Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Application of machine learning in the prediction of COVID-19 daily new cases: A Scoping Review

Soudeh Ghafouri-Fard¹, Hossein Mohammad-Rahimi², Parisa Motie², Mohammad A. S. Minabi³, Mohammad Taheri⁴*, Saeedeh Nateghinia⁵*

1. Department of Medical Genetics, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
2. Dental Research Center, Research Institute of Dental Science, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
3. Sirjan University of Technology, Kerman, Iran.
4. Urology and Nephrology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
5. Skull Base Research Center, Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Corresponding authors: Mohammad Taheri, Saeedeh Nateghinia

Mohammad_823@yahoo.com and s.nateghinia@sbmui.ac.ir
Abstract

COVID-19 has produced a global pandemic affecting all over of the world. Prediction of the rate of COVID-19 spread and modeling of its course have critical impact on both health system and policy makers. Indeed, policy making depends on judgments formed by the prediction models to propose new strategies and to measure the efficiency of the imposed policies. Based on the nonlinear and complex nature of this disorder and difficulties in estimation of virus transmission features using traditional epidemic models, artificial intelligence methods have been applied for prediction of its spread. Based on the importance of machine and deep learning approaches in the estimation of COVID-19 spreading trend, in the present study, we review studies which used these strategies to predict the number of new cases of COVID-19. Adaptive neuro-fuzzy inference system, long short-term memory, recurrent neural network and multilayer perceptron are among the mostly used strategies in this regard. We compared the performance of several machine learning methods in prediction of COVID-19 spread. Root means squared error (RMSE), mean absolute error (MAE), $R^2$ coefficient of determination ($R^2$), and mean absolute percentage error (MAPE) parameters were selected as performance measures for comparison of the accuracy of models. $R^2$ values have ranged from 0.64 to 1 for artificial neural network (ANN) and Bidirectional long short-term memory (LSTM), respectively. Adaptive neuro-fuzzy inference system (ANFIS), Autoregressive Integrated Moving Average (ARIMA) and Multilayer perceptron (MLP) have also have $R^2$ values near 1. ARIMA and LSTM had the highest MAPE values. Collectively, these models are capable of identification of learning parameters that affect dissimilarities in COVID-19 spread across various regions or populations, combining numerous intervention methods and implementing what-if scenarios by integrating data
from diseases having analogous trends with COVID-19. Therefore, application of these methods would help in precise policy making to design the most appropriate interventions and avoid non-efficient restrictions.

**Key words:** COVID-19, machine learning, artificial intelligence, Spread, Global Pandemic
Introduction

The novel coronavirus disease initiated in the late 2019 (COVID-19) is resulted from the infection with the severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) (1). Since late 2019, it has spread globally, leading to a persistent pandemic. COVID-19 spread is dependent on inter-individual close contacts and transmission of breath droplets. Prediction of the rate of COVID-19 spread and modeling of its course have critical impact not only for health systems but also for policy makers. In fact, policy making relies on discernments formed by prediction models to propose new strategies and to measure the efficiency of the imposed policies. Based on the nonlinear and complex nature of this disorder (2) application of artificial intelligence methods is an appropriate alternative to traditional epidemic models for prediction of its spread. Although some traditional epidemic models such as Susceptible-Exposed-Infective-Recovery has been used for prediction of epidemic course (3), these methods have some limitations. For instance, the validity of the Susceptible-Exposed-Infective-Recovery model relies on precise appraisal of virus transmission features including the basic reproductive quantity $R_0$ as well as incubation and infectious periods which are rather difficult to be estimated in real contexts (4). Figure 1 shows the role of artificial intelligence approaches for prediction of COVID-19 spread.

Machine learning methods usually use data sequences retrieved over a period of time as inputs to predict course of COVID-19 epidemic. Several strategies have been implemented for prediction of COVID-19 spread. Among the applied strategies is the Long short-term memory (LSTM) model. For instance, Multilayer perceptron (MLP) has also been applied for modeling of COVID-19 spread. This method has facilitated prediction of the highest number of persons who are affected by COVID-19, the highest number of
people who recovered, and the highest number of mortalities per place in each time division (5). LSTM with the Natural language processing (NLP) module has been used to assess the infection frequency and enhance the predictive accuracy of the model (6). LSTM can efficiently improve gradient explosion and gradient disappearance in the course of the training process by presenting the constant error carousel unit (6). LSTM is superior to the traditional Recurrent neural network (RNN) in term of its good enactment in apprehending the long-term dependency of sequences, thus being appropriate for the categorization, processing, and forecasting the long sequence data (7). Based on the importance of machine and deep learning methods in the prediction of COVID-19 spreading trend, in the current study, we reviewed studies which used these strategies to envisage the number of new cases of COVID-19. The research question was: “What are the applications of machine learning systems and their performances in the prediction of COVID-19 daily new cases?”. In the current study we were looking for publications that evaluate the performance of any machine learning or deep learning approaches based on the research question inclusion and exclusion criteria.

The following parameters were extracted: Root means squared error (RMSE), Mean absolute error (MAE), $R^2$ coefficient of determination ($R^2$), and Mean absolute percentage error (MAPE). These parameters are the main parameters which are applied to assess the error rates of forecasting and performance of the model in regression analysis. MAPE is calculated based on percentage errors.

**Materials and Methods**

We used PRISMA Scoping review guidelines and checklist.

*Protocol*
Reporting this scoping review is based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping reviews (8).

**Exclusion criteria**

1) Studies that did not report or evaluate their prediction regarding the daily confirmed cases or cumulative number of confirmed cases.

2) Studies that did not report at least one of the Root means squared error (RMSE), Mean absolute error (MAE), R² coefficient of determination (R²), and Mean absolute percentage error (MAPE) in their measurements.

**Information sources and search**

An electronic search was conducted in PubMed, Google Scholar, Scopus, Embase, arXiv, and medRxiv for finding the relevant literature from January 2020, to June 2021. Different combinations of the following keywords were used in the search procedure: “machine learning”, “deep learning”, “neural network”, “artificial intelligence”, “Covid-19”, “incidence”, “prevalence”, “spread*”, “new cases”, “predict*”, and “forecast*”.

**Selection of sources of evidence**

Duplicate studies were removed. Studies that were cited within the retrieved papers were reviewed for finding any missing studies. For identifying the proper journal papers and conference proceedings, our team members screened the title and abstracts based on inclusion and exclusion criteria independently. Finally, considering the inclusion and exclusion criteria, investigators identified the eligible publications in this stage independently. Figure 2 illustrates the flowchart of the protocol of systematic literature review.
Data charting process

Two investigators were responsible for extracting the data, separately. The charting process was followed by consensus to resolve any disagreements.

Data items

For the selected studies, the following data have been extracted: regions (e.g., countries, states, etc), data source, data structure, machine learning model and model performance including RMSE, MAE, $R^2$, MAPE (on the basis of the best model). These performance measures were selected, since they are the most common performance measurement among the selected studies.
Records identified through database searching (n=1842)

Additional records identified through other sources (n=0)

Records after duplicates removed (n=991)

Records screened (n=991)  Records excluded (n=841)

Full-text articles assessed for eligibility (n=150)  Full-text articles excluded, with reasons (n=75)

Full-text articles added by reviewing references (n=10)
Figure 2. The flowchart of the protocol of systematic literature review.

Results

Several artificial intelligence strategies have been used for prediction of COVID-19 spread using different models (Figure 3).
Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS is a type of artificial neural network being founded on Takagi–Sugeno fuzzy inference system. Architecture of ANFIS has five layers, namely fuzzification layer, the layer which generates the firing strengths for the rules (rule layer), the layer that normalizes the computed firing strengths, the layer which receives as input the normalized values and the consequence parameters, and the layer
that returns the final output (9). Al-Qanes et al. (10) have designed an upgraded kind of the ANFIS model to estimate the quantity of infected persons in four countries, namely Italy, Iran, Korea, and the USA. Their model has been founded on a novel nature-inspired optimizer, namely the marine predators algorithm (MPA). This algorithm has optimized the ANFIS variables, increasing its predicting performance. They have shown superiority of the MPA-ANFIS method to previously suggested predicting models in terms of better values for RMSE, MAE, MAPE, and $R^2$ (10). In another study, ANFIS was boosted using an improved flower pollination algorithm (FPA) by using the salp swarm algorithm (SSA). The suggested FPASSA-ANFIS model was then appraised using the official data retrieved from WHO site. Moreover, the accuracy of the suggested model was then appraised using two distinct datasets of weekly influenza cases (11). Alsayed et al. (12) have predicted the epidemic peak in Malaysia using the Susceptible-Exposed-Infectious-Recovered (SEIR) model. They have also used the ANFIS model short-time prediction of the amount of infected individuals. They have also demonstrated the impact of interventions on postponing the epidemic peak. Moreover, they have suggested that extension of the intervention period might decrease the epidemic magnitude at the peak. This study has reported RMSE, $R^2$ and MAPE values as 46.87, 0.9973 and 2.79, respectively (12). Thus, this study has reported the best performance measurements using this method. Behnood et al. (13) have used an integration of the virus optimization algorithm (VOA) and ANFIS to appraise the impact of numerous climate-associated parameters and population density on COVID-19 spread. They have demonstrated the remarkable influence of population density on the performance of their designed models, emphasizing on the prominence of social distancing in decreasing COVID-19 infection rate and spread. RMSE, MAE and $R^2$ values have been reported to be 22.47, 7.33 and 0.83, respectively (13).
Autoregressive Integrated Moving Average (ARIMA)

As a type of univariate regression analysis method, ARIMA forecasts upcoming values according to differences between values instead of actual figures. As a generalization of an autoregressive moving average (ARMA) model, ARIMA is fitted to time series data for better understanding of the data or predicting upcoming points in these series. Alzahrani et al. (14) have used ARIMA model to predict the estimated daily amounts of COVID-19 persons in Saudi Arabia. They have reported the superiority of ARIMA to Autoregressive Model, Moving Average and an integration ARMA and ARIMA. Using ARIMA, they have reported RMSE, MAE, $R^2$ and MAPE values as 21.17, 14.93, 0.99 and 2.16, respectively (14). Chakraborty et al. (15) have proposed a hybrid strategy founded on ARIMA model and Wavelet-based predictive model which could produce short-term predictions of the amount of daily cases for Canada, France, India, South Korea, and the UK. The obtained RMSE and MAE values ranged from 55.25-631.91 and 24-306.78 in different regions (15). Khan et al. (16) have used an ARIMA model for forecasting daily cases of COVID-19 in India. They selected the appropriate model according to the Bayesian Information Criteria parameters and the total maximum $R^2$ value of 0.95 (16). The best performance measurements using ARIMA has been reported in the study conducted by Adiga et al. (MAPE=999.1) (17).

Multilayer perceptron (MLP)

MLP is a type of feedforward artificial neural network (ANN). This model has three layers of nodes, namely an input layer, a hidden layer and an output layer. With the exception of the input node, other nodes are neurons that use a nonlinear activation function. MLP uses the backpropagation supervised learning method for training (18). Car et al. (5) have used a freely accessible time-series dataset for design of their model. They have used this dataset in training an MLP model. The finest designed models had 4 hidden layers with
4 neurons in each. This model had appropriate measures in the prediction of the deceased and confirmed cases, but it had low robustness for recovered patients (5). Pinter et al. (19) have used the hybrid machine learning strategies of ANFIS and MLP-imperialist competitive algorithm (MLP-ICA) for prediction of time series of COVID-19 cases and mortality amount. Short-term observation has confirmed the accuracy of the proposed model. Authors have suggested that the model keeps its exactness providing no substantial interruption happens (19).

**Long short-term memory (LSTM)**

LSTM is an artificial recurrent neural network (RNN) method utilized as a deep learning strategy. In contrast to standard feedforward neural networks, this model ensures feedback connection. In addition to processing single data points, LSTM can process complete sequences of data (20). Aora et al. (21) have used RNN-related LSTM variants on an Indian dataset of COVID-19 patients to forecast the amount of positive cases. Based on the lowest error rate, LSTM model was selected for prediction of daily and weekly new COVID-19 cases with approximate error rates of 3% and 8%, respectively. Subsequently, they classified Indian states into different zones based on the extent of positive cases and daily escalation for recognition of COVID-19 hot-spots (21). Fokas et al. (21) have applied a bidirectional LSTM network to yield a robust generalization of RNNs. This method has been used for predication of new cases of COVID-19 in Italy, Spain, France, Germany, USA and Sweden (22).

**Other models**

Yadav et al. (23) have used six regression analysis based methods including quadratic, third degree, fourth degree, fifth degree, sixth degree, and exponential polynomial for prediction of COVID-19 cases with the sixth degree polynomial regression method
representing the best model for prediction of short-term new cases (23). Kim et al. (24) have used geographic hierarchy to create Hi-COVIDNet according to a neural network with two-level machineries that are based on data collected from country-level and continent-level systems. This method apprehends the multifaceted relations among distant countries and relates their particular infection risk to the target country (24). Table S1 shows the application of machine learning methods for prediction of COVID-19 spread.

Table S1. Summary of the results of studies which used machine learning methods for prediction of COVID-19 spread.

| Author, Year | Region(s) | Data Source | Data Structure | Best Algorithm/Model Structure(s) | Performance Measurements (on the best model) | Ref |
|--------------|-----------|-------------|----------------|----------------------------------|----------------------------------------------|-----|
| Abbasimehr, H., et al. 2021 | Dataset 1: US, United Kingdom, Turkey, Spain, Mexico, Italy, Iran, Germany, France, Belgium | Humanitarian Data Exchange (Dataset 1 and 2 were used for short-term and long-term prediction, respectively) | Daily confirmed cases from 20 January to 1 August 2020 | Attention-based model using Bayesian Optimizer | RMSE: 2715.12, MAE: - , $R^2$: - , MAPE: 0.2157 | (25) |
| Dataset 2: US, Brazil, India, Russia, South Africa, Mexico, Peru, Chile, Colombia, Iran | | | Daily confirmed cases from 20 January to 3 August 2020 | LSTM using Bayesian Optimizer | RMSE: 25292.337, MAE: - , $R^2$: - , MAPE: 2.6606 | |
| Adiga, A., et al. 2021 | Maricopa AZ | Johns Hopkins University Center for Systems Science and Engineering | The 7-day smoothed version of confirmed cases | SEIR (not machine learning) | - , - , - | (17) |
| Los Angeles CA | | | | Spatial Autoregressive ARIMA | - , - , - | 1678.7 |
| San Bernardino | | | | | | 999.1 |
| Country          | Region                        | Dataset                          | Daily confirmed cases                                  | Model            | MAE     | MSE     | R      | RMSE   | (Ref) |
|------------------|-------------------------------|----------------------------------|--------------------------------------------------------|------------------|---------|---------|--------|--------|--------|
| Brazil           | Kings NY                     | Official WHO data                | Daily confirmed cases from 26 March to 1 June 2020     | LSTM             | 19.432  | 14.273  | -      | 0.3117 | 2085.4 |
| Russia           |                               |                                  |                                                        | MPA + ANFIS      | 493     | 379     | -      | 0.03223|        |
| Al-Qaness, M. A., A., 2021 |                               | Iran                             | Daily confirmed cases in the USA, Korea, Iran, and Italy from 22 January to 7 April 2020 | MPA-ANFIS        | 70.93   | 60.31   | 96.48% | 0.696  |        |
| Al-Qaness, M. A., A., et al., 2020 |                               | USA                              | Daily confirmed cases in the USA, Korea, Iran, and Italy from 22 January to 7 April 2020 | MPA-ANFIS        | 5465.66 | 3951.94 | 98.59% | 2.734  |        |
| Al-Qaness, M. A., A., et al., 2020 |                               | Korea                            | Daily confirmed cases from 21 January to 18 February 2020 | FPASSA-ANFIS     | 5779    | 4271    | 0.9645 | 4.79   |        |
| Al-Qaness, M. A., A., et al., 2020 |                               | Italy                            | Daily confirmed cases from 14 March to 14 May 2020    | MLP              | 191.27  | -       | 0.999  |        |        |
| Al-Qaness, M. A., A., et al., 2020 |                               | Iran                             | Daily confirmed cases from 22 March to 5 April 2020   | MLP              | 55.52   | -       | 0.995  |        |        |
| Alzahrani, S., et al., 2020 | Saudi Arabia                  | Ministry of Health website       | Daily confirmed cases in Saudi Arabia                  | ARIMA            | 21.17   | 14.93   | 0.99   | 2.16   |        |
| Ardabili, S. F., et al., 2020 | Italy                         | Worldometer website              | Daily confirmed cases over 30 days                     | MLP              | 391.8   | -       | 0.998  |        |        |
| Ardabili, S. F., et al., 2020 | Germany                       |                                  |                                                        | MLP              | 22.1    | -       | 0.999  |        |        |
| Ardabili, S. F., et al., 2020 | Iran                          |                                  |                                                        | MLP              | 2318.22 | 3.22    | -      |        |        |
| Arora, P., et al, 2020 | India                         | Johns Hopkins University Center for Systems Science and Engineering dataset | Daily confirmed cases from 14 March to 14 May 2020    | Stacked LSTM     | 3.22    | -       | -      | -      |        |
| Arora, P., et al. 2020 | India                         | Ministry of Health and Family Welfare (Government of) | State-wise daily confirmed cases from 14 March to 14 May 2020 | Bi-directional LSTM | -     | -       | -      | 3.22   |        |
| Country          | Dataset Description                                                                 | Cumulative daily confirmed cases | LSTM   | GRU   | Seasonal ARIMA |
|------------------|-------------------------------------------------------------------------------------|---------------------------------|--------|-------|----------------|
| USA              | Johns Hopkins University Center for Systems Science and Engineering dataset         | 1.69E + 06                      | 1.33E + 05 | 6.03E + 03 | - | - | - |
| Brazil           |                                                     | 4.44E + 03                      | 1.22E + 03 | 1.33E + 02 | - | - | - |
| South Africa     |                                                     | 1.33E + 05                      | - | - | - |
| Peru             |                                                     | - | - | - |
| Chile            |                                                     | 4.21E + 02                      | 9.37E + 02 | - | - | - |
| Iran             |                                                     | 9.00E + 02                      | - | - | - |
| Mexico           |                                                     | 2.80E + 03                      | - | - | - |
| UK               |                                                     | 2.34E + 03                      | 7.80E + 03 | - | - | - |
| Russia           |                                                     | 2.38E + 02                      | 1.16E + 02 | - | - | - |
| India            |                                                     | 1.21E + 04                      | 7.15E + 03 | - | - | - |
| Bangladesh       |                                                     | 1.48E+04                        | 1.17E+04 | 5.46 | - |
| Brazil           |                                                     | 1.30E+03                        | 2.92E+03 | 2.57 | - |
| Chile            |                                                     | 2.34E+03                        | 2.13E+03 | 3.12 | - |
| Colombia         |                                                     | 9.07E+03                        | 7.80E+03 | 2.53 | - |
| India            |                                                     | 2.38E+02                        | 1.16E+02 | 0.13 | - |
| Iran             |                                                     | 1.21E+04                        | 7.15E+03 | 0.87 | - |
| Italy            |                                                     | 4.65E+03                        | 3.80E+03 | 0.095 | - |
| Mexico           |                                                     | 1.94E+03                        | 1.79E+03 | 0.923 | - |
| Pakistan         |                                                     | 1.09E+04                        | 8.82E+03 | 3.24 | - |
| Peru             |                                                     | 3.07E+02                        | 1.58E+02 | 0.078 | - |
| Russia           |                                                     | 2.41E+03                        | 1.99E+03 | 0.77 | - |
| Saudi Arabia     |                                                     | 1.40E+04                        | 1.23E+04 | 97 | - |
| Spain            |                                                     | 1.15E+04                        | 1.08E+04 | 0.08 | - |
| UK               |                                                     | 2.00E−03                        | 2.00E−02 | 0.988 | - |
| USA              |                                                     | 1.45E+01                        | 1.25E+01 | 0.80 | - |
| USA              |                                                     | 2.46E+04                        | 1.72E+04 | 1.41 | - |

ArunKumar, K. E., et al. 2021a

ArunKumar, K. E., et al. 2021b

Ayyoubzadeh, S. M., et al., 2020
| Authors | Country | Dataset | Time Period | Method | MAPE | RMSE | R^2 | Source |
|---------|---------|---------|-------------|--------|------|------|-----|--------|
| Bedi, P., et al. 2020 | India | covid19india.org website | Daily confirmed cases 30 January to 6 September 2020 | LSTM | - | - | - | 0.03 | (31) |
| Behnood, A, et al., 2020 | USA | USAFacts Website | Daily confirmed cases in 1657 counties | ANFIS-VOA-II | 22.4744 | 7.3337 | 0.8339 | - | (13) |
| Borghi, P. H., et al., 2021 | Global (top 30 countries with the highest number of daily new cases) | Johns Hopkins University Center for Systems Science and Engineering dataset | Daily confirmed cases till 11 May 2020 | ANN | 2.082E+03 | 3.718E+06 | - | - | (32) |
| Car, Z., et al., 2020 | 406 locations | Johns Hopkins University Center for Systems Science and Engineering dataset | Daily confirmed, recovered, and deceased patients in a certain location (defined by the name of location, latitude, and longitude), from 22 January to 12 March 2020 | MLP | - | - | 0.98599 | - | (5) |
| Chakraborty, T., et al., 2020 | USA | Our World in Data Website | Daily confirmed cases | TARNN | 721.5658 | 468.6335 | - | - | (33) |
| Brazil | | | | | 178.0458 | 90.2053 | | | |
| India | | | | | 201.0696 | 128.7718 | | | |
| Russia | | | | | 443.4280 | 202.6083 | | | |
| South Africa | | | | | 243.5067 | 160.3598 | | | |
| Mexico | | | | | 24.4335 | 15.1298 | | | |
| Spain | | | | | 136.3910 | 87.5449 | | | |
| Iran | | | | | 319.9160 | 182.8744 | | | |
| Chakraborty, T., et al., 2020 | Canada | Our World in Data Website | Daily confirmed cases | Hybrid ARIMA-WBF Model | 149.60 | 40.05 | | | (15) |
| France | | | | | 631.91 | 306.78 | | | |
| India | | | | | 55.25 | 24.00 | | | |
| Country/Region | Source/Website | Data | Model | Predictions (Short-term) | Predictions (Long-term) | Other Notes |
|---------------|----------------|------|-------|-------------------------|-----------------------|-------------|
| **South Korea** | Our World in Data Website and a simulated dataset | Daily confirmed, and deceased patients from 1 January to 2 April, 2020 | Bidirectional LSTM | 8,649.154, 7,130.149 | 1 | - | (34) |
| **UK** | China, Italy, Spain, Germany, Iran, Switzerland, South Korea, Belgium, Netherlands, Austria, Singapore, Malaysia, France, Australia, USA, UK and Portugal | | | | | | |
| **CHATTERJEE, A., ET AL. 2020** | Worldwide | DataHub-Novel Coronavirus 2019-Dataset | Daily confirmed, recovered, and deceased patients from 22 January to 29 June, 2020 | ARIMA | 0.1517, 0.12044 | - | 0.0091 | (35) |
| **Chaurasia, V., ET AL., 2020** | Canada | Johns Hopkins University and Canadian Health authority | Daily confirmed, recovered, and deceased cases from 22 January to 31 March, 2020 | LSTM | 34.83 (short-term predictions) 45.70 (long-term predictions) | - | - | - | (36) |
| **CHIMMULA, V., K., R., ET AL., 2020** | Bangladesh | Worldometer website | Daily confirmed cases 10 April to 30 June 2020 | LSTM | 6.55 | - | - | 4.51 | (37) |
| **CHOWDHURY, A. A., ET AL. 2021** | Brazil | Brasil.io portal | Daily confirmed cases till 6 June 2020 | Linear Regression | 11.42% | - | - | - | (38) |
| **DASILVA, C. C., ET AL. 2021** | | | | | | | | |
| Study                                      | Country         | Source                  | Methodology                     | Cumulative Cases | R        | RMSE     |
|-------------------------------------------|-----------------|-------------------------|---------------------------------|------------------|---------|---------|
| de Souza, D. G. S., et al., 2020          | Amapa (A state in Brazil) | Health surveillance secretary of Amapa | Cumulative confirmed cases from 20 March to 31 August, 2020 | Holt-Winters | 162     | 0.98    | 0.34   | (39) |
| Dharani, N. P., et al. 2021               | India           | Kaggle website          | Daily confirmed cases 30 January to 21 May 2020 | Linear Regression | 223.89  | 157.78  | 1.0    | -    | (40) |
| Doe, S. W., et al., 2020                  | USA             | Johns Hopkins University confirmed cases data for US counties | Daily confirmed cases from 22 January to 31 May, 2020 and latitude, and longitude of each county | CLEIR-Net | 264.33  | -       | -      | -    | (41) |
| Fokas, A. S., et al., 2020                | Italy           | European CDC website    | Daily confirmed cases          | Bidirectional LSTM network | 538     | 0.9999  |        |
|                                          | Spain           |                         |                                 |                  | 1022    | 0.9998  |        |
|                                          | France          |                         |                                 |                  | 821     | 0.9997  |        |
|                                          | Germany         |                         |                                 |                  | 1128    | 0.9997  |        |
|                                          | USA             |                         |                                 |                  | 10754   | 0.9996  |        |
|                                          | Sweden          |                         |                                 |                  | 178     | 0.9997  |        |
| Ganiny, S., et al., 2020                  | India           | Worldometer website, India's Ministry of Health and Family Welfare, the Covindia website | Daily confirmed, recovered, and deceased cases from 1 March to 25 July, 2020 | ARIMA     | 457.61  | 330.79  | 0.99998| 0.2471| (42) |
| Ghany, K. K. A., et al. 2021              | Saudi Arabia    | Johns Hopkins University Center for Systems Science and Engineering | Daily confirmed cases 22 January to 24 July 2020 | LSTM      | 1768.35 | 375.21  |        |
|                                          | Qatar           |                         |                                 |                  | 735.21  | 730.53  | -      | -    | (43) |
|                                          | Oman            |                         |                                 |                  | 456.90  | -       | -      | -    |     |
|                                          | Kuwait          |                         |                                 |                  | 446.44  | -       | -      | -    |     |
|                                          | UAE             |                         |                                 |                  |         |         |        |      |     |
| Author(s)          | Country  | Dataset                                                                 | Data Description                                                                 | Method    | Validation   | Cross-Validation | Cross-Validation Error | Notes   |
|--------------------|----------|--------------------------------------------------------------------------|----------------------------------------------------------------------------------|-----------|--------------|-------------------|------------------------|---------|
| Ghazaly, N., M., et al., 2020 | Bahrain | WHO situation reports                                                   | Daily cases and deaths from 21 January to 2 April, 2020                          | NAR       | -            | -                 | -                      | 2.6521  |
| Guo, Q. and He, Z., 2021 | Global  | Official WHO data                                                        | Daily confirmed cases 21 January to 11 November 2020                            | ANN       | 3102.9       | 2090.6            | 0.9683                 | -       |
| Hasan, K. T., et al. 2021 | Bangladesh | Official WHO data and the Institute of Epidemiology, Disease Control and Research of Bangladesh | Daily confirmed cases till 3 August 2020 + Government control and people's compliance data + Information of how many people will be in contact with an infected person outside of their home when they move out | LSTM      | 10,368.318   | -                | 1                      | 5.96    |
| Hazarika, B. B., et al., 2020 | Brazil  | Our World in Data Website                                                | Cumulative number of confirmed cases from 11 April to 10 July, 2020             | WCRVFL    | 0.00323      | 0.99975           | -                      | 0.99999 |
|                    | India    |                                                                         |                                                                                  |           | 0.00147      | 0.99996           | -                      | -       |
|                    | Peru     |                                                                         |                                                                                  |           | 0.00197      | 0.99986           | -                      | -       |
|                    | Russia   |                                                                         |                                                                                  |           | 0.00029      | 0.99999           | -                      | -       |
|                    | USA      |                                                                         |                                                                                  |           | 0.00524      | 0.99999           | -                      | -       |
| Hawas, M, et al., 2020 | Brazil  | Johns Hopkins University Center for Systems Science and Engineering dataset | Daily cases from 22 January, 2020, to several dates                              | RNN       | -            | -                | 0.665                  | -       |
| Heni¸ B., et al., 2020 | Phase I: 79 countries phase II: China, European CDC website              | Phase I: Daily confirmed cases from the date of the first confirmed cases to 13 | LSTM      | -            | -                | Phase I: 0.999         | Phase II: 0.996 | -       |
| Country  | Region/Website                                                                 | Time Frame                                                                 | Method      | Accuracy   | R²          | Notes                  |
|----------|-------------------------------------------------------------------------------|---------------------------------------------------------------------------|-------------|------------|-------------|------------------------|
| South Korea, France, Germany, Iran, Iraq, United Kingdom, Italy, Japan, Singapore, Spain, Thailand | March, 2020 Phase II: Daily recovered cases from the date of the first confirmed cases to 19 March, 2020 | Stacked LSTM | 593.764    | 0.95       | 1.76%      | (50)                  |
| Bangladesh | Johns Hopkins University’s GitHub repository                                   | Daily confirmed, recovered, and deceased cases from 8 March to 13 June, 2020 | Stacked LSTM | 593.764    | 0.95       | 1.76%      | (50)                  |
| Global (42 countries) | Official WHO data, the World Bank website, the Weather Underground website | Daily confirmed cases 22 January to 24 March 2020 + infrastructure, environment, policies, and infection-related independent variables | Random forest | -         | 0.543 to 0.992 (country-wise) | - | (51) |
| Denmark | European CDC                                                                     | Daily confirmed cases from the date of the first confirmed cases to 3 May 2020 | LSTM        | 54.5398    | 0.999963324 | 0.5033    | (52)                  |
| Belgium |                                                                              |                                                                            | LSTM        | 274.0248   | 0.999967249 | 0.5422    | (52)                  |
| Germany |                                                                              |                                                                            | LSTM        | 569.4791   | 0.999980729 | 0.3083    | (52)                  |
| France |                                                                              |                                                                            | LSTM        | 455.7141   | 0.999987339 | 0.3155    | (52)                  |
| United Kingdom |                                                                              |                                                                            | LSTM        | 5482.2361  | 0.998923082 | 2.5025    | (52)                  |
| Finland |                                                                              |                                                                            | LSTM        | 49.4966    | 0.999897188 | 0.8492    | (52)                  |
| Switzerland |                                                                              |                                                                            | LSTM        | 55.8685    | 0.999996376 | 0.1640    | (52)                  |
| Turkey |                                                                              |                                                                            | LSTM        | 640.26257  | 0.999971407 | 0.4823    | (52)                  |
| Jharkhand | COVID19 INDIA Website                                                         | Daily confirmed cases and the cumulative confirmed cases up to 24 May, 2020 | SAIUQR      | 5.23       | -          | -         | (53)                  |
| Gujarat |                                                                              |                                                                            | SAIUQR      | 51.82      | -          | -         | (53)                  |
| Andhra Pradesh |                                                                              |                                                                            | SAIUQR      | 13.47      | -          | -         | (53)                  |
| Chandigarh |                                                                              |                                                                            | SAIUQR      | 3.82       | -          | -         | (53)                  |
| South Korea | Johns Hopkins University + Searched keyword                                   | Daily confirmed cases                                               | Hi-COVIDNet(A Customized) | May 6–12: | -  | -  | (24) |

Kasilingam, D., et al. 2021

Hridoy, A. E., et al., 2020

Kirbas, I., et al., 2020

Khajanchi, S., et al., 2020

Kim, M., et al., 2020
| Study            | Country       | Data Source                                                                 | Methodology                                                                 | Model Parameters | Confidence Interval | Additional Notes |
|------------------|---------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|------------------|---------------------|------------------|
| Khan, F., et al., 2020 | India         | Ministry of Health and Family Welfare, COVID19 INDIA Website               | Daily confirmed cases up to 4 April, 2020                                | ARIMA            | -                   | 0.95             |
| Kufel, T., et al., 2020 | 32 European countries | Johns Hopkins University Center for Systems Science and Engineering dataset | Daily confirmed, recovered, and deceased cases for 14 days in each country | ARIMA            | -                   | From 0.5577 to 1.0000 (depends on country and dates) |
| Kumar, S., et al., 2020 | India         | COVID19 INDIA Website                                                       | Cumulative number of confirmed, recovered, and deceased cases            | ARIMA            | 641.732 (new cases), 705.293 | 0.987 (new cases) |
| Kumar, N., et al., 2020 | Worldwide     | Johns Hopkins University Center for Systems Science and Engineering dataset | Daily confirmed, recovered, and deceased cases up to 20 May, 2020         | ARIMA            | 36992.53, 33109.68, 9774.06, 13078.23, 5853.29, 13901.04 | 34932.99, 31899.89, 9683.45, 12910.06, 5780.87, 13702.61 | 2.523, 15.635, 7.361, 12.78, 10.574, 9.808 |
| Country   | Daily confirmed cases | Daily recovered cases | Daily deaths | Amount of testing | Lockdown presence and its severity | Method     | Accuracy | RMSE   | MAE   | SE   |
|-----------|-----------------------|-----------------------|--------------|------------------|------------------------------------|------------|----------|--------|--------|------|
| Russia    | 3212.50               | 2376.69               | 5.103        |                  |                                    | ANN        | 0.9969   |        |        |      |
| Iran      | 4496.75               | 4213.14               | 4.933        |                  |                                    |            |          |        |        |      |
| UK        | 91.12                 | 78.19                 | 8.311        |                  |                                    |            |          |        |        |      |
| Turkey    | 4333.57               | 4242.09               | 4.321        |                  |                                    |            |          |        |        |      |
| India     | 1066.65               | 721.17                | 2.911        |                  |                                    |            |          |        |        |      |
| Liu, Z., et al., 2020 | Tencent news and Baidu migration websites | Daily confirmed cases | ANN | 0.9969 | - | (57) |
| Majhi, R., et al. 2020 | China (for training), India (for validation) | NA | Daily confirmed and recovered cases, daily deaths, Amount of testing, Lockdown presence and its severity | Random forest | - | - | 0.02 | (58) |
| Malki, Z., et al. 2021 | Johns Hopkins University, WHO and Worldometer official website | Daily confirmed cases | Decision Tree | 0.277 | 0.126 | (59) |
| Mishra, P., et al., 2020 | WHO daily situation reports | Daily new cases from 17 March to 1 July, 2020 | ANN | 0.3822 | 0.2312 | (60) |
| Moftakhar, L., et al., 2020 | Iran Ministry of Health and open datasets provided by Johns Hopkins University | Daily new cases from 19 February to 30 March 2020 | ARIMA | 1539.43 | 24.85 | (61) |
| Authors                  | Country       | Source/Website                                                                 | Type of Data                                                                 | Model Used                     | Parameters                                                                 | Notes |
|-------------------------|---------------|-------------------------------------------------------------------------------|------------------------------------------------------------------------------|--------------------------------|----------------------------------------------------------------------------|-------|
| Melin, P, et al., 2020  | Mexico        | Government of Mexico website                                                  | Daily confirmed and deceased cases                                            | Modular Neural Network with Fuzzy | From 8.6153 to 1554.0302 (Depends on the state) | (62)  |
| Mollalo, A. et al., 2020| USA           | USAFacts website                                                              | Cumulative number of confirmed cases from 22 January to 30 April 2020        | ANN                            | 0.722409, 0.355843, 0.645481, -                                           | (63)  |
| Nabi, K. N., et al., 2021| Brazil        | Johns Hopkins University Center for Systems Science and Engineering dataset   | Daily confirmed cases till 18 November 2020                                  | CNN                            | 0.086, 0.014, 0.048                                                     | 6.94  |
|                          |               |                                                                                |                                                                               |                                |                                                                            | 0.85  |
|                          | Russia        |                                                                                |                                                                               |                                |                                                                            |       |
|                          | UK            |                                                                                |                                                                               |                                |                                                                            | 3.75  |
| Neeraj, et al., 2020     | Canada        | Johns Hopkins University Center for Systems Science and Engineering dataset   | Daily confirmed, recovered, and deceased cases for 209 days in each country  | Attention-based encoder-decoder | 12.46, 209.23, 163.78, 281.03, -                                         |       |
|                          | Italy         |                                                                                |                                                                               |                                |                                                                            | 0.11  |
|                          | France        |                                                                                |                                                                               |                                |                                                                            | 1.71  |
|                          | Spain         |                                                                                |                                                                               |                                |                                                                            | 1.21  |
|                          |               |                                                                                |                                                                               |                                |                                                                            | 2.11  |
| Nikolopoulos, K, et al., 2020 | Germany, India, Singapore, UK, USA | Johns Hopkins university dataset, “covid19-report” website, Mayer Brown’s COVID-19 Global Travel Restrictions, the world population review, World Life Expectancy website, World Bank website | Daily confirmed, recovered and deceased cases, climate information, travel restrictions and curfews data, populations information, lung diseases data, coronary heart diseases data, diabetes prevalence data, GDP spent on healthcare data | Naive-d 0.1 for weekly prediction, GARCH(1,1) model with SGED for daily prediction | 1.0015 (Scaled), 0.2064 (Scaled) | 1.0022 |
| Pal, R., et al., 2020    | USA           | Johns Hopkins university dataset, Dark                                         | Daily confirmed, recovered and deceased cases from 22 January                | Shallow LSTM using used a Bayesian | 1103.5, -                                                                  |       |
| Author(s) | Source | Sky website | Data description | Optimization framework | Prediction Performance |
|-----------|--------|-------------|------------------|------------------------|-----------------------|
| Papastefanopoulos, V., et al., 2020 | US, Spain, Italy, UK, France, Germany, Russia, Turkey, Brazil, Iran | Novel Corona Virus 2019 Dataset and population-by-country dataset on Kaggle website | Daily confirmed, recovered and deceased cases as of 4 May 2020 | TBAT | 0.009873, 0.029295, 0.005810, 0.004310, 0.007003, 0.00389, 0.002193, 0.001946, 0.005621, 0.000425 |
| Peng, Y., et al. 2021 | Worldwide (215 countries) | Official WHO data, Google Trends service | Daily confirmed cases 10 January to 16 August 2020 + Infoveillance data (Google Trends (search volume of 28 COVID19-related features)) | Random forest | 9.27, 5.42, - , - |
| Pereira, I. G., et al., 2020 | USA, Brazil, Italy, Spain, France, UK | Johns Hopkins university dataset, Natalnet's Lab and Brazil ministry of health (just for Brazil), Italy—Official Covid Data Repository, | Daily confirmed and deceased cases | LSTM-SAE | - , - , 0.822 , 84 |
| Perone, G., 2020 | Italy, Russia | Worldometer website | Daily confirmed cases (Italy (February 22–) | ARIMA | 412.79, 283.49, 0.95, 13.039, 606.66, 430.83, 0.98, 11.39 |
| Author, Year | Country | Source | Time Frame | Methodology | Cases | Recovered | Mortality Rate | Deaths | R2 | Adjusted R2 | (Reference) |
|-------------|---------|--------|------------|-------------|-------|-----------|-------------|--------|----|-----------|------------|
| Perone, G., 2020 | USA | Italian Ministry of Health’s website | April 14, USA (March 9–May 16), Russia (March 22–May 22)) | ARIMA-NNAR | 2,411.6 | 1,631.3 | 0.95 | 9.59 | - | - | (72) |
| Pinter, G., et al., 2020 | Hungary | Worldometer website | Daily confirmed, and deceased cases from 4 March to 19 April, 2020 | MPL-ICA | 167.88 | - | - | - | - | - | (19) |
| Quintero, Y., et al. 2021 | Colombia | The National Institute of Health for Colombia and the National Administrative Department of Statistics | Daily confirmed cases from March to July 2020 + Socioeconomic data including people over 65, poverty index, total population, people per km², Average age, average morbidity | Gradient boosting regressor | 0.0157 | 0.0045 | 0.8986 | 1.5317 | - | - | (73) |
| Ribeiro, Mhdm, et al., 2020 | Brazil | WHO website | Daily confirmed, and deceased cases from 15 March to 19 April, 2020 | SVR (generally was the best algorithm, stacking-ensemble learning and ARIMA outperformed in some cases) | 18–409 (One day) | 8.5–59.67 (Three days) | 7.83–73.17 (Six days) | 0.87–3.51 (One day) | 1.02–5.63 (Three days) | 0.95–6.90 (Six days) | (74) |
| Rustam, F., et al., 2020 | Afghanistan, Australia, Algeria, Canada | Johns Hopkins University Center for Systems Science and Engineering dataset | Daily confirmed, recovered, and deceased patients | Exponential Smoothing | 16828.58 | 8867.43 | 0.98 | - | - | - | (75) |
| Saba, A. I., et al., | Egypt | The Egyptian Ministry of | Accumulated confirmed cases from 1 March to | Nonlinear Autoregressive | 10.410 | 7.752 | 0.999 | - | - | - | (76) |
| Year   | Data Source                      | Country       | Dataset Details                                                                 | Model                      | R² Value | RMSE Value | AIC Value | Notes |
|--------|----------------------------------|---------------|----------------------------------------------------------------------------------|----------------------------|----------|------------|-----------|-------|
| 2020   | Health Research                 | Worldwide data| Official WHO data + Demographic, socioeconomic, and health sector indicators data | Bidirectional LSTM         | 0.996    | 176.02     | 245.1     | (77)  |
|        |                                  |               | Daily confirmed cases                                                            |                            |          |            |           |       |
| 2021   | Saaid, A. B., et al.            | US            | Johns Hopkins University Center for Systems Science and Engineering dataset       | Hybrid polynomial-Bayesian ridge regression model | 0.996    | 418.36     | 723.75    | (78)  |
|        |                                  | Italy         | Daily confirmed patients                                                          |                            |          |            |           |       |
|        |                                  | Spain         | Daily confirmed patients                                                          |                            |          |            |           |       |
| 2021   | Saqib, M.                       | Assam (a state in India) | Kaggle website Daily confirmed cases till 30 October 2020 | Piecewise linear regression | -        | -          | 0.392     | (79)  |
|        |                                  | India         | “Our World in Data” website                                                       |                            |          |            |           |       |
| 2020   | Shyam Sunder Reddy, K., et al.  | East Midlands region of England | Public Health UK and NHS Digital Daily cumulative confirmed cases and the total number of daily phone calls received at the NHS 111 from 18 March to 19 September, 2020 | MLR-T                      | 19.37    | 14.16      | -         | (81)  |
| 2020   | Shyam Sunder Reddy, K., et al.  | East Midlands region of England | Public Health UK and NHS Digital Daily cumulative confirmed cases and the total number of daily phone calls received at the NHS 111 from 18 March to 19 September, 2020 | MLR-T                      | 19.37    | 14.16      | -         | (81)  |
| 2020   | Torrealba-Rodriguez, O. et al.  | Mexico        | Daily Technical Report" by the Mexican Ministry of Health Daily confirmed cases from 27 February to 8 May, 2020 | ANN                        | -        | -          | 0.9999    | (82)  |
| 2020   | Tuli, S., et al.                | World         | Our World in Data website                                                          | Generalized Inverse Weibull distribution fitting | -        | -          | 0.98      | 49.14 |
|        |                                  | India         | Daily confirmed cases                                                             |                            |          |            | 0.97      | 18.33 |
|        |                                  | USA           | Daily confirmed cases                                                             |                            |          |            | 0.95      | 24.33 |
|        |                                  | UK            | Daily confirmed cases                                                             |                            |          |            | 0.95      | 21.46 |
|        |                                  | Italy         | Daily confirmed cases                                                             |                            |          |            | 0.96      | 14.98 |
| Authors                          | Country/Region | Source Details                                                                 | Key Data                                                                 | Method/Approach                                      | World Daily Confirmed Cases | US Daily Confirmed Cases | UK Daily Confirmed Cases | Italy Daily Confirmed Cases | Notes |
|----------------------------------|----------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------|------------------------------------------------------|----------------------------|--------------------------|--------------------------|--------------------------|-------|
| Tuli, S., et al., 2020           | World          | Our World in Data website, Index Mundi, World Bank, Oxford Government Response Tracker | Daily confirmed cases till 19 May, 2020, Socioeconomic data, Virus Type data, Government Stringency Index | LSTM-based Robust Weibull approach                   | 559.46                     | 0.93                     | 39.29                    | 4.39                     |       |
|                                 | India          |                                                                                |                                                                           |                                                      | 4.39                       | 0.91                     | 22.00                    | 0.89                     |       |
|                                 | USA            |                                                                                |                                                                           |                                                      | 319.37                     | 0.86                     | 26.09                    | 0.86                     |       |
|                                 | UK             |                                                                                |                                                                           |                                                      | 99.60                      | 0.90                     | 14.93                    | 0.90                     |       |
|                                 | Italy          |                                                                                |                                                                           |                                                      | 45.93                      | 0.89                     | 12.78                    | 0.89                     |       |
| Wang, L., et al., 2020           | World          | UVA COVID-19 surveillance dashboard, Johns Hopkins University dataset, Google COVID-19 Aggregated Mobility Research dataset | Weekly confirmed and deceased cases, Case count growth rate, COVID-19 testing data, aggregated relative weekly mobility flows over google users, Flow Reduction Rate of connectivity before and after the pandemic, and Social Distancing Index, All from 7 March to 22 August, 2020 (25 weeks) | Ensemble of RNNs and SEIR | -                          | -                        | -                        | -                        |       |
|                                 | Global (Austria, Brazil, India, Italy, Nigeria, Singapore, the United Kingdom) |                                                                                |                                                                           |                                                      | -                          | -                        | -                        | -                        |       |
|                                 | US-States      |                                                                                |                                                                           |                                                      | -                          | -                        | -                        | -                        |       |
|                                 | US-Counties    |                                                                                |                                                                           |                                                      | -                          | -                        | -                        | -                        |       |
| Yadav, R. S., 2020              | India          | COVID-19 in India Kaggle dataset                                                 | Daily confirmed cases from 1 March to 11 April, 2020                    | Sixth degree polynomial regression analysis         | -                          | -                        | 0.9990                   | -                        |       |
| Zawbaa, H. M., et al. 2021      | China          | Johns Hopkins University Center for Systems Science and Engineering dataset and European Centre for Disease Prevention and Control | Daily confirmed cases 22 January to 13 December 2020                    | ANN                                                  | 1460.57                    | 105.94                   | 486.13                   | 318.29                   |       |
|                                 | Cote d'Ivoire  |                                                                                |                                                                           |                                                      | 486.13                     | 318.29                   | 371.40                   | 952.29                   |       |
|                                 | Kenya          |                                                                                |                                                                           |                                                      | 5632.84                    | 17922.58                 | 77822.38                 | 4398                     |       |
|                                 | Egypt          |                                                                                |                                                                           |                                                      | 77822.38                   | 4398                     | -                        | -                        | (86)  |
|                                 | Algeria        |                                                                                |                                                                           |                                                      | 19722.58                   | -                        | -                        | -                        | (86)  |
|                                 | Japan          |                                                                                |                                                                           |                                                      | 19722.58                   | -                        | -                        | -                        | (86)  |
|                                 | Iran           |                                                                                |                                                                           |                                                      | 19722.58                   | -                        | -                        | -                        | (86)  |
|                                 | Italy          |                                                                                |                                                                           |                                                      | 19722.58                   | -                        | -                        | -                        | (86)  |
|                                 | USA            |                                                                                |                                                                           |                                                      | 19722.58                   | -                        | -                        | -                        | (86)  |
| Zaremba, J., et al.             | Italy          | Johns Hopkins University Center for Systems Science and Engineering dataset and European Centre for Disease Prevention and Control | Daily confirmed cases 22 January to 13 December 2020                    | Bidirectional                                           | 1,041,374                  | 1,033,467                | 4398                     | -                        | (87)  |
### Discussion

#### Synthesis of results

Accurate prediction of the time of outbreak would help in reduction of the effect of COVID-19, permit governments to modify their preventive strategies and plan in advance for the protective steps required. Modeling of COVID-19 spread is particularly important in...
defining its potential future impacts. Artificial intelligence methods are superior to traditional statistical modeling methods in the terms of offering high-quality predictive models (89). These models are capable of identification of learning parameters that affect dissimilarities in COVID-19 spread across various regions or populations, combining numerous intervention methods and implementing what-if scenarios by integrating data from diseases having analogous trends with COVID-19. In the current scoping review, we compared the performance of several machine learning methods in prediction of COVID-19 spread. RMSE, MAE, $R^2$ and MAPE parameters were selected as performance measures for comparison of the accuracy of models. $R^2$ values have ranged from 0.64 to 1 for ANN and Bidirectional LSTM, respectively. ANFIS, ARIMA and MLP have also have $R^2$ values near 1. ARIMA and LSTM had the highest MAPE values. These prediction models could also appraise the impact of climate-associated factors in infection rate or COVID-19 spread facilitating implementation of specific strategies for each condition. Moreover, the data obtained from these models can be used for categorization of county regions and identification of hot spots for COVID-19 to organize region-specific preventive measures. Incorporation of data from health status of affected individuals including general health situation and related risk factors would enhance the accuracy of these models. Most of the proposed models have been effective in short-term forecasting of the COVID-19-related parameters. However, their efficacy in long-term should be validated in further studies.

Modeling of the COVID-19 is practically important in defining the possible upcoming impact of this disorder and artificial intelligence methods have especial situation in this regard. These modeling strategies have implications in disease management by policy makers as they can predict the future course of the pandemic. Moreover, the impact of large-scale screening strategies and application of disease-controlling modalities can be considered in these modeling methods. ARIMA and LSTM have good
performance values in this regard. In fact, ARIMA model is the furthermost extensively used forecasting method for prediction of trends in time series. However, it is not possible to compare the results of these studies, as these methods have not been applied and trained on the same data. Moreover, although artificial intelligence strategies have been promising in prediction of COVID-19 course during the pandemics, COVID-19 continues to be an unknown disease with no historic information to predict its spreading. Therefore, integration of these methods and implementation of the results in larger populations consisting of different ethnicities would help in design of better predictive models.

ARIMA method of time has been used to predict the stability and growth of COVID-19. Recent studies have suggested that the performance of this model can be enhanced or the model can provide more precise data if more numbers of datasets are accessible (90). The model provides results according to the data established by information provided by health organizations. Therefore, prediction may not be completely precise, yet it can confidently be used as a corrective tool (90). Combination of new factors and algorithms with ARIMA can lead of enhancement of accuracy.

Accordingly, Abbasimehr and Paki have proposed three hybrid methods for prediction of COVID-19 time series methods according to conjoining three deep learning models, namely multi-head attention, LSTM, and CNN with the Bayesian optimization algorithm. Their analyses have shown higher performance of deep learning models compared with the benchmark model both for short-term prediction and long-term prediction. Particularly, the mean SMAPE of the best deep learning model has 0.25 and 2.59 for the short-term and long-term predictions, respectively (25).
Deep Neural Networks (DNNs) has also been suggested as method for approximation. This method is an important alternative to estimate the solution of a Partial Differential Equation (91). DNN has been used for detection of COVID-19 based on CT scan and chest X rays (92). Application of unsupervised learning methods in which algorithm training is achieved using unlabeled data is another approach which is less studies in this context. A recent study has used the k-means algorithm to divide the countries into clusters based on the spread of COVID-19 in three time spans (93).

**Summary of evidence**

These forecasts are just built on past trends of COVID-19 spread, so forecast values are not definite. Nevertheless, these predicted estimates of events can assist authorities to establish resource planning for better management of this pandemic. Moreover, these methods can be used for prediction of need for preventive measures in each geographical region, thus helping vaccine manufacturers for designing appropriate plans.

**Limitations**

Impossibility of accurate comparison of methods, lack of consistency between study variables

**References**

1. Ghafouri-Fard S, Noroozi R, Vafaei R, Branicki W, Pośpiech E, Pyrc K, et al. Effects of host genetic variations on response to, susceptibility and severity of respiratory infections. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie. 2020;128:110296.
2. Ivanov D. Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. Transportation research Part E, Logistics and transportation review. 2020;136:101922.

3. Yang M, Ke H, Zhou W. LncRNA RMRP Promotes Cell Proliferation and Invasion Through miR-613/NFAT5 Axis in Non-Small Cell Lung Cancer. OncoTargets and therapy. 2020;13:8941.

4. Sun J, Chen X, Zhang Z, Lai S, Zhao B, Liu H, et al. Forecasting the long-term trend of COVID-19 epidemic using a dynamic model. Scientific reports. 2020;10(1):1-10.

5. Car Z, Baressi Šegota S, Andelić N, Lorencin I, Mrzljak V. Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron. Comput Math Methods Med. 2020;2020:5714714.

6. Zheng N, Du S, Wang J, Zhang H, Cui W, Kang Z, et al. Predicting COVID-19 in China Using Hybrid AI Model. IEEE Trans Cybern. 2020;50(7):2891-904.

7. Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J. LSTM: A search space odyssey. IEEE transactions on neural networks and learning systems. 2016;28(10):2222-32.

8. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Annals of internal medicine. 2018;169(7):467-73.

9. Karaboga D, Kaya E. Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. Artificial Intelligence Review. 2019;52(4):2263-93.
10. Al-Qaness MAA, Ewees AA, Fan H, Abualigah L, Elaziz MA. Marine predators algorithm for forecasting confirmed cases of COVID-19 in Italy, USA, Iran and Korea. International Journal of Environmental Research and Public Health. 2020;17(10).

11. Al-Qaness MA, Ewees AA, Fan H, Abd El Aziz M. Optimization method for forecasting confirmed cases of COVID-19 in China. Journal of Clinical Medicine. 2020;9(3):674.

12. Alsayed A, Sadir H, Kamil R, Sari H. Prediction of epidemic peak and infected cases for COVID-19 disease in Malaysia, 2020. International Journal of Environmental Research and Public Health. 2020;17(11):1-15.

13. Behnood A, Mohammadi Golafshani E, Hosseini SM. Determinants of the infection rate of the COVID-19 in the U.S. using ANFIS and virus optimization algorithm (VOA). Chaos, Solitons and Fractals. 2020;139.

14. Alzahrani SI, Aljamaan IA, Al-Fakih EA. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions, J Infect Public Health. 2020;13(7):914-9.

15. Chakraborty T, Ghosh I. Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis. Chaos, Solitons & Fractals. 2020:109850.

16. Khan FM, Gupta R. ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. Journal of Safety Science and Resilience. 2020;1(1):12-8.

17. Adiga A, Wang L, Hurt B, Peddireddy A, Porebski P, Venkatramanan S, et al. All Models Are Useful: Bayesian Ensembling for Robust High Resolution COVID-19 Forecasting. medRxiv. 2021.
18. Rosenblatt F. Principles of neurodynamics, perceptrons and the theory of brain mechanisms. Cornell Aeronautical Lab Inc Buffalo NY; 1961.

19. Pinter G, Felde I, Mosavi A, Ghamisi P, Gloaguen R. COVID-19 pandemic prediction for Hungary; A hybrid machine learning approach. Mathematics. 2020;8(6).

20. Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation. 1997;9(8):1735-80.

21. Arora P, Kumar H, Panigrahi BK. Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. Chaos Solitons Fractals. 2020;139:110017.

22. Fokas AS, Dikaios N, Kastis GA. Mathematical models and deep learning for predicting the number of individuals reported to be infected with SARS-CoV-2. J R Soc Interface. 2020;17(169):20200494.

23. Yadav RS. Data analysis of COVID-2019 epidemic using machine learning methods: a case study of India. Int J Inf Technol. 2020:1-10.

24. Kim M, Kang J, Kim D, Song H, Min H, Nam Y, et al. Hi-COVIDNet: Deep Learning Approach to Predict Inbound COVID-19 Patients and Case Study in South Korea. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; Virtual Event, CA, USA: Association for Computing Machinery; 2020. p. 3466–73.

25. Abbasimehr H, Paki R. Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. Chaos Solitons Fractals. 2021;142:110511.
26. Al-Qaness MAA, Saba AI, Elsheikh AH, Elaziz MA, Ibrahim RA, Lu S, et al. Efficient artificial intelligence forecasting models for COVID-19 outbreak in Russia and Brazil. Process Saf Environ Prot. 2021;149:399-409.

27. Ardabili SF, MOSAVI A, Ghamisi P, Ferdinand F, Varkonyi-Koczy AR, Reuter U, et al. COVID-19 Outbreak Prediction with Machine Learning. medRxiv. 2020:2020.04.17.20070094.

28. ArunKumar KE, Kalaga DV, Kumar CMS, Kawaji M, Brenza TM. Forecasting of COVID-19 using deep layer Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cells. Chaos Solitons Fractals. 2021;146:110861.

29. ArunKumar KE, Kalaga DV, Sai Kumar CM, Chilkoor G, Kawaji M, Brenza TM. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). Appl Soft Comput. 2021;103:107161.

30. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, SRN K. Predicting COVID-19 Incidence Through Analysis of Google Trends Data in Iran: Data Mining and Deep Learning Pilot Study. JMIR Public Health Surveill. 2020;6(2):e18828.

31. Bedi P, Dhiman S, Gole P, Gupta N, Jindal V. Prediction of COVID-19 Trend in India and Its Four Worst-Affected States Using Modified SEIRD and LSTM Models. SN Comput Sci. 2021;2(3):224.

32. Borghi PH, Zakordonets O, Teixeira JP. A COVID-19 time series forecasting model based on MLP ANN. Procedia Comput Sci. 2021;181:940-7.
33. Chakraborty T, Bhattacharyya A, Pattnaik M. Theta autoregressive neural network model for COVID-19 outbreak predictions. medRxiv. 2020:2020.10.01.20205021.

34. Chatterjee A, Gerdes MW, Martinez SG. Statistical Explorations and Univariate Timeseries Analysis on COVID-19 Datasets to Understand the Trend of Disease Spreading and Death. Sensors (Basel). 2020;20(11).

35. Chaurasia V, Pal S. COVID-19 Pandemic: ARIMA and Regression Model-Based Worldwide Death Cases Predictions. SN Computer Science. 2020;1(5):288.

36. Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos Solitons Fractals. 2020;135:109864.

37. Chowdhury AA, Hasan KT, Hoque KKS. Analysis and Prediction of COVID-19 Pandemic in Bangladesh by Using ANFIS and LSTM Network. Cognit Comput. 2021:1-10.

38. da Silva CC, de Lima CL, da Silva ACG, Silva EL, Marques GS, de Araújo LJB, et al. Covid-19 Dynamic Monitoring and Real-Time Spatio-Temporal Forecasting. Front Public Health. 2021;9:641253.

39. de Souza DGB, Alves Júnior FT, Soma NY. Forecasting COVID-19 cases at the Amazon region: a comparison of classical and machine learning models. bioRxiv. 2020:2020.10.09.332908.

40. Dharani NP, Bojja P, Raja Kumari P. Evaluation of Performance of an LR and SVR models to predict COVID-19 Pandemic. Mater Today Proc. 2021.
41. Doe SW, Seekins TR, Fitzpatrick D, Blanchard D, Sekeh SY. Adaptive County Level COVID-19 Forecast Models: Analysis and Improvement. arXiv preprint arXiv:200612617. 2020.

42. Ganiny S, Nisar O. Mathematical Modeling and a Month Ahead Forecast of the Coronavirus Disease 2019 (COVID-19) Pandemic: An Indian Scenario. medRxiv. 2020:2020.09.10.20192195.

43. Ghany KKA, Zawbaa HM, Sabri HM. COVID-19 prediction using LSTM algorithm: GCC case study. Inform Med Unlocked. 2021;23:100566.

44. Ghazaly NM, Abdel-Fattah MA, Abd El-Aziz AA. Novel coronavirus forecasting model using nonlinear autoregressive artificial neural network. International Journal of Advanced Science and Technology. 2020;29(5 Special Issue):1831-49.

45. Guo Q, He Z. Prediction of the confirmed cases and deaths of global COVID-19 using artificial intelligence. Environ Sci Pollut Res Int. 2021;28(9):11672-82.

46. Hasan KT, Rahman MM, Ahmmed MM, Chowdhury AA, Islam MK. 4P Model for Dynamic Prediction of COVID-19: a Statistical and Machine Learning Approach. Cognit Comput. 2021:1-14.

47. Hazarika BB, Gupta D. Modelling and forecasting of COVID-19 spread using wavelet-coupled random vector functional link networks. Applied Soft Computing. 2020;96:106626.

48. Hawas M. Generated Time-series Prediction Data of COVID-19's Daily Infections in Brazil by Using Recurrent Neural Networks. Data Brief. 2020:106175.
49. Heni B. COVID-19, Bacille Calmette-Guerin (BCG) and tuberculosis: Cases and recovery previsions with deep learning sequence prediction. Ingenierie des Systemes d'Information. 2020;25(2):165-72.

50. Hridoy A-EE, Naim M, Emon NU, Tipo IH, Alam S, Al Mamun A, et al. Forecasting COVID-19 Dynamics and Endpoint in Bangladesh: A Data-driven Approach. medRxiv. 2020;2020.06.26.20140905.

51. Kasilingam D, Sathiya Prabhakaran SP, Rajendran DK, Rajagopal V, Santhosh Kumar T, Soundararaj A. Exploring the growth of COVID-19 cases using exponential modelling across 42 countries and predicting signs of early containment using machine learning. Transbound Emerg Dis. 2021;68(3):1001-18.

52. Kırbaş İ, Sözen A, Tuncer AD, Kazancıoğlu F. Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. Chaos Solitons Fractals. 2020;138:110015.

53. Khajanchi S, Sarkar K. Forecasting the daily and cumulative number of cases for the COVID-19 pandemic in India. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2020;30(7):071101.

54. Kufel T. ARIMA-based forecasting of the dynamics of confirmed Covid-19 cases for selected European countries. Equilibrium Quarterly Journal of Economics and Economic Policy. 2020;15(2):181-204.

55. Kumar S, Agiwal V, Kumar A, Kumar J. Modeling and Prediction of COVID-19 Outbreak in India. 2020.

56. Kumar N, Susan S, editors. COVID-19 Pandemic Prediction using Time Series Forecasting Models. 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT); 2020 1-3 July 2020.
57. Liu Z, Huang S, Lu W, Su Z, Yin X, Liang H, et al. Modeling the trend of coronavirus disease 2019 and restoration of operational capability of metropolitan medical service in China: a machine learning and mathematical model-based analysis. Glob Health Res Policy. 2020;5:20.

58. Majhi R, Thangeda R, Sugasi RP, Kumar N. Analysis and prediction of COVID-19 trajectory: A machine learning approach. J Public Aff. 2020:e2537.

59. Malki Z, Atlam ES, Ewis A, Dagnew G, Ghoneim OA, Mohamed AA, et al. The COVID-19 pandemic: prediction study based on machine learning models. Environ Sci Pollut Res Int. 2021:1-11.

60. Mishra P, Fatih C, Rawat D, Sahu S, Pandey SA, Ray M, et al. Trajectory of COVID-19 Data in India: Investigation and Project Using Artificial Neural Network, Fuzzy Time Series and ARIMA Models. Annual Research & Review in Biology. 2020:46-54.

61. Moftakhar L, Seif M, Safe MS. 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. 2020;49:92-100.

62. Melin P, Monica JC, Sanchez D, Castillo O. Multiple Ensemble Neural Network Models with Fuzzy Response Aggregation for Predicting COVID-19 Time Series: The Case of Mexico. Healthcare (Basel). 2020;8(2).

63. Mollalo A, Rivera KM, Vahedi B. Artificial Neural Network Modeling of Novel Coronavirus (COVID-19) Incidence Rates across the Continental United States. Int J Environ Res Public Health. 2020;17(12).
64. Nabi KN, Tahmid MT, Rafi A, Kader ME, Haider MA. Forecasting COVID-19 cases: A comparative analysis between Recurrent and Convolutional Neural Networks. medRxiv. 2021.

65. Neeraj, Mathew J, Behera RK, Panthakkalakath ZE. A Deep Learning Framework for COVID Outbreak Prediction. ArXiv. 2020;abs/2010.00382.

66. Nikolopoulos K, Punia S, Schäfers A, Tsinopoulos C, Vasilakis C. Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. Eur J Oper Res. 2020.

67. Pal R, Sekh AA, Kar S, Prasad DK. Neural network based country wise risk prediction of COVID-19. arXiv preprint arXiv:200400959. 2020.

68. Papastefanopoulos V, Linardatos P, Kotsiantis S. COVID-19: A comparison of time series methods to forecast percentage of active cases per population. Applied Sciences (Switzerland). 2020;10(11).

69. Peng Y, Li C, Rong Y, Pang CP, Chen X, Chen H. Real-time Prediction of the Daily Incidence of COVID-19 in 215 Countries and Territories Using Machine Learning: Model Development and Validation. J Med Internet Res. 2021;23(6):e24285.

70. Pereira IG, Guerin JM, Júnior AGS, Garcia GS, Piscitelli P, Miani A, et al. Forecasting covid-19 dynamics in brazil: A data driven approach. International Journal of Environmental Research and Public Health. 2020;17(14):1-26.

71. Perone G. ARIMA forecasting of COVID-19 incidence in Italy, Russia, and the USA. arXiv preprint arXiv:200601754. 2020.

72. Perone G. Comparison of ARIMA, ETS, NNAR and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy. arXiv preprint arXiv:201011617. 2020.
73. Quintero Y, Ardila D, Camargo E, Rivas F, Aguilar J. Machine learning models for the prediction of the SEIRD variables for the COVID-19 pandemic based on a deep dependence analysis of variables. Comput Biol Med. 2021;134:104500.
74. Ribeiro M, da Silva RG, Mariani VC, Coelho LDS. Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. Chaos Solitons Fractals. 2020;135:109853.
75. Rustam F, Reshi AA, Mehmood A, Ullah S, On BW, Aslam W, et al. COVID-19 Future Forecasting Using Supervised Machine Learning Models. IEEE Access. 2020;8:101489-99.
76. Saba AI, Elsheikh AH. Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks. Process Saf Environ Prot. 2020;141:1-8.
77. Said AB, Erradi A, Aly HA, Mohamed A. Predicting COVID-19 cases using bidirectional LSTM on multivariate time series. Environ Sci Pollut Res Int. 2021:1-10.
78. Saqib M. Forecasting COVID-19 outbreak progression using hybrid polynomial-Bayesian ridge regression model. Applied Intelligence. 2020.
79. Senapati A, Nag A, Mondal A, Maji S. A novel framework for COVID-19 case prediction through piecewise regression in India. Int J Inf Technol. 2020:1-8.
80. Shyam Sunder Reddy K, Padmanabha Reddy YCA, Mallikarjuna Rao C. Recurrent neural network based prediction of number of COVID-19 cases in India. Mater Today Proc. 2020.
81. Tabar BR, Rendon-Sanchez JF. Forecasting COVID-19 daily cases using phone call data. ArXiv. 2020;abs/2010.02252.
82. Torrealba-Rodriguez O, Conde-Gutiérrez RA, Hernández-Javier AL. Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. Chaos Solitons Fractals. 2020:109946.

83. Tuli S, Tuli S, Tuli R, Gill SS. Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. Internet of Things. 2020;11:100222.

84. Tuli S, Tuli S, Verma R, Tuli R. Modelling for prediction of the spread and severity of COVID-19 and its association with socioeconomic factors and virus types. medRxiv. 2020.

85. Wang L, Adiga A, Venkatramanan S, Chen J, Lewis B, Marathe M. Examining Deep Learning Models with Multiple Data Sources for COVID-19 Forecasting. arXiv preprint arXiv:201014491. 2020.

86. Zawbaa HM, El-Gendy A, Saeed H, Osama H, Ali AMA, Gomaa D, et al. A study of the possible factors affecting COVID-19 spread, severity and mortality and the effect of social distancing on these factors: Machine learning forecasting model. Int J Clin Pract. 2021;75(6):e14116.

87. Zeroual A, Harrou F, Dairi A, Sun Y. Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. Chaos Solitons Fractals. 2020;140:110121.

88. Zivkovic M, Bacanin N, Venkatachalam K, Nayyar A, Djordjevic A, Strumberger I, et al. COVID-19 cases prediction by using hybrid machine learning and beetle antennae search approach. Sustain Cities Soc. 2021;66:102669.

89. Wong ZSY, Zhou J, Zhang Q. Artificial Intelligence for infectious disease Big Data Analytics. Infection, disease & health. 2019;24(1):44-8.
90. Painuli D, Mishra D, Bhardwaj S, Aggarwal M. Forecast and prediction of COVID-19 using machine learning. Data Science for COVID-19. 2021:381-97.

91. Samaniego E, Anitescu C, Goswami S, Nguyen-Thanh VM, Guo H, Hamdia K, et al. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. Computer Methods in Applied Mechanics and Engineering. 2020;362:112790.

92. Mukherjee H, Ghosh S, Dhar A, Obaidullah SM, Santosh K, Roy K. Deep neural network to detect COVID-19: one architecture for both CT Scans and Chest X-rays. Applied Intelligence. 2021;51(5):2777-89.

93. Silva R, Xavier F, Saraiva A, Cugnasca C, editors. Unsupervised machine learning and pandemics spread: the case of COVID-19. Anais do XX Simpósio Brasileiro de Computação Aplicada à Saúde; 2020: SBC.