Coronavirus Disease (COVID-19) Global Prediction Using Hybrid Artificial Intelligence Method of ANN Trained with Grey Wolf Optimizer

Sina Ardabili
Department of Biosystem Engineering,
University of Mohaghegh Ardabili,
Ardabil, Iran. s.ardabili@ieee.org

Amir Mosavi
Institute of Automation, Obuda
University, 1034 Budapest, Hungary
amir.mosavi@kvk.uni-obuda.hu

Shahab S. Band
National Yunlin University of Science
and Technology, Yunlin 64002,
Taiwan,
shamshirbands@yuntech.edu.tw

Annamaria R. Varkonyi-Koczy
Institute of Automation, Obuda
University, 1034 Budapest, Hungary
varkonyi-koczy@uni-obuda.hu

Abstract— advancement of the novel models for time-series prediction of COVID-19 is of utmost importance. Machine learning (ML) methods have recently shown promising results. The present study aims to engage an artificial neural network-integrated by grey wolf optimizer for COVID-19 outbreak predictions by employing the Global dataset. Training and testing processes have been performed by time-series data related to January 22 to September 15, 2020 and validation has been performed by time-series data related to September 16 to October 15, 2020. Results have been evaluated by employing mean absolute percentage error (MAPE) and correlation coefficient (r) values. ANN-GWO provided a MAPE of 6.23, 13.15 and 11.4% for training, testing and validating phases, respectively. According to the results, the developed model could successfully cope with the prediction task.

Keywords— COVID-19, Machine learning (ML), Grey wolf optimizer (GWO), artificial neural network (ANN), time-series, outbreak prediction

I. INTRODUCTION

Coronavirus Disease (COVID-19) pandemic broke out in December 2019 in Wuhan, China [1]. COVID-19 outbreak has shown irregularity and uncertainties, which pose significant challenges for the conventional epidemiological models. The advancement of the novel models for time-series prediction of COVID-19 is of utmost importance to better support the strategy of combating against the consequences of COVID-19 [2]. Recently, several techniques, including mathematical models, have been successfully employed for pandemic prediction of disease. Singhal et al. (2020) used the Fourier decomposition method (FDM) and susceptible-infected-recovered (SIR) model for the prediction of the COVID-19 pandemic. According to the results, the total number of cases and death was predicted to be 12.7×106 and 5.27×105, respectively [3]. Sarkar et al. (2020) employed a mathematical model for the pandemic prediction of the COVID-19 in India. The proposed model was a combination of susceptible (S), recovered (R), asymptomatic (A), infected (I), quarantined susceptible (Sq), and isolated infected (Iq) models [4]. Anirudh (2020) employed SIR, SEIRU, SEIR, SLIAR, SIRD, ARIMA and SIDARTHE for the COVID-19 pandemic prediction [5]. But, mathematical models have disadvantages such as complexity, time-consuming and lower reliability [6, 7].

Machine Learning (ML) based models have been successfully employed in different fields of science [8]. Pandemic prediction can be considered one of the ML-based models' fields that provided good accuracy and reliability. Recently, researchers have developed studies for the prediction of COVID-19 cases and death rates using ML-based methods. Tabrizchi et al. (2020) employed a Deep Learning-based estimation model integrated with an image-based diagnosis method to detect coronavirus infection [9]. Dutta et al. (2020) used a DL based prediction model for the estimation of the COVID-19 outbreak [10]. Ardabili et al. (2020) employed different single and hybrid ML-based methods to predict the COVID-19 epidemic in USA, Germany, China, Iran and Italy. According to the results, ML-based prediction models could successfully cope with the prediction task [11]. Pinter et al. (2020) employed hybrid ML-based models (ANFIS and ANN-ICA) for the prediction of total cases of COVID-19 in Hungary using time-series data

[6]. Accordingly, the present study aimed to predict the global COVID-19 outbreak using a robust hybrid Artificial Neural Network integrated by Grey Wolf Optimizer (ANN-GWO) in the presence of the time-series total and daily COVID-19 cases.

II. METHODOLOGY

The required data for developing the prediction model was generated from https://www.worldometers.info/. The nature of the dataset is time-series based. Accordingly, the total and the daily number of cases are presented in Figure I from January 22 to November 15, 2020. Figure I has two vertical axes. One is for total cases and the other is related to the daily cases.
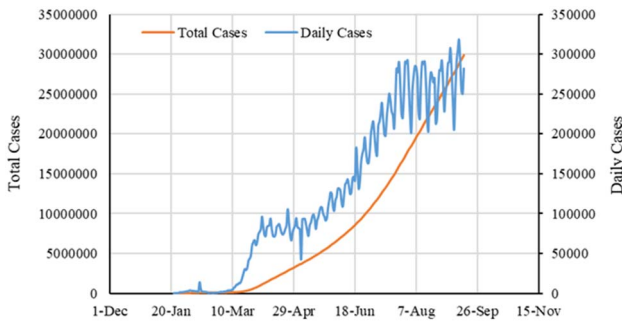

FIGURE I. THE EMPLOYED TIME-SERIES DATASET FOR DEVELOPING THE ML-BASED PREDICTION MODEL

Due to the use of time-series data, it has been decided to choose the input dataset according to Table I for the prediction of the global COVID-19 outbreak as the output parameter.

TABLE I. DESCRIPTION OF THE INPUTS AND OUTPUT OF MODELING PROCESS

Inputs	Output
$X_{(t-1)}$, $X_{(t-2)}$, $X_{(t-3)}$, $X_{(t-4)}$, $X_{(t-5)}$, $X_{(t-6)}$ and $X_{(t-7)}$	Global COVID-19 Outbreak

ANN-GWO performed modeling in the presence of input and output datasets. ANN-GWO can be considered as one of the robust hybrid ML-based methods [12]. The modeling phase was started by employing the different architectures of ANN and different population sizes for GWO. Finally, according to the model's accuracy, ANN with the architecture of 7-10-4-1 in the presence of population size 150 for GWO was selected as the best prediction model for the global COVID-19 prediction with a high accuracy. 70% of total data were chosen for the training phase, 30% of total data were employed for the testing phase and the predicted COVID-19 outbreak for September 16 to October 20, 2020 were selected for the validation phase. The description of ANN-GWO has been comprehensively presented in our previous study in [11]. Figure II presents the main schematic of the developed model.
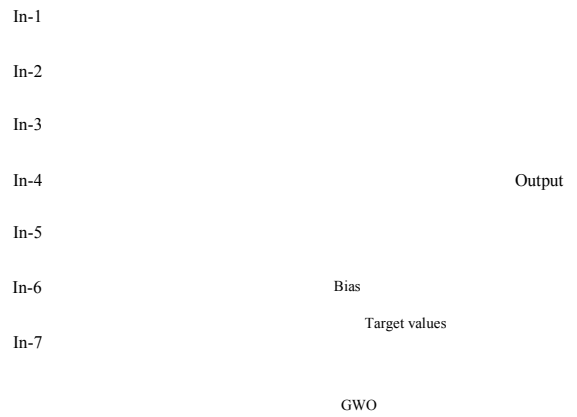

FIGURE II. THE MAIN SCHEMATIC OF ANN-GWO

According to Figure II, the main mechanism of the GWO is to update the bias and weights of the desired ANN architecture to reach the highest accuracy.

The performance and accuracy of the developed method was evaluated using the Mean Absolute Percentage Error (MAPE) and correlation coefficient (r) according to Equations 1 and 2:

$$MAPE = \frac{\frac{\sum_{k=1}^{N} \left| \frac{y_k^{pred.} - y_k^{exp.}}{y_k^{exp.}} \right|}{N}}{\times 100} \tag{1}$$

$$R = \sqrt{1 - \frac{\sum_{k=1}^{N} \left(y_k^{exp.} - y_k^{pred.} \right)^2}{\sum_{k=1}^{N} \left(y_k^{exp.} - y_{ave.} \right)^2}} \tag{2}$$

Where $y_k^{exp.}$ Refers to the target values, $y_k^{pred.}$ Refers to the predicted values and N refers to the number of data.

III. RESULTS

Table II presents the training results, testing and validating phases into two terms: MAPE and r values. According to the results, the selected ANN architecture with a 150 at the maximum iteration of 500 provided a good accuracy for training, testing and validating phases.

TABLE II. RESULTS OF THE TRAINING, TESTING AND VALIDATING PHASES

ANN architecture	Pop. size	Maximum iteration	Phase	MAPE	Correlation coefficient
7-10-4-1	150	500	Training	6.23 %	0.999
			Testing	13.15%	0.994
			Validating	11.4%	0.994

Figure III (a) and (b) present the plot diagram of the predicted and target values and the deviation from the testing phase's target values. As is clear from Figure 3a and according to the testing phase's determination coefficient, the testing phase's performance is acceptable.
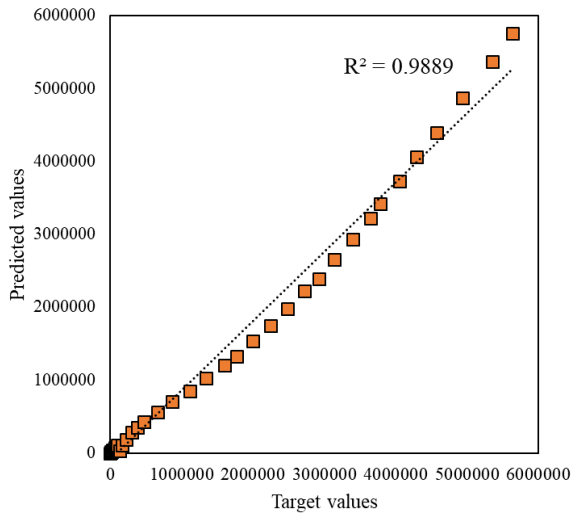
(A)
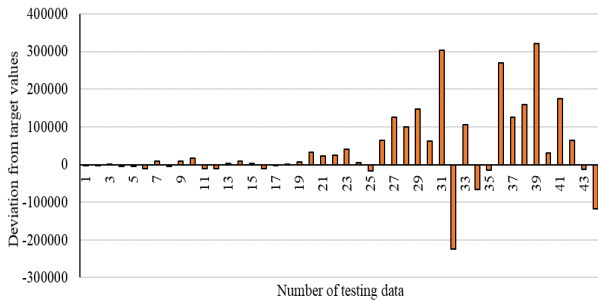
(B)

FIGURE III. (A) THE PLOT DIAGRAM, (B) DEVIATION FROM TARGET VALUES.

Figure 4 presents the outbreak prediction by ANN-GWO from October 16 to June 4, 2021 (for 232 days. As is clear according to the projections, the outbreak of the COVID-19 is still increasing and the trend is rising. This can be a severe warning to the World Health Organization.

FIGURE IV. THE PREDICTION OF THE GLOBAL COVID-19 OUTBREAK FROM OCTOBER 16 TO JUNE 4, 2021

IV. CONCLUSION

In this paper, the COVID-19 outbreak is modeled as a complex time series. We have developed a hybrid machine learning model based on the artificial neural network. We trained the system with the grey wolf optimization algorithm to get the highest performance. Based on the testing, we projected the outbreak until late May. Training and testing processes have been performed by time-series data related to January 22 to September 15, 2020 and validation has been performed by time-series data associated with September 16 to October 15, 2020. Results have been evaluated by employing mean absolute percentage error (MAPE) and correlation coefficient (r) values. ANN-GWO provided a MAPE of 6.23, 13.15 and 11.4% for training, testing and validating phases, respectively. According to the results, the developed model could successfully cope with the prediction task. The model provided promising results, and it can be used for outbreak prediction.

References

1. Fanelli, D., F.J.C. Piazza, Solitons, and Fractals, *Analysis and forecast of COVID-19 spreading in China, Italy and France*. 2020. **134**: p. 109761.
2. Shen, Z., et al., *UV-thermal dual-cured polymers with degradable and anti-bacterial function*. 2020. **148**: p. 105783.
3. Singhal, A., et al., *Modeling and prediction of COVID-19 pandemic using Gaussian mixture model*. 2020. **138**: p. 110023.
4. Sarkar, K., et al., *Modeling and forecasting the COVID-19 pandemic in India*. 2020. **139**: p. 110049.
5. Anirudh, A.J.I.D.M., *Mathematical modeling and the transmission dynamics in predicting the Covid-19-What next in combating the pandemic*. 2020. **5**: p. 366-374.
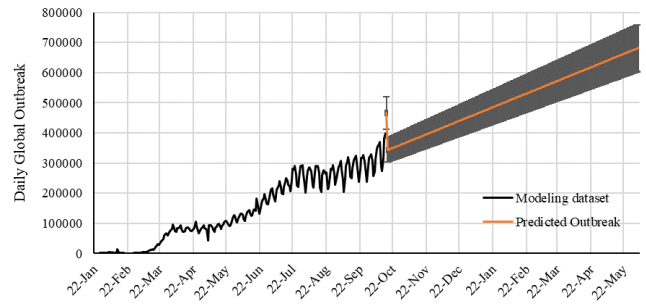6. Pinter, G., et al., *COVID-19 Pandemic Prediction for Hungary; a Hybrid Machine Learning Approach*. 2020. **8**(6): p. 890.
7. Fallah, S.N., et al., *Computational Intelligence Load Forecasting: A Methodological Overview*. 2018.
8. Shamshirband, S., T. Rabczuk, and K.-W.J.I.A. Chau, *A survey of deep learning techniques: application in wind and solar energy resources*. 2019. **7**: p. 164650-164666.
9. Tabrizchi, H., et al., *Rapid COVID-19 Diagnosis Using Deep Learning of the Computerized Tomography Scans*. 2020.
10. Dutta;, A., A. Gupta;, and F.H. Khan;, *COVID-19: Detailed Analytics & Predictive Modelling using Deep Learning*. International Journal of Scientific Research in Science and Technology, 2020. **7**(5): p. 85-94.

11. Ardabili, S.F., et al., *COVID-19 Outbreak Prediction with Machine Learning.* Algorithms 2020. **13**: p. 249.

12. Ahmadi, M.H., et al., *Machine learning prediction models of electrical efficiency of photovoltaic-thermal collectors.* 2019.