google[c]		Continuous
X_{13}	Workplaces percent change	Google[c]		Continuous
X_{14}	Residential percent change	Google[c]		Continuous
X_{15}	Weekend or not	-		Categorical

[a] Turkish Republic Ministry of Health, 2021., [b] Our World in Data, 2021b., [c] Google, 2021. [39, 40, 37]

Taking into account 15 variables in Table 2, we attempt to predict mortality as an output variable. There is a lag between symptom severity and death [41], and an average of 7- day ICU stay is reported in the literature [42]. Hence, we predicted mortality as the number of deaths in four ways: on day, lag 1 day, lag 7 day, and lag 14 day in ANN models.

3. Methods

Three different prediction models were presented in this section considering the disease situation after vaccination started in Turkey. In the first model, time series analysis was performed with ARIMA in Section 3.1. Next, ANN models with different input variables related to mortality were created in Section 3.2. Last, to improve the prediction performance by handling complex data structures, we proposed an Augmented-ANN approach in Section 3.3.

3.1 ARIMA Model

Among many time series analysis models, ARIMA is preferred in different applications due to its advantages over other stochastic models, such as its superior forecasting capability and ability to provide greater information concerning the time-related change [43]. Particularly, when the explanatory variables describing the prediction variable are limited or unsatisfactory, the ARIMA modeling approach is instrumental since the only data input is the previous data records.

ARIMA models consist of three components: Autoregressive (*AR*), moving average (*MA*), and integrated (*I*). *AR* component indicates that the future values of the data vary over its past values, while *MA* component reflects the purely random regression errors. Also, the last component *I* refers to the differencing period used to forecast future values [6]. These three components are usually represented in the form of ARIMA(*p,d,q*) where *p* refers to the order of *AR* polynomials, *d* is the degree of difference, and *q* is the order of *MA* polynomials [44]. Given in a time series (Z_t), ARIMA models can be in the form of AR(*p*), MA(*q*), ARMA(*p,q*), or ARIMA (*p,d,q*) depending on the values of *p*, *d*, and *q*. Equations (1-3) show the general formula of AR (*p*), MA (*q*), and the ARMA(*p,q*) respectively where ϕ and θ are autoregressive and moving average parameters, α is a constant, and ε is the random errors [5].

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t \tag{1}$$

$$Z_t = \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{2}$$

$$Z_t = \alpha + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \tag{3}$$

In Box-Jenkins methodology [45], building an ARIMA model consists of three stages: (i) model identification, (ii) parameter estimation, and (iii) diagnostic testing. The identification process includes applying transformation algorithms to convert data to the stationary form. Stationary is a pre-requirement for ARIMA models, which refers to a constant statistical characteristic (mean, variance, autocorrelation structure) over time. In this stage, the data should be tested in terms of stationarity. Augmented Dickey-Fuller (ADF) test is a stationarity test used in addition to a time-series plot. When data is in a nonstationary form due to trend or heteroscedasticity, logarithmic transformation, power transformation, or differencing should be applied to stabilize the data. After a tentative model is defined, the following stage includes the estimation of the autoregressive and the moving average parameters presented in a way that the overall error is minimized. The parameters are decided according to the lowest Akaike Information Criterion (AIC), which is presented in Equation (4), where *n* is the data number of observations, *r=p+q+1* and $\hat{\sigma}_a^2$ is a maximum likelihood prediction.

$$AIC_{p,q} = \frac{-2\ln(maximized\ likehood) + 2r}{n} \approx \ln\left(\hat{\sigma}_a^2\right) + r\frac{2}{n} + constant \tag{4}$$

The final stage includes the performance evaluation of the forecasting. Commonly used performance evaluation criteria in the literature are Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) presented as follows in Equations (5-7).

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(z¿¿t - \hat{z}_t)^2¿} \tag{5}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|z_t - \hat{z}_t\right| \tag{6}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{z_t - \hat{z}_t}{z_t}\right| \tag{7}$$

3.2 Artificial Neural Networks

The ANNs are nature-inspired computational modeling tools that simulate human learning to predict future data and make decisions. The ANNs are widely accepted in many disciplines for modeling complex real-world problems [46]. Because first, ANN models have flexible nonlinear function mapping capability that helps approximate the variable in interest with high accuracy [47, 48]. Second, as a data-driven approach, ANN can capture the uncertain hidden relationships between variables [49]. Third, ANNs have an adaptive structure that provides more robust generalizable models for nonstationary environments. However, in contrast to the advantages mentioned earlier, the performance of the ANN compared to the linear model is inconsistent. Some researchers reported that the linear structure of the data without much disruption might be the reason for outperforming linear time series. Therefore, instinctively developing ANN models for any type of data does not guarantee good performance.

In the literature, feed-forward ANNs were preferred mostly for a wide range of prediction applications [50], and also multi-layered perceptron feed-forward ANNs were found favorable for application in predicting deadly infectious disease outbreaks [46]. In this paper, a two-layer feed-forward backpropagation neural network architecture was preferred to predict mortality. The ANN model has 15 input variables as presented in Table 1, two hidden layers, and one output. The same ANN model is replicated for four outputs: as on day, lag 1 day, lag 7 day and lag 14 day. The data set was decomposed into three sets: 80% for training, 10% for testing, and 10% for validation. The tangent sigmoid function was preferred as a neural transfer function in the hidden layer, and the Purelin function is employed in the output layer. The neuron numbers of hidden layers were determined as 8 and 10. Levenberg-Marquardt (LM), Bayesian Regularization (BR), and Scaled Conjugate Gradient (SCG) backpropagation algorithms were deployed as training algorithms.

Preferred training algorithms have different superior features across each other. LM algorithm, which has high accuracy prediction recognition rate in supervised learning networks [51], is known to be a fast and stable convergence training algorithm for moderate-sized feed-forward neural networks (up to several hundred weights) [52]. Based on the success of conjugate gradient methods in solving large-scale unconstrained optimization problems, the SCG algorithm is usually efficient for large multilayer networks, which have more than a thousand or more weights and biases [53]. The BR algorithm can result in good generalization for small or noisy data sets. It is quite challenging to collect all COVID-19 data with complete accuracy. Therefore, assuming that the obtained data is noisy, the BR training algorithm is expected to give better results than other algorithms for the COVID-19 data set. Besides, the BR algorithm updates the squared weights E_W and squared errors E_D by regularization parameters α and β. It limits the size of network parameters by regularization to create a network with better generalization ability. Regularization forces the network to keep the weight and bias values at smaller values; this helps the network reduce overfitting and capture noise [54]. For detailed information on the BR training algorithm, see Hagan et al. [54].

3.3 Augmented-ANN

Real-world problems are seldom pure linear or nonlinear and usually merge both structures, implying that neither a single technique is sufficient for modeling [55]. Besides, due to their complex structure, there has been agreed-upon prediction literature that no universal model acquires distinct patterns evenly. Both theoretical and empirical findings indicated that an effective way to improve their predictive performance is particularly combinations of quite different models [6]. Hence, to adequately capture both the linear and nonlinear components of the data, the ARIMA and ANN models for linear and nonlinear data structure can be combined in one model.

COVID-19 mortality data is complex by nature since neither all the responsible factors nor their impact can be truly captured yet. Various studies utilized different factors, such as age, gender, and chronic disease history, to understand the spread and consequences of the COVID-19 pandemic. However, there are still many unknowns in terms of underlying factors and their impact. Besides, a concern about the nonlinear structure of COVID-19 data was mentioned by [56]. To adequately

acquire different patterns in the data, we propose an Augmented-ANN model. The integration of ANN and ARIMA allows us to provide a better fit for the complex mortality data. The ARIMA model captures the linear relationships between present and prior mortality values, while ANN captures the nonlinear relationships.

The framework of the proposed prediction methodology is presented in Figure 3. The methodology initiates by creating the best performing ARIMA (*p,d,q*) model. The outcomes of the ARIMA model are considered as a baseline prediction that has been built upon the linear correlation structures of the current and past deaths. Then, in addition to other responsible factors in Table 2 (X_1-X_{15}), the ARIMA predictions (X_{16}) are included as input variables to the ANN model for predicting mortality.

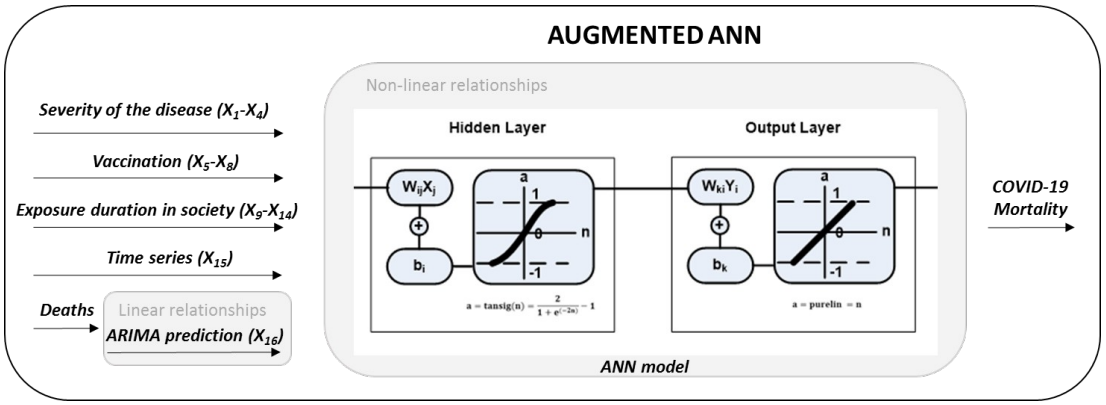


Figure 3 The framework of the proposed Augmented-ANN models

4. Results and Discussion

This section discusses the results of the ARIMA, ANN, and Augmented-ANN models employed to predict the number of deaths under subsections Section 4.1 to 4.3, respectively. The modeling performances of the three models were evaluated in Section 4.4 to demonstrate the appropriateness and effectiveness of Augmented-ANN. Besides, a sensitivity analysis was conducted to identify the impact of each factor on mortality prediction.

4.1 ARIMA Results

The original mortality time series data were tested by Augmented Dickey-Fuller (ADF) test for stationarity. After the stationary assumption was satisfied (p= 0.0269), the parameters of the ARIMA model were determined based on the smallest AIC value that was chosen among the possible combinations of values of 0, 1, and 2 for p and q, respectively. Accordingly, AR (1) model with "1541.503" AIC value was the best model to represent the mortality data. It is worth to mention that AR (1) is a special version of ARIMA model with $p=1$, $d=0$, $q=0$ values. The estimations of the AR(1) model were presented in Table 3.

Table 3 AR (1) model parameters

	Coefficient	Standard error	z-Statistic	p-Value	95% CI for the coefficient	
					Lower limit	Upper Limit
ϕ_1	0.998	0.0056	178.32	0.000	0.9870	1.0089

4.2 ANN Results

The parameters such as type of training algorithms, the number of hidden neurons affect the performance of ANNs. Considering their impact, three training algorithms (LM, BR, and SCG) and two levels of hidden layers (8 and 10) were included in the experimental design. A total of 24 experiments (i.e., 3 training algorithms × 2 levels of hidden layers × 4 outputs) were run based on full factorial design. Performance comparisons of all ANN models in terms of MAE, RMSE, and MAPE were performed. However, due to the conflictive scores of performance parameters (MAE, MSE, and RMSE) across each experiment, the performance parameters were normalized to the range of 0-1 and the overall score is calculated by weighting all parameters equally (w_{MAE}= w_{MSE} = w_{RMSE}= 0.33). The model with the lowest overall score shows the minimum error, thus the best performance. Table 4 presents the results of ANN experiments.

Table 4 ANN models in detail

Training Alg.	# of Hidden Neurons	Output	# of output data	Performance Measures			Overall Score	Correlation Coefficients			
				MAE	RMSE	MAPE		Training	Validation	Test	All
LM	10	On day	170	6.346	8.880	0.049	0.271	.997	.992	.994	.997
LM	8			3.369	6.584	0.027	0.118	1	.994	.985	.998
BR	10			2.362	4.563	0.022	0.043	1	-	.985	.998
BR	8			3.174	6.602	0.027	0.115	1	-	.994	.999
SCG	10			13.553	17.509	0.116	0.759	.983	.987	.981	.983
SCG	8			13.098	16.740	0.103	0.693	.988	.978	.953	.984
LM	10	Lag 1 day	169	4.360	9.486	0.037	0.228	1	.987	.981	.995
LM	8			4.802	8.441	0.042	0.222	.999	.997	.971	.996
BR	10			2.574	9.516	0.016	0.144	1	-	.968	.995
BR	8			3.182	5.109	0.029	0.086	.999	-	.992	.999
SCG	10			9.449	14.163	0.071	0.495	.992	.980	.983	.989
SCG	8			8.698	12.548	0.065	0.432	.993	.985	.991	.991
LM	10	***Lag 7 day***	***163***	4.357	10.712	0.035	0.252	1	.987	.959	.994
LM	8			4.680	8.178	0.036	0.196	.999	.989	.989	.996
BR	***10***			***2.121***	***3.707***	***0.020***	***0.013***	***1***	***-***	***.99***	***1***
BR	8			2.930	4.947	0.024	0.065	1	-	.992	.999
SCG	10			8.804	12.473	0.063	0.422	.992	.988	.989	.991
SCG	8			11.190	14.587	0.087	0.572	.989	.990	.982	.988
LM	10	Lag 14 day	156	6.579	9.584	0.051	0.294	.996	.991	.992	.995
LM	8			3.884	10.603	0.028	0.220	1	.975	.972	.994
BR	10			4.330	8.534	0.027	0.172	.999	-	.980	.996
BR	8			4.351	7.932	0.028	0.476	.990	-	.974	.996
SCG	10			10.171	14.049	0.076	0.514	.993	.952	.982	.989
SCG	8			9.710	12.866	0.072	0.471	.990	.992	.992	.991

In Table 4, the best performing ANN model was indicated as italic and bold. Among all outputs, the performance of the lag 7 day mortality prediction (MAE= 2.121, RMSE= 3.707, MAPE= 0.02, overall score= 0.013) was the highest. Our finding complies with the literature stating the presence of several days lag between severity of disease and the number of deaths [41, 42]. Additionally, it was also observed that the BR algorithm succeeded a better fit in comparison to other training algorithms for all the outcomes. In contrast, the performance of the SCG algorithm was the worst. This might be because of the BR algorithm's superiority for modeling the noisy structure of COVID-19 mortality data. Moreover, it was observed that a higher number of neurons were effective to improve the prediction accuracy, and 10 hidden neurons outperformed 8 hidden neurons for on-day, lag 7 day, and lag 14 day predicted death cases.

4.3 Augmented-ANN Results

This subsection reports the results of the Augmented-ANN in which the ARIMA predictions were included as an input to strengthen ANNs. Because the performance of the Augmented-ANN also depends on the training algorithm and the number of hidden neurons, as mentioned in Section 4.2, we

repeated the same set of experimental designs for the augmented-ANN models. The performance of each different Augmented-ANN model was presented in Table 5.

Table 5 Augmented- ANN models in detail

Training Alg.	# of Hidden Neurons	Output	# of output data	Performance Measures			Overal l Score	Correlation Coefficients			
				MAE	RMSE	MAPE		Training	Validation	Test	All
LM	10	On day	170	3.862	7.622	0.032	0.448	1	.997	.983	.997
LM	8			3.775	7.070	0.033	0.433	.999	.996	.981	.997
BR	10			0.938	3.150	0.007	0.119	1	-	.995	.999
BR	8			2.638	5.259	0.023	0.304	1	-	.979	.999
SCG	10			7.062	10.264	0.053	0.719	.994	.999	.989	.994
SCG	8			7.989	11.148	0.060	0.806	.983	.993	.994	.993
LM	10	***Lag 1 day***	***169***	0.811	1.138	0.006	0.060	1	1	1	1
LM	8			0.488	0.800	0.004	0.030	1	1	1	1
BR	***10***			***0.17***	***0.542***	***0.001***	***0.000***	***1***	***-***	***1***	***1***
BR	8			0.271	0.598	0.002	0.008	1	-	1	1
SCG	10			5.323	6.810	0.043	0.528	.998	.998	.994	.997
SCG	8			2.711	3.600	0.024	0.269	.999	.999	.998	.999
LM	10	Lag 7 day	163	3.946	8.739	0.030	0.471	.999	.993	.962	.996
LM	8			3.974	7.974	0.030	0.452	1	.992	.981	.997
BR	10			3.073	10.982	0.025	0.472	1	-	.940	.993
BR	8			3.706	11.591	0.023	0.500	1	-	.954	.992
SCG	10			9.638	13.550	0.071	0.973	.993	.985	.973	.990
SCG	8			9.961	13.297	0.074	0.994	.989	.994	.992	.990
LM	10	Lag 14 day	156	5.563	12.402	0.045	0.685	.999	.949	.974	.991
LM	8			5.109	10.625	0.033	0.571	.999	.984	.970	.994
BR	10			1.367	5.548	0.007	0.196	1	-	.982	.998
BR	8			2.984	11.810	0.023	0.481	1	-	.935	.992
SCG	10			9.042	12.361	0.066	0.898	.993	.991	.976	.991
SCG	8			9.384	12.685	0.070	0.936	.992	.982	.994	.991

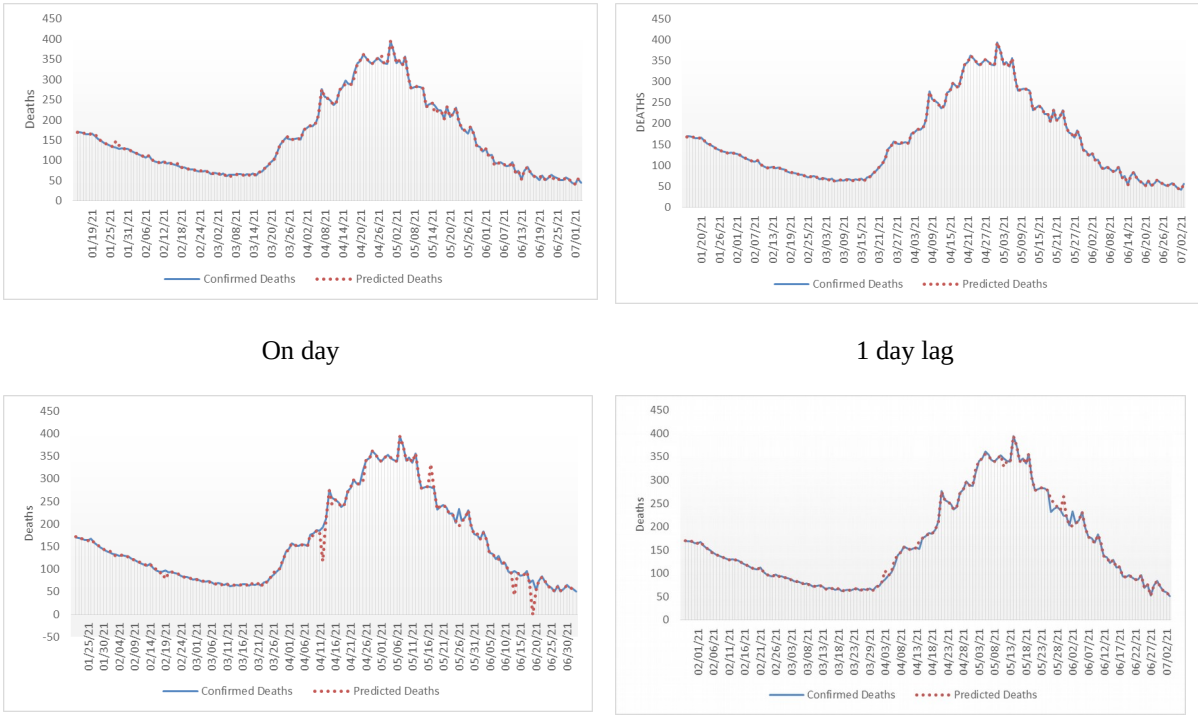
In Table 5, the highest prediction accuracy was accomplished for lag 1 day death prediction with 0.171 MAE, 0.542 RMSE, 0.001 MAPE, and 0.000 overall score. The best performing Augmented-ANN model was succeeded with BR training algorithm and 10 hidden neurons.

4.4 Performance Comparison

The performance of the ARIMA, ANN, and Augmented-ANN are compared in Table 6. It is important to mention that in Table 6 ARIMA model – AR(1) – presents the results of time series and solely considers the linear relationship between current and previous COVID-19 deaths. The ANN row reports the performance of the ANN, considering the impact of different inputs on mortality prediction. Finally, the performance of the Augmented-ANN is presented in the last row. Based on the results, the lowest MAE, RMSE, and MAPE indicated that the Augmented-ANN predicts mortality with the lowest error. Confirmed deaths and predictions of on-day, lag 1 day, lag 7 day, and lag 14 day are presented in Figure 4 for Augmented-ANN.

Table 6 Comparison of the best performing models

Methods	MAE	RMSE	MAPE
AR(1)	8.670	13.545	0.064
ANN	2.121	3.707	0.020
Augmented-ANN	0.171	0.542	0.001



On day 1 day lag

7 day lag 14 day lag

Figure 4 Confirmed and predicted deaths for best performing Augmented-ANN model

4.5 Impact of model parameters on prediction performance

The prediction performance of the model intends to improve as the number of variables increase. However, the question is to identify which input variables have the strongest association with mortality. In this section, we performed a sensitivity analysis to discuss the impact of the input variables on the prediction performance of the Augmented-ANN model. Four new Augmented-ANN models were created considering the categories of the input variables, namely severity of the disease variables (X_1-X_4), vaccination policy as a preventive strategy variables (X_5-X_8), exposure duration in society variables (X_9-X_{14}), and time series variable (X_{15}). Each category's impact was explored by dropping the associated variables while remaining the other three categories in the prediction model. The procedure is repeated for all possible combinations of input variables (categorized in four subsets) to examine the weakest predictors for mortality. The results of the four models were presented in Table 7 in addition to the base model so that their influence on mortality could be discussed.

Table 7 Impact of each category on explaining the mortality for Augmented-ANN model

Impact of category	**Variables added**	**Variables dropped**	**MAE**	**MAPE**	**RMSE**	**Overall score**	**Coefficient of determination (R^2)**
Base model	X_1-X_{16}	-	0.171	0.001	0.542	0.236	0.999
Time series	X_1-X_{14}, X_{16}	X_{15}	0.287	0.002	0.472	0.529	1.000
Severity of disease	X_5-X_{16}	X_1-X_4	0.267	0.002	0.570	0.724	0.999
Vaccination policy	X_1-X_4, X_9-X_{16}	X_5-X_8	0.278	0.002	0.565	0.787	0.999
Exposure duration in society	X_1-X_8, X_{15}-X_{16}	X_9-X_{14}	0.338	0.003	0.560	0.963	0.999

According to the results in Table 7, the R^2- values of all models are above 0.999, which indicates that more than 99.9% of the variation can be explained in COVID-19 mortality with the variables considered. The base model in which all input variables were taken into account outperforms the other models (MAE=0.171, MAPE=0.001, RMSE=0.542, and overall score= 0.236). The second-best model was observed with a MAE of 0.287, a MAPE of 0.002, a RMSE of 0.472, and an overall score of 0.529 after dropping the time series variable. This remarked that the time series had the lowest impact

on model performance, i.e., prediction accuracy. Thus, dropping associated variables resulted in a minimal change in MAE, MAPE, RMSE, and R^2. The performance of the model was the worst when the variables of exposure duration in the society category were removed. Additionally, dropping vaccination policy variables accounted for higher errors. These two categories (exposure duration in society and vaccination policy) have higher impact than the severity of disease category in mortality prediction. Maintaining social distance, applying stay-at-home orders to limit community movement, and getting vaccinated appear to be a safe solution for community health.

5. Conclusion and Limitations

Despite the massive amount of studies available to predict impacts of COVID-19, the research for improving the effectiveness of the prediction has been the focus of the research because of the pandemic's enormous effect on society, the healthcare system, and the economy. Prediction models with increased accuracy assist authorities in taking the complete picture of adverse effects; thus, helping them prompt proactive strategies to fight the pandemic. In this regard, we presented ARIMA, ANN, and augmented –ANN models to predict the number of on-day, lag 1 day, lag 7 day, and lag 14 day COVID-19 deaths in Turkey, considering the data after vaccination has started. ARIMA model encapsulated the linear relationships between current and past data. ANN model accommodated four groups of responsible factors, namely (i) severity of the disease, (ii) vaccination policy, (iii) time series, and (iv) exposure duration in society in the prediction model. Finally, augmented-ANN integrated ARIMA and ANN models to incorporate the ARIMA predictions as a baseline to the established ANN model. Integrating two methods improved prediction accuracy; thus, augmented-ANN achieved the best performance, followed by ANN and ARIMA. The dominance of the ANN models (pure ANN and augmented-ANN) may be attributed to the nonlinear data structure of the COVID-19 mortality.

In order to examine the impact of the responsible factors on mortality, we set up an experimental design and compared the performance of the models by either adding or dropping the responsible categories. Our experiments revealed that the basic model that includes all four categories outperforms the other models. The mobility (integrated through exposure duration in society) has the highest impact on prediction performance so it should not be ignored in the model. Additionally, vaccination policy has the second-highest impact on mortality prediction. On the other hand, dropping time series variables accounted for a minimal change in the error, stating the lowest impact on prediction performance.

Since the accuracy of prediction models plays a crucial role in adopting preventive measures, it is vital to exploit the one with the desirable precision. The results of this research can assist authorities in effectively planning their resources such as ICU bed and staff planning, developing and implementing a shutdown policy that limits the public movement. The model can also be used as a simulation tool to investigate the impact of different regularizations, such as restricting the internal movement, closing nonessential shops, etc.

Even though our study's novelties in terms of incorporating vaccination into the prediction model and integrating ARIMA with ANN to boost accuracy, there are some limitations of the study. First, we could not employ the daily COVID-19 cases in the prediction model since it was not clear whether the daily numbers published by the health ministry of Turkey indicate the number of new "patients" or "cases". Second, the model is based on the current stay-at-home regularizations. The change on these orders will directly influence mobility data, which has the highest impact on mortality. Nevertheless, the prediction model can be updated according to the changes in mobility to provide more precise predictions. In future work, we aim to investigate the effect of various movement restrictions on COVID-19 mortality under the influence of vaccination.

Conflict of Interest Statement

The authors declare no competing interests.

Author’s Contributions

All authors contributed equally to the study.

References

[1] WHO, "World Health Organization COVID-19 Dashboard," 2021. [Online] Available: https://covid19.who.int/ [Accessed: 02-Feb-2022]

[2] A. Hernandez-Matamoros, H. Fujita, T. Hayashi, and H. Perez-Meana, "Forecasting of COVID-19 per regions using ARIMA models and polynomial functions," *Appl. Soft Comp*., vol. 96, 106610, 2020.

[3] S. Zhang, M. Diao, W. Yu, L. Pei, Z. Lin, and D. Chen, " International Journal of Infectious Diseases Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: a data-driven analysis," *Int. J. Infect. Dis.*, vol. 93, pp. 201–204, 2020.

[4] R. Pal, A. A. Sekh, S. Kar, and D. K. Prasad, " Neural network based country wise risk prediction of COVID-19," *Appl. Sci.*, vol. 10, no. 18, 6448, 2020.

[5] Z. Ceylan, "Estimation of COVID-19 prevalence in Italy, Spain, and France," *Sci. Total Environ*., vol. 729, 138817, 2020.

[6] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting," *Appl. Soft Comp*., vol. 11, no. 2, pp. 2664-2675, 2011.

[7] A. Mollalo, K. M. Rivera, and B. Vahedi, " Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States," *Int. J. Environ. Res. Public Health*, vol. 17, no. 12, 4204, 2020.

[8] I. E. Agbehadji, B. O. Awuzie, A. B. Ngowi, and R. C. Millham, "Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing," *Int. J. Environ. Res. and Public Health*, vol. 17, no. 5, 5330, 2020.

[9] F. N Khan, A. A. Khanam, A. Ramlal, and S. Ahmad, A review on predictive systems and data models for COVID-19. In Computational Intelligence Methods in COVID-19: Surveillance, Prevention, Prediction and Diagnosis. *Springer*, Singapore, pp. 123-164, 2021.

[10] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons & Fractals*, vol. 139, 110059, 2020.

[11] Y. Mohamadou, A. Halidou, P. T. Kapen, "A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19," *Appl. Intell.*, vol. 50, no. 1, pp. 3913-3925, 2020.

[12] I. Rahimi, F. Chen, A. H. Gandomi, " A review on COVID-19 forecasting models," *Neural Comput. and Applic*., pp. 1-11, 2021.

[13] H. Swapnarekha, H. S. Behera, J. Nayak, and B. Naik, " Role of intelligent computing in COVID-19 prognosis: A state-of-the-art review," *Chaos, Solitons & Fractals*, vol. 138, 109947, 2020.

[14] L. Peng, W. Yang, D. Zhang, C. Zhuge, and L. Hong, " Epidemic analysis of COVID-19 in China by dynamical modeling," *arXiv* preprint arXiv:2002.06563, 2020.

[15] F. M. Khan and R. Gupta, "ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India," *J. Saf. Sci. Resilience*, vol. 1, no. 1, pp. 12-18, 2020.

[16] Z. Malki, E. S. Atlam, A. E. Hassanien, G. Dagnew, M. A. Elhosseini, and I. Gad, "Association between weather data and COVID-19 pandemic predicting mortality rate: Machine learning approaches," *Chaos, Solitons and Fractals*, vol. 138, 110137, 2020.

[17] M. A. Achterberg, B. Prasse, L. Ma, S. Trajanovski, M. Kitsak, and P. Van Mieghem, "Comparing the accuracy of several network-based COVID-19 prediction algorithms," *Int. J. Forecasting*, In Press.

[18] S. Dhamodharavadhani, R. Rathipriya, and J. M. Chatterjee, "COVID-19 mortality rate prediction for India using statistical neural network models," *Front. Public Health*, vol. 8, 2020.

[19] G. Pinter, I. Felde, A. Mosavi, P. Ghamisi, R. Gloaguen, "COVID-19 pandemic prediction for Hungary; a hybrid machine learning approach," *Mathematics*, vol. 8, no. 6, 890, 2020.

[20] L. Moftakhar, S.E.I.F., Mozhgan, and M. S. Safe, "Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting

models," *Iranian Journal of Public Health*, 49, 2020.
[21] O. Torrealba-Rodriguez, R. A. Conde-Gutiérrez, and A. L. Hernández-Javier, "Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models," *Chaos, Solitons and Fractals,* 138, 109946, 2020.
[22] 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 and Fractals*, vol. 139, 110017, 2020.
[23] C. P. Kuo and J. S. Fu, "Evaluating the impact of mobility on COVID-19 pandemic with machine learning hybrid predictions," *Sci. Total Environ*., vol. 758, 144151, 2021.
[24] I. Ahmad and S. M. Asad, "Predictions of coronavirus COVID-19 distinct cases in Pakistan through an artificial neural network," *Epidemiology and Infection*, vol. 148, 2020.
[25] M. A. Al-qaness, A.A. Ewees, H. Fan, and M. Abd El Aziz, "Optimization method for forecasting confirmed cases of COVID-19 in China," *J. Clin. Med.*, vol. 9, no. 3, 674, 2020.
[26] A. K. Sahai, N. Rath, V. Sood, and M. P. Singh, "ARIMA modelling & forecasting of COVID-19 in top five affected countries," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews,* vol. 14, no. 5, pp. 1419-1427, 2020.
[27] L. Bayyurt and B. Bayyurt, "Forecasting of COVID-19 cases and deaths using ARIMA models," *Medrxiv*, 2020.
[28] I. 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 and Fractals*, 138, 110015, 2020.
[29] M. Niazkar, T. G. Eryılmaz, H. R. Niazkar, and Y. A. Türkkan, "Assessment of Three Mathematical Prediction Models for Forecasting the COVID-19 Outbreak in Iran and Turkey," *Comp. Math. Methods in Med.*, 2020.
[30] S. Arslan, M. Y. Ozdemir, and A. Ucar, "Nowcasting and Forecasting the Spread of COVID-19 and Healthcare Demand in Turkey, a Modeling Study," *Front. Public Health*, vol. 8, 2020.
[31] R. Salgotra, M. Gandomi, and A. H. Gandomi, "Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming," *Chaos, Solitons & Fractals*, vol. 138, 109945, 2020.
[32] A. I. Saba and A. H. Elsheikh, "Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks," Process Saf. Environ. *Protection*, vol. 141, pp. 1-8, 2020.
[33] K. C. Santosh, "COVID-19 prediction models and unexploited data," *J. Med. Sys*., vol. 44, no. 170, 2020.
[34] O. Sharma, A. A. Sultan, H. Ding, and C. R. Triggle, "A Review of the Progress and Challenges of Developing a Vaccine for COVID-19," *Frontiers in Immunology*, vol. 11, 2413, 2020.
[35] F. Jung, V. Krieger, F. T. Hufert, and J. H. Küpper, "Herd immunity or suppression strategy to combat COVID-19," *Clinical Hemorheology and Microcirculation*, pp. 1-5, 2020.
[36] Our World in Data, "Statistics and Research: Coronavirus (COVID-19) Vaccinations," 2021. [Online] Available: https://ourworldindata.org/covid-vaccinations [Accessed: 09-Sep-2021]
[37]Google, "Community Mobility Reports," 2021. [Online] Available: https://support.google.com/covid19-mobility/answer/9824897?hl=en [Accessed: 20-Jul-2021]
[38] F. Pan, et al., " Factors associated with death outcome in patients with severe coronavirus disease-19 (COVID-19): A case-control study," *Int. J. Med. Sci*., vol. 17, no. 9, 1281, 2020.
[39] Turkish Republic Ministry of Health, "COVID-19 Information Platform," 2021. [Online] Available: https://covid19.saglik.gov.tr/TR-66935/genel-koronavirus-tablosu.html [Accessed: 20-Jul-2021]
[40] Our World in Data, "COVID-19 Vaccinations Public Data: Turkey," 2021b. [Online] Available: https://github.com/owid/covid-19-data/blob/master/public/data/vaccinations/country_data/Turkey.csv [Accessed: 20-Jul-2021]
[41] P. Chrusciel and S. Szybka, "On the lag between deaths and infections in the first phase of the COVID-19 pandemic," *medRxiv*, 2021.
[42] G. Grasselli, et al., 1Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy," *Jama*, vol. 323, no. 16, pp. 1574-1581,

2020.

[43] S. Morid, and V. Smakhtin, "Drought forecasting using artificial neural networks and time series of drought indices," *Int. J. Climatol.*, vol. 27, no. 15, pp. 2103-2111, 2007.

[44] X. Li, C. Zhang, B. Zhang, and K. Liu, "A comparative time series analysis and modeling of aerosols in the contiguous United States and China," *Sci. Total Environ*., vol. 690, pp. 799-811, 2019.

[45] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time series analysis: forecasting and control. *John Wiley & Sons*, 2015.

[46] M. D. Philemon, Z. Ismail, and J. Dare, "A review of epidemic forecasting using artificial neural networks," *Int. J. Epidemiol*. Res., vol. 6, no. 3, pp. 132-143, 2019.

[47] G. Cybenko, "Approximations by superpositions of a sigmoidal function," *Math. Control Signal Systems*, vol. 2, pp. 303–314, 1989.

[48] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feed forward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[49] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: the state of the art," *Int. J. Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[50] R. M. Rizk-Allah, and A. E. Hassanien, "COVID-19 forecasting based on an improved interior search algorithm and multi-layer feed forward neural network," *arXiv* preprint arXiv:2004.05960, 2020.

[51] A. A. Suratgar, M. B. Tavakoli, A. Hoseinabadi, "Modified Levenberg–Marquardt method for neural networks training," *World Acad. Sci. Eng. Technol*., vol. 6, no. 1, pp. 46-48, 2005.

[52] H. Yu, and B. M. Wilamowski, "Levenberg-Marquardt training," *Industrial Electronics Handbook*, vol. 5, no. 12, 1, 2011.

[53] N. Andrei, "Scaled conjugate gradient algorithms for unconstrained optimization," *Comput. Optim. Appl*., vol. 38, no. 3, pp. 401-416, 2007.

[54] M. T. Hagan, H. B. Demuth, and M. Beale, Neural network design. PWS Publishing Co., 1997.

[55] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159-175, 2003.

[56] 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. Inf. and Public Health*, vol. 13, no. 7, pp. 914-919, 2020.