

From data to action: Empowering COVID-19 monitoring and forecasting with intelligent algorithms

Vincent Charles^a, Seyed Muhammad Hossein Mousavi^b, Tatiana Gherman^c and S. Muhammad Hassan Mosavi^d

^aQueen's Business School, Queen's University Belfast, Belfast, UK; ^bIndependent Scientist, Tehran, Iran; ^cFaculty of Business and Law, University of Northampton, Northampton, UK; ^dUniversity of Applied Sciences, Tehran, Iran

ABSTRACT

The COVID-19 pandemic has profoundly impacted every aspect of our lives, from economic to the social facets of contemporary society. While the new COVID-19 waves may not be anticipated to be as severe as previous ones, it would be unreasonable to assume that they will cease any time soon. Consequently, forecasting the number of future infections, recovered patients, and death cases remains a very much logical step in trying to fight against further waves, in conjunction with ongoing vaccination efforts. In this paper, we investigate the efficiency of three intelligent machine learning algorithms, namely GMDH, Bi-LSTM, and GA + NN, for COVID-19 forecasting, with an application to Iran and the United Kingdom. The experimental results show that the algorithms can be used to forecast the next six months of COVID-19 in terms of confirmed, recovered, and death cases, which gives a more ample timeframe for using the results to make better practical yet strategic decisions and take appropriate actions or measures to deploy resources effectively to contain or curb the spread of the coronavirus. Despite the distinct dynamics observed in the data, our analysis proves the robustness of the employed models.

ARTICLE HISTORY

Received 13 March 2021
Accepted 15 July 2023

KEYWORDS

OR in health services;
artificial intelligence;
machine learning; COVID-
19; pandemics; time series
forecasting

1. Introduction

The starting point of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) or COVID-19 was reported as December 2019, in the Hubei province and its capital, Wuhan (Acter et al., 2020). The virus became a pandemic on 11 March 2019 and spread worldwide. As of 28 January 2022, Iran had already entered a sixth wave, while in the UK, the latest surge in COVID-19 cases over Spring 2022 put the country on the brink of a sixth wave. According to the World Health Organization (WHO, 2022), as of 09 June 2022, a total of 531,550,610 people had been infected worldwide (confirmed cases) and as of 07 June 2022, a total of 11,854,673,610 vaccine doses were administered. In particular, as of 09 June 2022, 7,233,117 confirmed cases, 141,342 deaths, and 149,357,848 vaccine doses, and 7,057,938 recovered cases of COVID-19 were reported for Iran, while for the United Kingdom, these values were 22,363,071 confirmed cases, 179,083 deaths, 143,183,637 vaccine doses, and 22,013,928 recovered cases (JHU CSSE, 2022; WHO, 2022). So, the answer to what seems to be an everlasting question “When will the COVID-19 pandemic end?” is not a simple one.

While the new COVID-19 waves are not anticipated to be as severe as previous ones, it seems unreasonable to assume that they will cease any time soon. Moreover, new variants continue to emerge frequently, and it is not yet clear whether and to what extent the current available vaccines will be effective against these. At the same time, there is still a significant portion of the population that has not been vaccinated, some who simply have not developed immunity, and others who have a compromised immune system. For all of these groups, the virus remains a highly dangerous one, and in some cases even fatal. For all of the above reasons, the COVID-19 pandemic requires continuous special attention. Despite vaccination efforts, it remains important that we design forecasting and control strategies to restrain further waves. Forecasting the numbers of future infections, recovered patients, and death cases remains a very much logical step in trying to fight against the pandemic and plan for the future. It remains an essential part of many decision-making processes, a view that is widely shared (Nikolopoulos et al., 2021; Petropoulos & Makridakis, 2020).

Accurate modelling plays a key role in this process, as it can help the healthcare system and relevant policymakers in better planning the provision

CONTACT Tatiana Gherman  tatiana.gherman@northampton.ac.uk  Faculty of Business and Law, University of Northampton, Northampton, UK.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

of services for patients (Katris, 2021). Artificial Intelligence (AI) and its sub-field machine learning (ML), along with data mining techniques, can help in this regard (Mousavi, Charles, & Gherman, 2020; Mousavi et al., 2017). On the one hand, data mining is the process of finding specific patterns and correlations within large amounts of data (Dezfoulan et al., 2016). ML, on the other hand, involves a computer programme that is trained on a training dataset and can then adapt and learn by itself to make sense of the data without the need of further human intervention (Mousavi & Lyashenko, 2017). In other words, ML can keep a computer's built-in algorithms current regardless of changes in the real-world. Through ML algorithms, it is possible to mine and learn from available datasets and produce new meaningful information, which can subsequently be used in decision-making processes to increase efficiency, decrease costs, improve customer relationships, and reduce risks, among others (Duda et al., 2012; Nielsen, 2015).

In this paper, we explore three ML algorithms, namely Group Method of Data Handling (GMDH) (Ivakhnenko, 1966), Bi-directional Long Short-Term Memory (Bi-LSTM) (Graves & Schmidhuber, 2005), and Genetic Algorithm (GA) (Gu et al., 2011; Holland, 1975) for COVID-19 forecasting, with an application to the cases of Iran and the United Kingdom. An observation to be made at this point concerns the difference between forecast and prediction. A forecast refers to a calculation or an estimation which uses data from previous events, combined with recent trends to come up with a future event outcome. Therefore, forecasting relies on time-series data, while prediction does not. Instead, a prediction is a statement which tries to explain a possible outcome or future event, being concerned with estimating the outcomes for unseen data. Our interest in the present paper is to forecast the number of confirmed, recovered, and death cases.

In the context of AI being a new paradigm for healthcare systems, intelligent ML algorithms can be employed for analysing COVID-19 data and informing decision-making processes. This means that AI-driven tools can help identify new COVID-19 waves, as well as forecast their nature of spread across the Globe. One of the fundamental requirements is that sufficient data are available to train the respective models. Earlier in the pandemic, most of the AI-driven tools used by prior studies to forecast the pandemic were limited to proof-of-concept models. However, as more and more data are being generated every day, this opens the possibility to recheck existing algorithms for their robustness.

The selection of GMDH, Bi-LSTM, and GA + NN forecasting approaches among a plethora of available intelligent algorithms is justified by their unique strengths and capabilities, as well as proven effectiveness in addressing the complexities of COVID-19 forecasting. These algorithms were indeed chosen after an extensive review of the existing literature and a thorough evaluation of their performance in various forecasting tasks, including epidemic forecasting. GMDH is known for its inductive nature and ability to automatically select relevant features and handle complex, non-linear relationships in the data, which is particularly crucial for modelling infectious diseases like COVID-19. Bi-LSTM, on the other hand, leverages its ability to incorporate information from both the past and present in input sequences, allowing it to capture temporal dependencies and dynamic patterns that are crucial in understanding the evolving nature of the pandemic. Finally, GA + NN offers the potential for optimisation, enhancing the accuracy and robustness of the forecasts. By relying on the strengths and capabilities of these approaches, we aim to ensure a comprehensive analysis that accounts for various factors influencing COVID-19 dynamics, resulting in more accurate and reliable forecasts.

Many algorithms are being created frequently. And while we do acknowledge the importance of creating new and perhaps better algorithms, it is also right that we balance such a view with an approach where we can also make use of what we already have and has been proven to work. In this sense, then, our work complements the existing literature. The fact that we use three well known methods that have a track record of proven robustness (*i.e.*, GMDH, Bi-LSTM, and GA + NN) for predictive analytics (Charles et al., 2022) to forecast the pandemic is an advantage which works to counter-balance the still yet little understood phenomenon called COVID-19. As mentioned, we were able to forecast the next six months of COVID-19 in terms of confirmed, recovered, and death cases, which gives a more ample amount of time to be able to use the results to make better practical yet strategic decisions and take appropriate actions or measures to contain or curb the spread of the coronavirus.

This study makes several contributions. First, it provides an original contribution in terms of a research process that employs neural networks (NNs) and evolutionary-based systems to derive more intelligent, nature-inspired results. This answers to calls for the evaluation of more intelligent algorithms to study and understand the evolution of the COVID-19 pandemic (*e.g.*, Agbehadji et al., 2020). Our experimental results show that the

mentioned forecasting approaches are deemed to be the most reliable among other possible methods. Second, and more importantly, the paper makes an empirical contribution. The three ML algorithms are tested in the context of the COVID-19 pandemic using the most up-to-date and reliable databases, which also works to decrease the error. We forecast the next six months of COVID-19 in terms of confirmed, recovered, and death cases. The results have practical value for a range of stakeholders, particularly for healthcare administrators and decision-makers, governments, and policymakers. Moreover, the novelty of the present work lies in the forecasting of three categories of confirmed, recovered, and death cases, as opposed to most studies that consider one or two categories. We also provide a more comprehensive evaluation of the accuracy of the models by means of using eight performance metrics.

The paper consists of six main sections. [Section 1](#) lays out the background and fundamentals of the proposed research. [Section 2](#) explores relevant prior related works on the topic of COVID-19 forecasting. [Section 3](#) relays details regarding the dataset used and introduces the three forecasting methods of our interest, along with their immediate variants, and two traditional time series forecasting methods. Implementation and validation metrics are discussed in [Section 4](#). Subsequently, [Section 5](#) proffers a discussion of results and implications for practice and policy. Finally, [Section 6](#) concludes the paper with a summary of key points, additional reflections and implications, and a discussion of avenues for future research.

2. Literature review

Due to its severity and significant impact on economies and the society, a myriad of research studies on COVID-19 have already been performed and published and many others are on their way. This is because despite bearing similarities with two previous diseases, namely the severe acute respiratory syndrome (SARS) and the Middle Eastern Respiratory Syndrome (MERS), this pandemic also has singularities: it is a novel type of coronavirus, and although vaccines have been developed to tackle it, the behaviour of the virus is still yet to be fully understood by scientists. In this sense, COVID-19 remains a public health issue of international proportions.

COVID-19 has been the subject of numerous studies. Some research studies have taken a macro-level approach to the pandemic, focusing on the strategic management of the virus and its implications for the economy and society (e.g., Block et al.,

2020; Mandal et al., 2020). Others have focused on the various factors (such as frequency of testing, delay in receiving the results, etc.) affecting the prevalence of the disease over time (e.g., Chang & Kaplan, 2023), while others have focused on drug and vaccine development (e.g., Dhama et al., 2020; Liu et al., 2020). The vast majority of the studies, however, have focused on forecasting the growth and trend of the virus (e.g., Hansun et al., 2021, 2022). The reason behind the growth in the third category of studies resides in their potential to be conducive to pre-emptive governmental strategies for minimising the number of COVID-19 cases (number of infections, recovered patients, and death cases). Various types of methods have been used for prediction purposes, such as mathematical models, time series analyses, and even sophisticated soft computing algorithms (Hansun et al., 2021). While conducting an exhaustive literature review is beyond the scope of this paper, we will highlight some relevant studies that focus on predicting and forecasting COVID-19, aiming to shed light on the methods utilised in this domain.

Roosa et al. (2020) used daily reported cumulative case data from the National Health Commission of China to generate short-term forecasts using a generalised logistic growth model, the Richards growth model, and a sub-epidemic wave model. While being a worthwhile endeavour that takes into account the pre-vaccination situation, the forecasts are only valid for a period of 15 days. Other works have used traditional (yet widely used) time series forecasting methods. For example, Petropoulos and Makridakis (2020) used exponential smoothing with multiplicative error and multiplicative trend components to forecast confirmed cases, deaths, and recoveries; however, they were only able to produce 10-days ahead point forecasts. The study was extended by Petropoulos et al. (2022) to more rounds of forecasts and two variables; nevertheless, the model remains a short-term forecasting model (with 10-day ahead forecasts). Al-Qaness et al. (2020) introduced a new forecasting model to estimate and forecast the number of confirmed cases of COVID-19 in the upcoming ten days based on the previously confirmed cases recorded in China. The proposed model is an improved adaptive neuro-fuzzy inference system (ANFIS) using an enhanced flower pollination algorithm (FPA) by using the salp swarm algorithm (SSA). Fanelli and Piazza (2020) used a modified susceptible-infected-recovered-deaths (SIRD) model, wherein the infection rate was allowed to vary with time, to analyse and forecast the COVID-19 spreading in China, Italy, and France. The research used data gathered from the GitHub repository associated with the interactive

dashboard hosted by JHU CSSE. This research extended the forecasted period to one month, from peak start to peak end, with proper accuracy. However, this study is from 2020, when very little data were available. Sarkar et al. (2020) proposed a mathematical model that predicted the dynamics of COVID-19 in 17 provinces of India and the overall India, based on data obtained from a wide range of official sources. The proposed model monitors the dynamics of six compartments, namely: susceptible (S), asymptomatic (A), recovered (R), infected (I), isolated infected (Iq), and quarantined susceptible (Sq), collectively expressed as SARIqSq. In addition, their data cover two months of forecasts and pre-vaccination period.

Ghanbari (2020) forecasted the number of infected people in the second wave of the COVID-19 outbreak in Iran by viewing the problem in the framework of the mathematical modelling of a system of differential equations. The study covered two months of COVID-19 forecasting. Kapoor et al. (2020) examined a novel forecasting approach for COVID-19 case prediction that uses Graph Neural Networks and mobility data and evaluated the approach on the US county level COVID-19 dataset, New York Times (NYT) COVID-19 dataset, the Google COVID-19 Aggregated Mobility Research Dataset, and the Google Community Mobility Reports. Zeroual et al. (2020) used five deep learning algorithms to forecast the number of new and recovered cases of COVID-19. Specifically, the authors used simple Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Bi-directional LSTM (Bi-LSTM), Gated Recurrent Units (GRUs), and Variational AutoEncoder (VAE) algorithms, which they applied on daily data in six countries, namely Italy, Spain, France, China, USA, and Australia. This research covered up to two months of COVID-19 forecasting before vaccination. Abdulaal et al. (2020) aimed to create a point-of-admission mortality risk scoring system using an Artificial Neural Network (ANN). The proposed ANN analyses a set of patient features including demographics, comorbidities, smoking history, and presenting symptoms and predicts patient-specific mortality risk during the current hospital admission.

Shahid et al. (2020) proposed forecast models comprising AutoRegressive Integrated Moving Average (ARIMA), Support Vector Regression (SVR), GRU, LSTM, and Bi-LSTM, which were assessed for time series prediction of confirmed cases, deaths, and recoveries in ten major countries affected by COVID-19. Also, their system predicted less than two months. Chumachenko et al. (2020) proposed a hybrid NN-based Group Method of

Data Handling (GMDH) forecasting algorithm, which they used to forecast the second wave COVID-19 cases in China and the USA. Another mentionable research using GMDH belongs to Vaishnav and Vajpai (2020), who focused on assessing the impact of relaxation in lockdown to forecast the number of active cases for the period of six months. However, this research pertains to the year 2020 and is not applicable to recent data that includes post-vaccination information. Rizk-Allah and Hassanien (2020) proposed an improved Interior Search Algorithm (ISA) based on Chaotic Learning (CL) strategy, named ISACL, which was implemented to improve the performance of the Multi-layer Feed-forward Neural Network (MFNN) by finding its optimal structure regarding the weights and biases. The proposed forecasting method was employed to forecast the number of confirmed cases of COVID-19 based on official daily data from WHO. Their study covered 38 days of forecasting. Also, Salgotra et al. (2020) used Genetic Programming (GP) to model the COVID-19 pandemic in terms of confirmed and death cases for the 15 most infected countries. By combining genetic algorithm with LSTM, Zhang and Liu (2021) further proposed an intelligent model to predict the number of new confirmed cases in Brazil; their system, however, covered less than a month.

Our research differs from prior studies in the following ways: (1) Most of the existing studies have used just one dataset for forecasting and, generally, for the second wave, whereas this study integrates two datasets, with the most up-to-date data available at the time of writing (*i.e.*, up to 09 January 2022). Moreover, our research considered post vaccination situation and data. (2) Additionally, prior studies have generally performed forecasts just for the next two months at best, which are hardly of practical value, whereas in this study we work on a six-month forecast. (3) Also, prior studies have generally forecasted one or two categories of cases, that is, either confirmed, recovered, or death cases, whereas in this study we forecast all the three categories. As previously mentioned, also, the present research answers calls for the evaluation of more intelligent algorithms to study the evolution of the COVID-19 pandemic; to this aim, the paper makes an original contribution in terms of a research process that employs NNs and evolutionary-based systems to derive more intelligent, nature-inspired results. Our approach offers several advantages over more traditional forecasting methods like exponential smoothing when it comes to forecasting pandemics; for example, it captures non-linear patterns, handles temporal dependencies, and shows adaptability to evolving data, among others.

3. Dataset and methods

There is a considerable number of up-to-date datasets on COVID-19 available, which contain time-series data. This research employs Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) (Dong et al., 2020) and World Health Organization (WHO, 2020) data as input, in view of the fact that these datasets are not only reliable, but also the most up to date. Data are gathered from these two databases and used as one dataset in our paper. Of course, we should note that despite the quality of these datasets, no dataset is perfect; differences and errors between these datasets will always be present and when such differences are significant, they may influence the models (Kalgotra et al., 2021; Sokol, 2020). However, the discrepancies are relatively minimal in our case; moreover, we rely on three categories of data ("confirmed", "death", and "recovered"), so when they are all trending up together, for example, we can get a pretty good sense that there is a new wave approaching. Dataset discrepancies were conducted by comparing our employed data with other available COVID-19 datasets and also datasets available on the Kaggle platform. So, while there is always a trade-off in using multiple datasets, combining the datasets improves the quality of the data. Moreover, understanding complex issues and responding to global challenges (such as is the case of the global pandemic) requires combining datasets collected by different organisations, sometimes even by organisations that use different standards and different ways to describe the pandemic.

For some countries like Iran, there are no data available for the first 32 days (± 5 , depending on dataset and category) and for all categories. This value is different for each country. In the case of the United Kingdom, this value is 21 days (± 5 , again depending on dataset and category). So, these first data points were removed as they can affect the training data and cause a significant amount of error. On the positive side, however, more than two years of data are available at the time of writing, which fixes this issue. Apart from the above regarding the initial period, there were no missing data; hence, we did not perform any imputation procedures.

Three algorithms, namely Group Method of Data Handling (GMDH) (Ivakhnenko, 1966), Bi-directional Long Short-Term Memory (Bi-LSTM) (Graves & Schmidhuber, 2005), and Genetic Algorithm (GA) (Holland, 1975) forecasting systems (Gu et al., 2011) are employed to forecast the number of confirmed (infected), recovered, and death cases of COVID-19 in Iran and the United Kingdom. The present research expands the

forecasting steps into the future by four extra months to reach six months, with proper accuracy on up-to-date data, alongside with using more creative and robust algorithms.

3.1. Group Method of Data Handling (GMDH)

Group Method of Data Handling (GMDH) (Ivakhnenko, 1966) is a sub-model of ANN that has been traditionally used for deep learning and knowledge discovery, forecasting, data mining, optimisation, and pattern recognition. Inductive GMDH algorithms foster the possibility to find interrelations in the data automatically, to select an optimal structure of model or network, and to increase the accuracy of existing algorithms (Farlow, 1984).

NN forecasting is more flexible than typical linear or polynomial approximations and is thus more precise. With NNs one can discover and take into account non-linear connections and relationships between data and build a candidate model with high prediction strength. GMDH Shell is a professional neural network software that solves time series forecasting and data mining tasks by building ANNs and applying them to the input data. The hybridisation of GMDH involves the use of various neurons, including classical, nonlinear Adaline, R-neuron, W-neuron, and wavelet-neurons. The choice of the number of neurons and layers is crucial for the hybridisation of a GMDH neural network. Figure 1 represents an example of a typical GMDH neural network and a hybrid GMDH neural network. We used Matlab code for generating the final results and GMDH Shell for validation purposes.

3.2. Bi-directional, Long Short-Term Memory (Bi-LSTM)

Bi-LSTM is a combination of LSTM (Hochreiter and Schmidhuber, 1997) and Bi-directional Recurrent Neural Networks (Bi-RNN) (Schuster & Paliwal, 1997). On the one hand, RNN is a special development of ANN that processes sequences and time series data. RNN has the advantage to encode dependencies between inputs. On the other hand, LSTM is a special kind of RNN (Selvin et al., 2017), capable of learning long-term dependencies. It was introduced by Hochreiter and Schmidhuber (1997) and was refined and popularised in time by many researchers. LSTM works tremendously well on a large variety of problems and is now widely used across application areas. LSTMs are explicitly designed to avoid the long-term dependency problem of the RNN. Remembering information for long periods of time is practically their default behaviours.

A Bi-directional LSTM (Bi-LSTM) (Graves & Schmidhuber, 2005) is a sequence processing model that consists of two LSTMs: one taking the input in a forward direction, and the other in a backwards direction (Shahid et al., 2020). Bi-LSTMs effectively increase the amount of information available to the network, improving the context available to the algorithm. The purpose of the Bi-LSTM is to look at a particular sequence both from front-to-back as well as from back-to-front. In this way, the network creates a context for each character in the text that depends on both its past as well as its future. These types of networks are highly efficient in forecasting and predicting according to their structure. Both LSTM and Bi-LSTM network structures are presented in Figure 2.

3.3. Nature-based genetic forecasting algorithm

Genetic algorithms (GAs) (Holland, 1975) are problem-solving methods (or heuristics) that simulate the process of natural evolution, harnessing the power of nature. Unlike ANNs, which have been designed to function like neurons in the brain, GAs determine the best solution for a problem by means of relying on the concepts of natural selection. In consequence, GAs are commonly used as optimisers that adjust parameters to minimise or maximise some feedback measure, which can then be first, be used independently or in the construction of an ANN, as this research used in the second form (Meng, 2012). There are three types of genetic operations that can then be performed (Kuepper, 2020):

- Crossovers, which represent the reproduction and crossover seen in biology, whereby a child takes on certain characteristics of its parents.
- Mutations, which represent biological mutations and are used to maintain genetic diversity from one generation of a population to the next by introducing random small changes.
- Selections, which are the stage at which individual genomes are chosen from a population for later breeding (recombination or crossover).

These three operations are then used in a five-step process:

1. Initialise a random population.
2. Select the chromosomes.
3. Apply mutation or crossover operators.
4. Recombine the offspring and the current population to form a new population.
5. Repeat steps two to four.

The process terminates when stopping criteria are met, which can include running time, fitness, number of generations, or other criteria.

Here, the rules of the employed variables (confirmed, recovered, and death cases) involve the use of parameters like Moving Average Convergence Divergence (MACD), an Exponential Moving Average (EMA), and stochastics (Kuepper, 2020). A GA would then input values into these parameters with the goal of maximising the values of the variables. Over time, small changes are introduced, and those that make a desirable impact are retained for the next generation in the network. In PSO + NN forecasting, NN is optimised to a better position during each iteration by PSO. A typical NN is very much dependent on the number of hidden layers and the number of nodes in each layer, and these parameters need to be adjusted manually, resulting in operation complexity. Therefore, PSO is useful for resolving this issue by finding the optimal parameter combination for NN forecasting. Figure 3

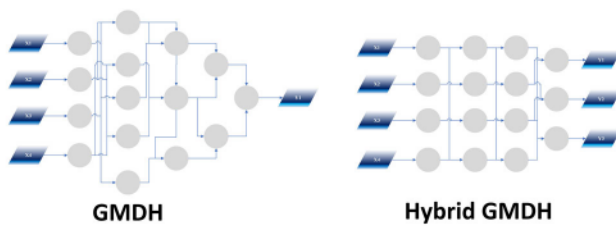


Figure 1. Structure of typical GMDH and hybrid-GMDH neural networks.

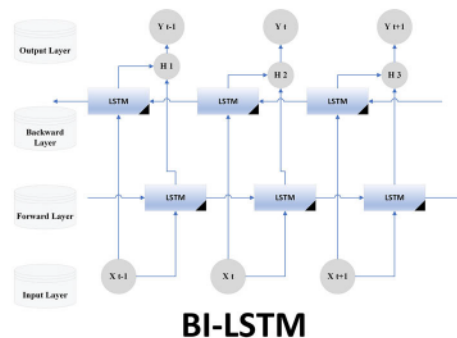
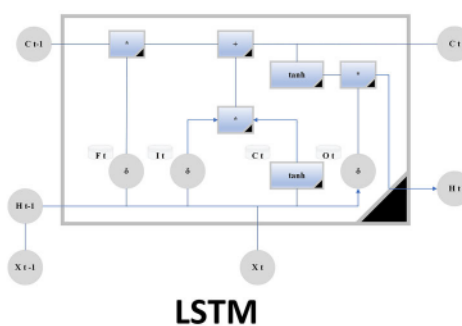


Figure 2. LSTMs and Bi-LSTMs network structures.

depicts the structure of the GA + NN and PSO + NN forecasting systems.

Also, Figure 4 shows the flowchart of the proposed research process that employs NNs and evolutionary-based systems to derive more intelligent results. GeneXproTools is a software for forecasting time series data using GA. However, we used Matlab code for generating the final results and GeneXproTools software for validation purposes.

4. Implementation and validation metrics

After the algorithms were developed and the three categories of cases were forecasted, their results were compared and evaluated using a series of performance or validation metrics. The validation metrics used are Mean Squared Error (MSE) (Equation 1), Root Mean Squared Error (RMSE) (Equation 2), Mean Absolute Error (MAE) (Equation 3), Mean Absolute Percentage Error (MAPE) (Equation 4), Explained Variance (EV) (Equation 5), and Root

Mean Squared Log Error (RMSLE) (Equation 6), Correlation Coefficient (CC) (Equation 7), Coefficient of Determination or R-squared (Equation 8) (Asuero et al., 2006).

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (2)$$

$$MAE = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n} \quad (3)$$

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \% \quad (4)$$

$$EV = 1 - \frac{Var(\hat{y} - y)}{Var(y)} \quad (5)$$

$$RMSLE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\log(y_t) - \log(\hat{y}_t))^2} \quad (6)$$

$$CC = r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (7)$$

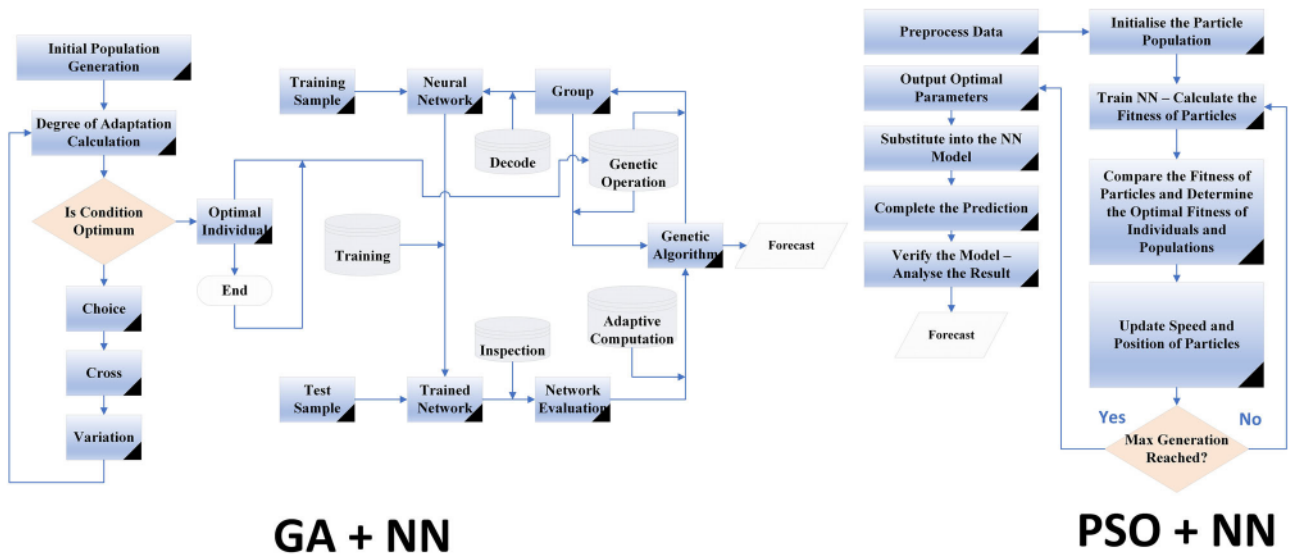


Figure 3. GA + NN and PSO + NN forecasting system structures.

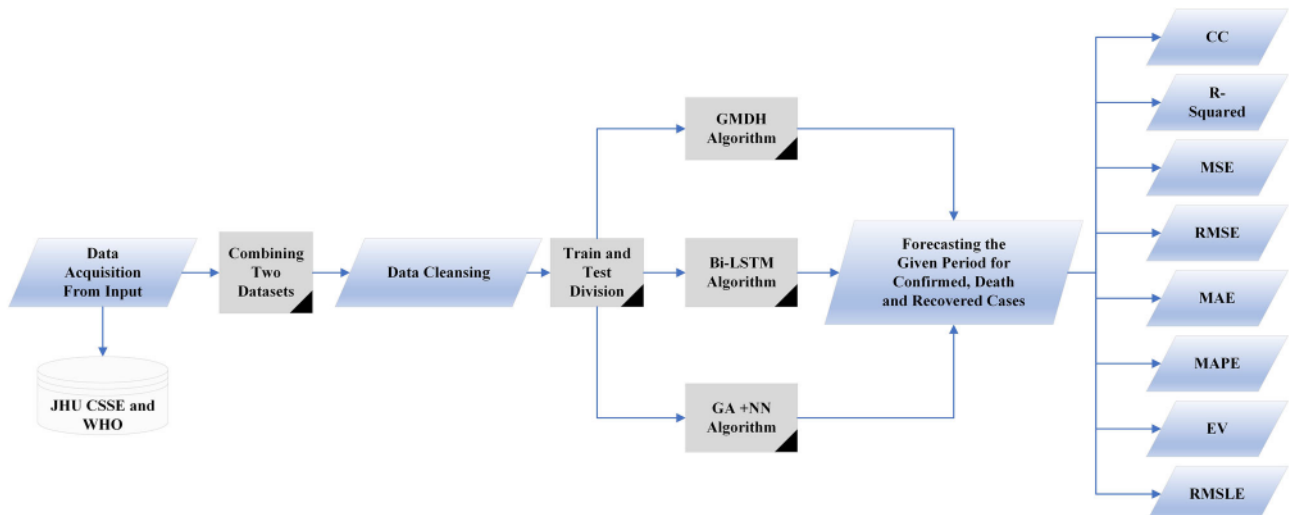


Figure 4. Proposed forecasting method flowchart.

$$R - Squared = r^2 = \left(\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \right)^2 \quad (8)$$

wherein y_i represents the actual values and $\hat{y}_i$ represents the corresponding estimated values. Also, n is the number of measurements. The benefit of using RMSLE as a statistical indicator lies in its robustness to outliers. Lower MSE, RMSE, MAE, and MAPE values and EV, CC, and R-squared values closer to 1 are representative of more accurate forecasting performances. The employed data for both datasets comprise 690 days of confirmed cases, 673 days of death cases, and 525 days of recovered cases until 09 Jan 2022. We adopted the split strategy, with 3

scenarios: 60%-40%, 70%-30%, and 80%-20% train and test, respectively.

All performance metrics are normalised in the 0 to 1 range for better understanding. Table 1 depicts the validation metrics results for “COVID-19 Confirmed cases” forecasting using the GMDH algorithm, for Iran and the UK, as well as a comparison with the results obtained by applying hybrid-GMDH, as in Chumachenko et al. (2020). Table 1 also includes the same details for “COVID-19 Confirmed cases” forecasting using the Bi-LSTM method, and a comparison with the results obtained by applying LSTM, as in Shahid et al. (2020). Lastly, Table 1 also shows the same validation metrics for “COVID-19 Confirmed cases” forecasting implemented using the GA + NN method, alongside a comparison with the results obtained by employing PSO + NN, as in Rizk-Allah and Hassanien (2020).

Table 1. Validation metrics for “COVID-19 Confirmed cases” forecasting in Iran and the UK using GMDH, Bi-LSTM, and GA + NN methods.

Partition	Region	Model	CC	R-squared	MSE	RMSE	MAE	MAPE	EV	RMSLE
60 % Train 40 % Test	Iran	GMDH	0.9853	0.9702	0.0228	0.1513	0.0611	0.1713	0.8288	0.2224
	UK	GMDH	0.9813	0.9624	0.0146	0.1212	0.0773	0.1925	0.8123	0.1895
	Iran	Hybrid-GMDH	0.9885	0.9761	0.0185	0.1363	0.1211	0.1961	0.8016	0.2602
	UK	Hybrid-GMDH	0.9794	0.9584	0.0169	0.1305	0.0996	0.2003	0.8164	0.2058
	Iran	Bi-LSTM	0.9917	0.9821	0.0256	0.1601	0.0914	0.1887	0.9666	0.1226
	UK	Bi-LSTM	0.9936	0.9860	0.0121	0.110	0.1115	0.1204	0.9700	0.1100
	Iran	LSTM	0.9903	0.9801	0.0289	0.1708	0.1211	0.1935	0.8746	0.1975
	UK	LSTM	0.9913	0.9821	0.0488	0.2216	0.1298	0.1202	0.8908	0.1378
	Iran	GA + NN	0.9835	0.9663	0.0144	0.1204	0.0917	0.1787	0.9001	0.1874
	UK	GA + NN	0.9808	0.9604	0.0123	0.1115	0.0976	0.1551	0.9121	0.1901
	Iran	PSO + NN	0.9717	0.9428	0.0137	0.1172	0.0824	0.1612	0.9002	0.2013
	UK	PSO + NN	0.9754	0.9506	0.0125	0.1121	0.0947	0.1591	0.9103	0.1874
	Iran	ARIMA	0.9676	0.9362	0.0491	0.2216	0.1299	0.2022	0.8000	0.2939
	UK	ARIMA	0.9644	0.9300	0.0536	0.2315	0.1398	0.2065	0.7821	0.2894
	Iran	ES	0.9677	0.9364	0.0577	0.2402	0.1325	0.2001	0.7880	0.2798
	UK	ES	0.9632	0.9278	0.0599	0.2447	0.1302	0.2006	0.7902	0.3025
	Iran	GMDH	0.9872	0.9746	0.0249	0.1577	0.0518	0.1477	0.8509	0.2032
	UK	GMDH	0.9832	0.9667	0.0163	0.1276	0.0680	0.1689	0.8344	0.1703
	Iran	Hybrid-GMDH	0.9904	0.9809	0.0204	0.1427	0.1118	0.1725	0.8237	0.2410
	UK	Hybrid-GMDH	0.9813	0.9630	0.0187	0.1369	0.0903	0.1767	0.8385	0.1866
70 % Train 30 % Test	Iran	Bi-LSTM	0.9936	0.9873	0.0277	0.1665	0.0821	0.1651	0.9887	0.1034
	UK	Bi-LSTM	0.9955	0.9910	0.0136	0.1164	0.1022	0.0968	1.0021	0.0808
	Iran	LSTM	0.9922	0.9845	0.0314	0.1772	0.1118	0.1699	0.8967	0.1783
	UK	LSTM	0.9932	0.9865	0.0520	0.2280	0.1205	0.0966	0.9129	0.1186
	Iran	GA + NN	0.9854	0.9710	0.0161	0.1268	0.0824	0.1551	0.9222	0.1682
	UK	GA + NN	0.9827	0.9657	0.0139	0.1179	0.0883	0.1315	0.9342	0.1709
	Iran	PSO + NN	0.9736	0.9479	0.0153	0.1236	0.0731	0.1376	0.9223	0.1821
	UK	PSO + NN	0.9773	0.9551	0.0140	0.1185	0.0854	0.1355	0.9324	0.1682
	Iran	ARIMA	0.9761	0.9527	0.0690	0.2627	0.1403	0.1908	0.7924	0.2739
	UK	ARIMA	0.9695	0.9400	0.0688	0.2623	0.1590	0.1911	0.7855	0.2711
	Iran	ES	0.9714	0.9436	0.0651	0.2551	0.1578	0.1827	0.7988	0.2998
	UK	ES	0.9777	0.9558	0.0778	0.2789	0.1614	0.2007	0.7930	0.2801
	Iran	GMDH	0.9761	0.9528	0.0205	0.1432	0.0472	0.1434	0.8572	0.2123
	UK	GMDH	0.9721	0.9450	0.0117	0.1082	0.0634	0.1646	0.8407	0.1794
	Iran	Hybrid-GMDH	0.9793	0.9591	0.0159	0.1261	0.1072	0.1682	0.8300	0.2501
	UK	Hybrid-GMDH	0.9702	0.9413	0.0142	0.1192	0.0857	0.1724	0.8448	0.1957
	Iran	Bi-LSTM	0.9825	0.9654	0.0234	0.1530	0.0775	0.1608	0.9750	0.1125
	UK	Bi-LSTM	0.9844	0.9691	0.0089	0.0943	0.0976	0.0925	0.9684	0.0899
	Iran	LSTM	0.9811	0.9626	0.0271	0.1646	0.1072	0.1656	0.9030	0.1874
	UK	LSTM	0.9821	0.9646	0.0480	0.2191	0.1159	0.0923	0.9192	0.1277
80 % Train 20 % Test	Iran	GA + NN	0.9743	0.9493	0.0115	0.1072	0.0778	0.1508	0.9285	0.1773
	UK	GA + NN	0.9716	0.9441	0.0093	0.0964	0.0837	0.1272	0.9405	0.1800
	Iran	PSO + NN	0.9625	0.9265	0.0107	0.1034	0.0685	0.1333	0.9286	0.1912
	UK	PSO + NN	0.9662	0.9336	0.0094	0.0970	0.0808	0.1312	0.9387	0.1773
	Iran	ARIMA	0.9493	0.9011	0.0669	0.2587	0.1354	0.1891	0.8012	0.2547
	UK	ARIMA	0.9497	0.9019	0.0681	0.2610	0.1358	0.1997	0.8124	0.2687
	Iran	ES	0.9492	0.9010	0.0770	0.2775	0.1274	0.1924	0.8055	0.2770
	UK	ES	0.9486	0.8998	0.0710	0.2665	0.1432	0.1990	0.8222	0.2559

Note. The validation metrics include comparisons with hybrid-GMDH (Chumachenko et al., 2020), LSTM (Shahid et al., 2020), and PSO + NN (Rizk-Allah & Hassanien, 2020), as well as traditional ES and ARIMA methods.

Table 2. Validation metrics for “COVID-19 Recovered cases” forecasting in Iran and the UK using GMDH, Bi-LSTM, and GA + NN methods.

Partition	Region	Model	CC	R-squared	MSE	RMSE	MAE	MAPE	EV	RMSLE
60 % Train 40 % Test	Iran	GMDH	0.9820	0.9643	0.0219	0.1481	0.0537	0.1673	0.8345	0.2116
	UK	GMDH	0.9780	0.9565	0.0139	0.1180	0.0699	0.1885	0.8180	0.1787
	Iran	Hybrid-GMDH	0.9852	0.9706	0.0177	0.1331	0.1137	0.1921	0.8073	0.2494
	UK	Hybrid-GMDH	0.9761	0.9528	0.0162	0.1273	0.0922	0.1963	0.8221	0.1950
	Iran	Bi-LSTM	0.9884	0.9769	0.0246	0.1569	0.0840	0.1847	0.9723	0.1118
	UK	Bi-LSTM	0.9903	0.9807	0.0114	0.1068	0.1041	0.1164	0.9757	0.0992
	Iran	LSTM	0.9870	0.9742	0.0281	0.1676	0.1137	0.1895	0.8803	0.1867
	UK	LSTM	0.9880	0.9761	0.0477	0.2184	0.1224	0.1162	0.8965	0.1270
	Iran	GA + NN	0.9802	0.9608	0.0137	0.1172	0.0843	0.1747	0.9058	0.1766
	UK	GA + NN	0.9775	0.9555	0.0117	0.1083	0.0902	0.1511	0.9178	0.1793
	Iran	PSO + NN	0.9684	0.9378	0.0130	0.1140	0.0750	0.1572	0.9059	0.1905
	UK	PSO + NN	0.9721	0.9450	0.0119	0.1089	0.0873	0.1551	0.9160	0.1766
	Iran	ARIMA	0.9552	0.9125	0.0547	0.2339	0.1491	0.2158	0.7541	0.2935
	UK	ARIMA	0.9620	0.9254	0.0551	0.2347	0.1632	0.2006	0.7638	0.2640
	Iran	ES	0.9500	0.9025	0.0590	0.2429	0.1477	0.2001	0.7990	0.2150
	UK	ES	0.9564	0.9147	0.0667	0.2583	0.1528	0.2009	0.7125	0.2660
	Iran	GMDH	0.9793	0.9591	0.0159	0.1260	0.0632	0.1430	0.8381	0.1932
	UK	GMDH	0.9865	0.9732	0.0199	0.1411	0.1070	0.1642	0.8274	0.1603
	Iran	Hybrid-GMDH	0.9774	0.9553	0.0183	0.1353	0.0855	0.1678	0.8422	0.2310
	UK	Hybrid-GMDH	0.9897	0.9795	0.0272	0.1649	0.0773	0.1720	0.9924	0.1766
	Iran	Bi-LSTM	0.9916	0.9833	0.0132	0.1148	0.0974	0.1604	0.9991	0.0934
	UK	Bi-LSTM	0.9883	0.9768	0.0308	0.1756	0.1070	0.0921	0.9004	0.0708
70 % Train 30 % Test	Iran	LSTM	0.9893	0.9787	0.0513	0.2264	0.1157	0.1652	0.9166	0.1683
	UK	LSTM	0.9815	0.9634	0.0157	0.1252	0.0776	0.0919	0.9259	0.1086
	Iran	GA + NN	0.9788	0.9581	0.0135	0.1163	0.0835	0.1504	0.9379	0.1582
	UK	GA + NN	0.9697	0.9403	0.0149	0.1220	0.0683	0.1268	0.9260	0.1609
	Iran	PSO + NN	0.9734	0.9475	0.0137	0.1169	0.0806	0.1329	0.9361	0.1721
	UK	PSO + NN	0.9736	0.9479	0.0193	0.1388	0.0396	0.1308	0.8637	0.1582
	Iran	ARIMA	0.9609	0.9234	0.0688	0.2623	0.1247	0.1897	0.8042	0.2861
	UK	ARIMA	0.9615	0.9245	0.0697	0.2640	0.1369	0.1820	0.8063	0.2843
	Iran	ES	0.9554	0.9127	0.0711	0.2666	0.1300	0.1834	0.8024	0.2882
	UK	ES	0.9644	0.9300	0.0700	0.2646	0.1497	0.1955	0.8011	0.2987
	Iran	GMDH	0.9760	0.9526	0.0151	0.1229	0.0998	0.1394	0.8357	0.2015
	UK	GMDH	0.9669	0.9349	0.0135	0.1160	0.0783	0.1606	0.8505	0.1686
	Iran	Hybrid-GMDH	0.9792	0.9588	0.0224	0.1498	0.0701	0.1642	0.9807	0.2393
	UK	Hybrid-GMDH	0.9811	0.9626	0.0083	0.0911	0.0902	0.1684	0.9741	0.1849
	Iran	Bi-LSTM	0.9778	0.9561	0.0260	0.1614	0.0998	0.1568	0.9087	0.1017
	UK	Bi-LSTM	0.9788	0.9580	0.0466	0.2159	0.1085	0.0885	0.9249	0.0791
	Iran	LSTM	0.9710	0.9428	0.0108	0.1040	0.0704	0.1616	0.9342	0.1766
	UK	LSTM	0.9683	0.9376	0.0087	0.0932	0.0763	0.0883	0.9462	0.1169
	Iran	GA + NN	0.9592	0.9201	0.0100	0.1002	0.0611	0.1468	0.9343	0.1665
	UK	GA + NN	0.9629	0.9272	0.0088	0.0938	0.0734	0.1232	0.9444	0.1692
80 % Train 20 % Test	Iran	PSO + NN	0.9840	0.9600	0.0241	0.1551	0.0527	0.1293	0.9345	0.1804
	UK	PSO + NN	0.9730	0.9465	0.0392	0.1980	0.0779	0.1272	0.9180	0.1665
	Iran	ARIMA	0.9550	0.9120	0.0666	0.2581	0.1132	0.1894	0.8011	0.2541
	UK	ARIMA	0.9517	0.9058	0.0669	0.2587	0.1249	0.1925	0.8169	0.2562
	Iran	ES	0.9527	0.9077	0.0670	0.2588	0.1099	0.1742	0.8245	0.2794
	UK	ES	0.9540	0.9101	0.0600	0.2449	0.1097	0.1988	0.8002	0.2566

Note. The validation metrics include comparisons with hybrid-GMDH (Chumachenko et al., 2020), LSTM (Shahid et al., 2020), PSO + NN (Rizk-Allah & Hassanien, 2020), as well as traditional ES and ARIMA methods.

The reason behind comparing our results with those obtained by running the other three methods mentioned above (namely, hybrid-GMDH, LSTM, and PSO + NN) lies in the fact that these methods are among the most recent ones to have been shown to be superior for COVID-19 forecasting purposes. Table 2 reports the validation metrics results for forecasting using the same three methods, but for the variable “COVID-19 Recovered cases”; and Table 3 provides them for the variable “COVID-19 Death cases”. In addition, two traditional and well-known algorithms, Exponential Smoothing (ES) (e.g., De Livera et al., 2011) and the Autoregressive Integrated Moving Average (ARIMA) model (Ho & Xie, 1998), are included in the comparisons to provide a basis for understanding the advantages of using intelligent forecasting methods.

Table 4 shows the properties of each algorithm in the experiment, for which the output results are derived. The figures for each of the three methods (GMDH, Bi-LSTM, and GA + NN) are presented for Iran and the UK, but each method is presented for one variable only. In other words, the outputs are presented for the variable “Confirmed cases” in the case of GMDH, for the variable “Recovered cases” in the case of Bi-LSTM, and for the variable “Death cases” in the case of “GA + NN.”

Figure 5 shows the forecasted results using the GMDH algorithm for the next six months, from 09 Jan 2022 until 09 Jul 2022 for Iran; and Figure 6 shows the same for the UK. Figures 7 and 8 then present the forecasted results using the Bi-LSTM algorithm for the next six months for Iran and the UK, respectively. Lastly, Figures 9 and 10 provide

Table 3. Validation metrics for “COVID-19 Death cases” forecasting for Iran and the UK using GMDH, Bi-LSTM, and GA + NN methods.

Partition	Region	Model	CC	R-squared	MSE	RMSE	MAE	MAPE	EV	RMSLE
60 % Train 40 % Test	Iran	GMDH	0.9866	0.9734	0.0198	0.1408	0.0465	0.1658	0.8437	0.1953
	UK	GMDH	0.9826	0.9655	0.0123	0.1107	0.0627	0.1870	0.8272	0.1624
	Iran	Hybrid-GMDH	0.9898	0.9797	0.0158	0.1258	0.1065	0.1906	0.8165	0.2331
	UK	Hybrid-GMDH	0.9807	0.9618	0.0144	0.1200	0.0850	0.1948	0.8313	0.1787
	Iran	Bi-LSTM	0.9930	0.9860	0.0224	0.1496	0.0768	0.1832	0.9815	0.0955
	UK	Bi-LSTM	0.9949	0.9898	0.0099	0.0995	0.0969	0.1149	0.9849	0.0829
	Iran	LSTM	0.9916	0.9833	0.0257	0.1603	0.1065	0.1880	0.8895	0.1704
	UK	LSTM	0.9926	0.9853	0.0446	0.2111	0.1152	0.1147	0.9057	0.1107
	Iran	GA + NN	0.9848	0.9698	0.0121	0.1099	0.0771	0.1732	0.9150	0.1603
	UK	GA + NN	0.9821	0.9645	0.0102	0.1010	0.0830	0.1496	0.9270	0.1630
	Iran	PSO + NN	0.9730	0.9467	0.0114	0.1067	0.0678	0.1557	0.9151	0.1742
	UK	PSO + NN	0.9767	0.9539	0.0103	0.1016	0.0801	0.1536	0.9252	0.1603
	Iran	ARIMA	0.9611	0.9237	0.0498	0.2232	0.1327	0.2103	0.7950	0.2579
	UK	ARIMA	0.9672	0.9354	0.0471	0.2170	0.1391	0.2206	0.7813	0.2451
	Iran	ES	0.9719	0.9445	0.0512	0.2263	0.1400	0.2019	0.7522	0.2637
	UK	ES	0.9695	0.9400	0.0583	0.2415	0.1402	0.2022	0.7708	0.2973
70 % Train 30 % Test	Iran	GMDH	0.9946	0.9893	0.0181	0.1347	0.0468	0.1396	0.8527	0.1705
	UK	GMDH	0.9855	0.9713	0.0166	0.1289	0.0906	0.1608	0.8420	0.1376
	Iran	Hybrid-GMDH	0.9978	0.9957	0.0251	0.1585	0.0691	0.1644	0.8568	0.2083
	UK	Hybrid-GMDH	0.9997	0.9995	0.0117	0.1084	0.0609	0.1686	1.0070	0.1539
	Iran	Bi-LSTM	0.9964	0.9929	0.0286	0.1692	0.0810	0.1570	1.0204	0.0707
	UK	Bi-LSTM	0.9974	0.9949	0.0484	0.2200	0.0906	0.0887	0.9150	0.0481
	Iran	LSTM	0.9896	0.9794	0.0141	0.1188	0.0993	0.1618	0.9312	0.1456
	UK	LSTM	0.9869	0.9741	0.0121	0.1099	0.0612	0.0885	0.9405	0.0859
	Iran	GA + NN	0.9778	0.9562	0.0134	0.1156	0.0671	0.1470	0.9525	0.1355
	UK	GA + NN	0.9815	0.9634	0.0122	0.1105	0.0519	0.1234	0.9406	0.1382
	Iran	PSO + NN	0.9817	0.9638	0.0175	0.1324	0.0642	0.1295	0.9507	0.1494
	UK	PSO + NN	0.9849	0.9701	0.0133	0.1153	0.0232	0.1274	0.8783	0.1355
	Iran	ARIMA	0.9539	0.9100	0.0696	0.2639	0.1129	0.1937	0.8235	0.2547
	UK	ARIMA	0.9551	0.9122	0.0585	0.2419	0.1197	0.1902	0.8325	0.2461
	Iran	ES	0.9593	0.9203	0.0667	0.2582	0.1025	0.1987	0.8461	0.2691
	UK	ES	0.9647	0.9307	0.0778	0.2790	0.1101	0.1920	0.8411	0.2450
80 % Train 20 % Test	Iran	GMDH	0.9838	0.9679	0.0203	0.1425	0.0926	0.1379	0.8449	0.1852
	UK	GMDH	0.9857	0.9716	0.0070	0.0838	0.0711	0.1591	0.8597	0.1523
	Iran	Hybrid-GMDH	0.9824	0.9651	0.0237	0.1541	0.0629	0.1627	0.9899	0.2230
	UK	Hybrid-GMDH	0.9834	0.9671	0.0435	0.2086	0.0830	0.1669	0.9833	0.1686
	Iran	Bi-LSTM	0.9756	0.9518	0.0094	0.0967	0.0926	0.1553	0.9179	0.0854
	UK	Bi-LSTM	0.9729	0.9465	0.0074	0.0859	0.1013	0.0870	0.9341	0.0628
	Iran	LSTM	0.9638	0.9289	0.0086	0.0929	0.0632	0.1601	0.9434	0.1603
	UK	LSTM	0.9675	0.9361	0.0075	0.0865	0.0691	0.0868	0.9554	0.1006
	Iran	GA + NN	0.9886	0.9773	0.0218	0.1478	0.0539	0.1453	0.9435	0.1502
	UK	GA + NN	0.9732	0.9557	0.0364	0.1907	0.0662	0.1217	0.9536	0.1529
	Iran	PSO + NN	0.9776	0.9554	0.0187	0.1499	0.0335	0.1278	0.9437	0.1641
	UK	PSO + NN	0.9800	0.9611	0.0134	0.1112	0.0567	0.1257	0.9272	0.1502
	Iran	ARIMA	0.9487	0.9000	0.0575	0.2397	0.1537	0.1981	0.8130	0.2784
	UK	ARIMA	0.9496	0.9018	0.0504	0.2246	0.1699	0.1803	0.8195	0.2836
	Iran	ES	0.9500	0.9025	0.0493	0.2221	0.1470	0.1936	0.8234	0.2594
	UK	ES	0.9541	0.9103	0.0620	0.2490	0.1327	0.1775	0.8082	0.2639

Note. The validation metrics include comparisons with hybrid-GMDH (Chumachenko et al., 2020), LSTM (Shahid et al., 2020), PSO + NN (Rizk-Allah & Hassanien, 2020), as well as traditional ES and ARIMA methods.

the forecasted results using the GA + NN algorithm for the next six months for Iran and the UK, respectively. Figures include regressions, training stage, curves, and loss functions during iterations, for better readability. Also, to note that “target” and “output” variables in the figures are “actual” and “predicted” variables, respectively. So, “actual or target” is original data and “output or predicted” is what we get as the result from the system. The difference between these two provides the final error.

5. Discussion of results and managerial implications

According to the performance metrics results provided in Tables 1–3, both the employed and the

comparison methods showed nice performance in view of all the eight performance metrics. This indicates that all the algorithms can be exploited for pandemic forecasting for better planning and management (although it should be noted at the outset that ML models largely outperform the traditional approaches of ES and ARIMA). Among the ML algorithms, “PSO + NN” exhibits slightly better performance compared to the counterpart method across all three categories of cases (confirmed, death, and recovered). This is because PSO is capable to effectively perform global search during network training, resulting in better forecast models while avoiding easy entrapment in a local minimum. However, the difference is insignificant.

GA is a traditional and common optimisation solver in AI that operates with values encoded in binary form in most cases. PSO does the same, but

Table 4. Algorithms' properties.

ML Model	Properties	
GMDH	Recorder Observations	Desc. + Odd / Even
	Validation Strategy	Whole Data Testing – 60 to 40 (time partitioned)
	Validation Criterion	RMSE
	Variable Ranking	By Error
	Drop Variables After Rank	600
	Core Algorithm	Neural-Type
	Neuron Function	Linear
	Max Number of Layers	33
	Round to the Nearest Integer	Yes
	Train / Test	60 to 40, 70 to 30, 80 to 20
Bi-LSTM	Layers in Bi-LSTM Unit One and Two	250
	Epochs	500
	Batch Size	120
	Neural Type	Deep
	Input Delay	[1 2 3 4 5 6 12 16 18 20 24]
	Train Test Division	75 to 25
	Hidden Layers	[30 50]
	Function Fitting Network	Bayesian Regularisation
	Train / Test	60 to 40, 70 to 30, 80 to 20
	Generations	200
GA + NN	No of Chromosomes	30
	No of Gens	20
	Embedding Dimension	9
	Fitness Function	MAE, MSE, RMSE
	Lower and Upper Bounds	[-10 10]
	Mutation	0.0138
	Recombination	0.0755
	Train / Test	60 to 40, 70 to 30, 80 to 20

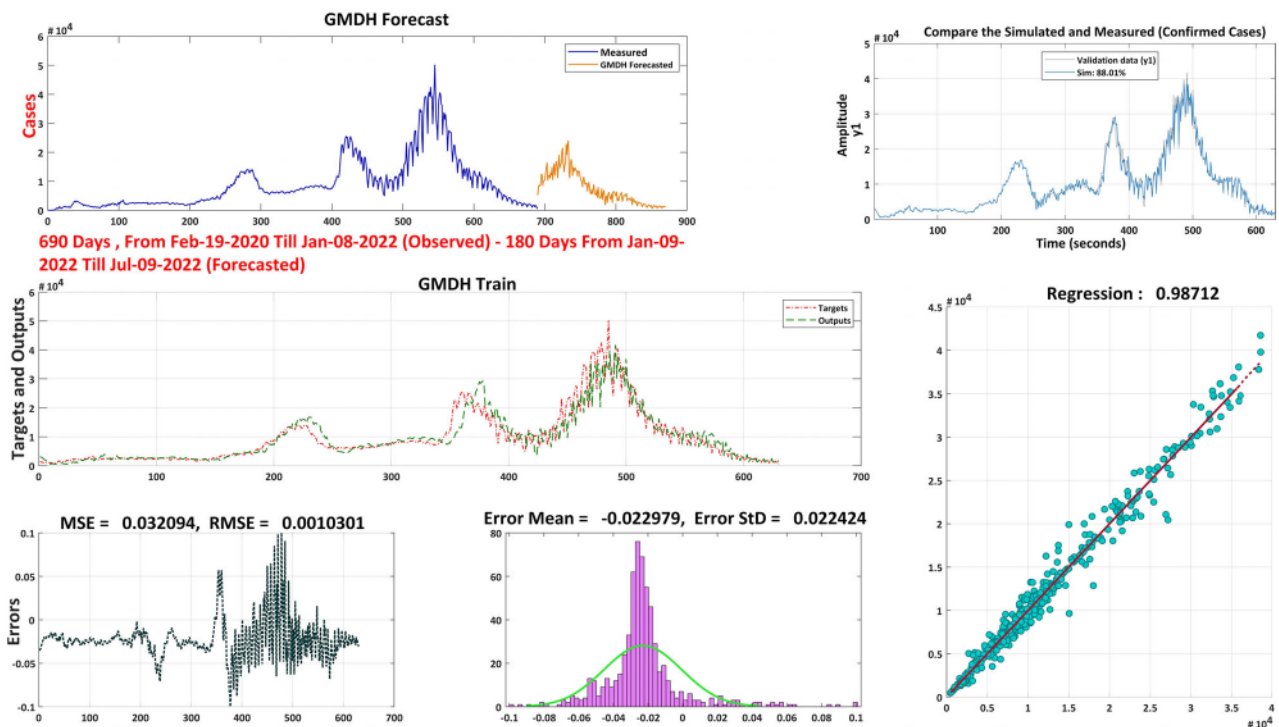


Figure 5. Forecasted result for "Confirmed cases" using the GMDH algorithm for Iran (next 180 days). Note: MSE, RMSE, Error Mean, Error StD, and R-squared are included. Also, the high similarity between Targets and Outputs in GMDH Train indicates very good performance.

it is superior in terms of complexity, accuracy, iteration, and programme simplicity in finding the optimal solution. GMDH outperformed hybrid-GMDH in long-term forecasting (more than two months) in the experiment. However, in the short-term, the hybrid version performed better. Also, as Bi-LSTM is an extension of traditional LSTM, it can improve model performance on sequence forecast problems

here. In problems where all timesteps of the input sequence are available, Bi-LSTMs train two instead of one LSTM on the input sequence.

The three algorithms used in this research approach the solution in different ways, resulting in distinct patterns of newly generated waves compared to traditional methods. For example, GMDH algorithms employ an inductive approach, gradually

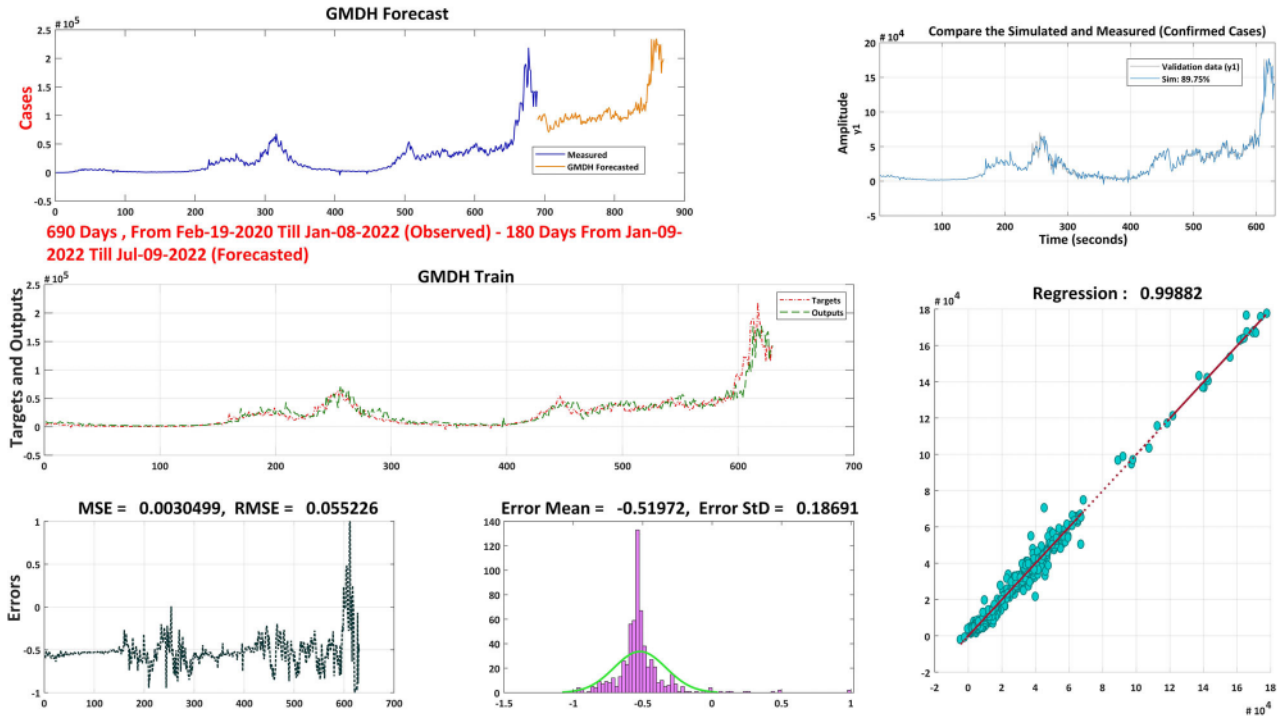


Figure 6. Forecasted result for "Confirmed cases" using the GMDH algorithm for the UK (next 180 days). Note: MSE, RMSE, Error Mean, Error StD, and R-squared are included. Also, the high similarity between Targets and Outputs in GMDH Train indicates very good performance.

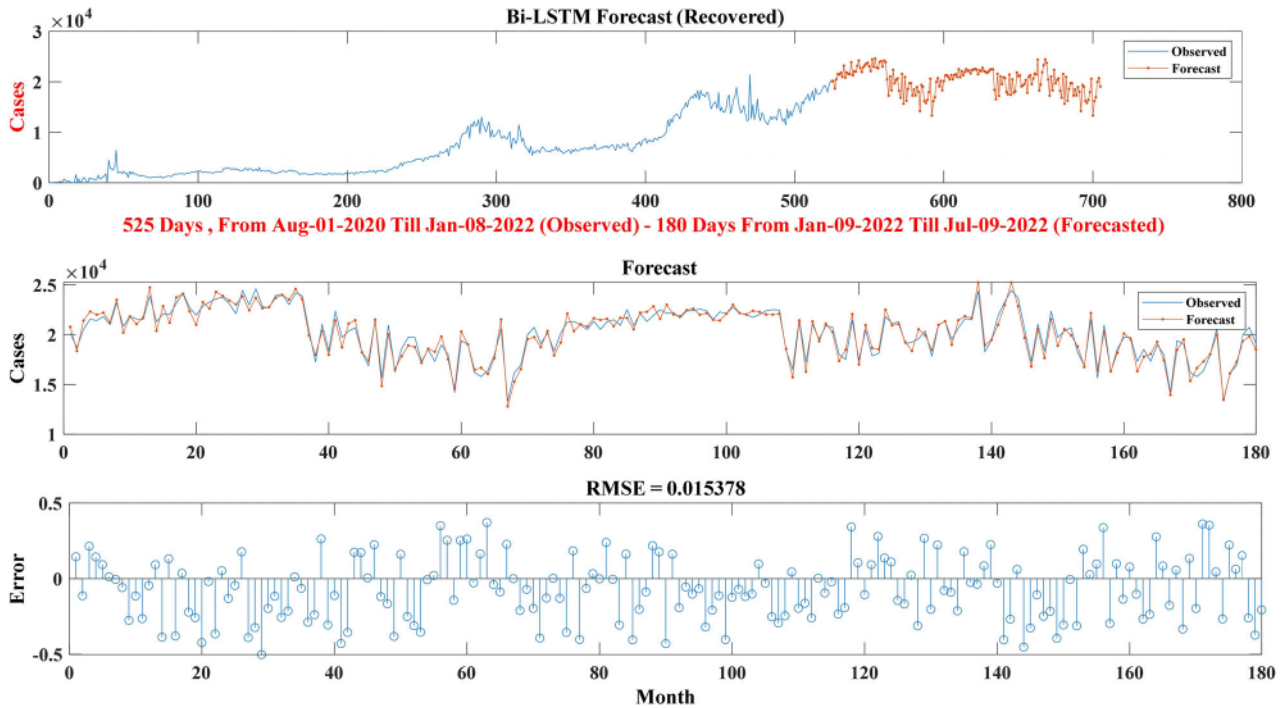


Figure 7. Forecasted result for "Recovered cases" using the Bi-LSTM algorithm for Iran (next 180 days). Note: The high similarity between Observed and Forecasted in the forecast part indicates very good performance.

organising complex polynomial models and selecting the best solution based on external criteria. As a result, Figures 5 and 6 exhibit gradually generated waves. In the Bi-LSTM algorithm, each component of an input sequence incorporates information from both the past and present, explaining the similarity between the generated waves in Figures 7 and 8 compared to previous waves. The GA + NN, on the other hand, is based on

natural selection, favouring the fittest entities. In our case, the fittest represents the average of optimised cases from previous waves. Figures 9 and 10 demonstrate how the GA + NN attempts to produce new offspring (cases) resembling combinations of previous cases, thus shaping the waves.

The goal behind using performance metrics is to understand how each algorithm works according to

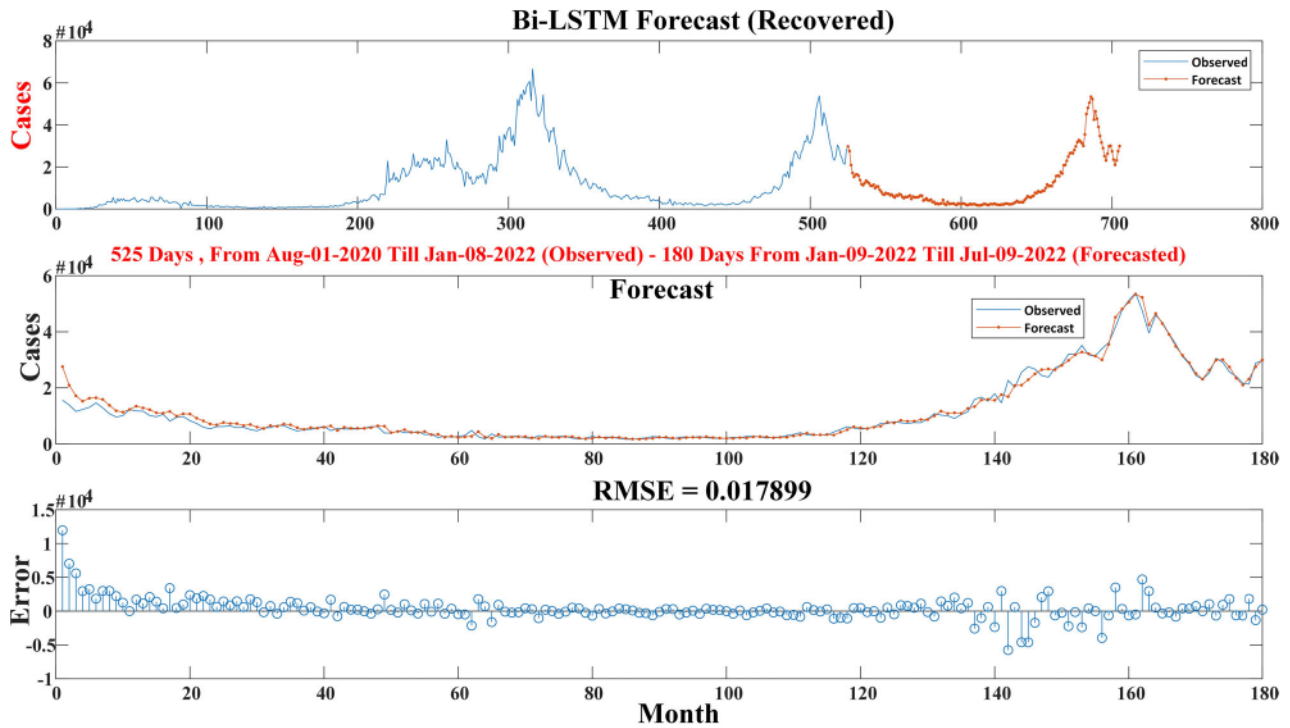


Figure 8. Forecasted result for “Recovered cases” using the Bi-LSTM algorithm for UK the (next 180 days). Note: The high similarity between Observed and Forecasted in the forecast part indicates very good performance.

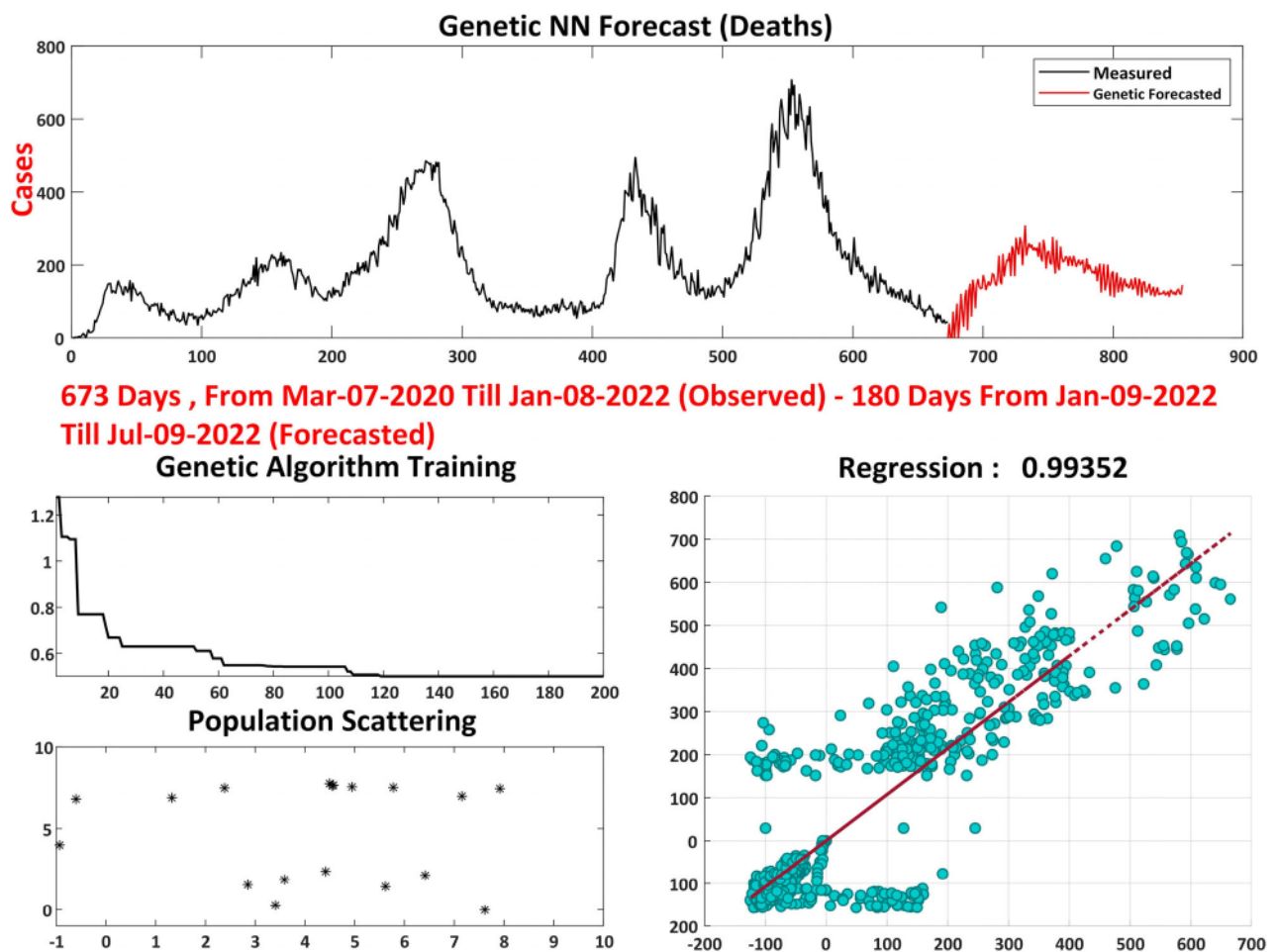


Figure 9. Forecasted result for “Death cases” using the GA + NN algorithm for Iran (next 180 days).

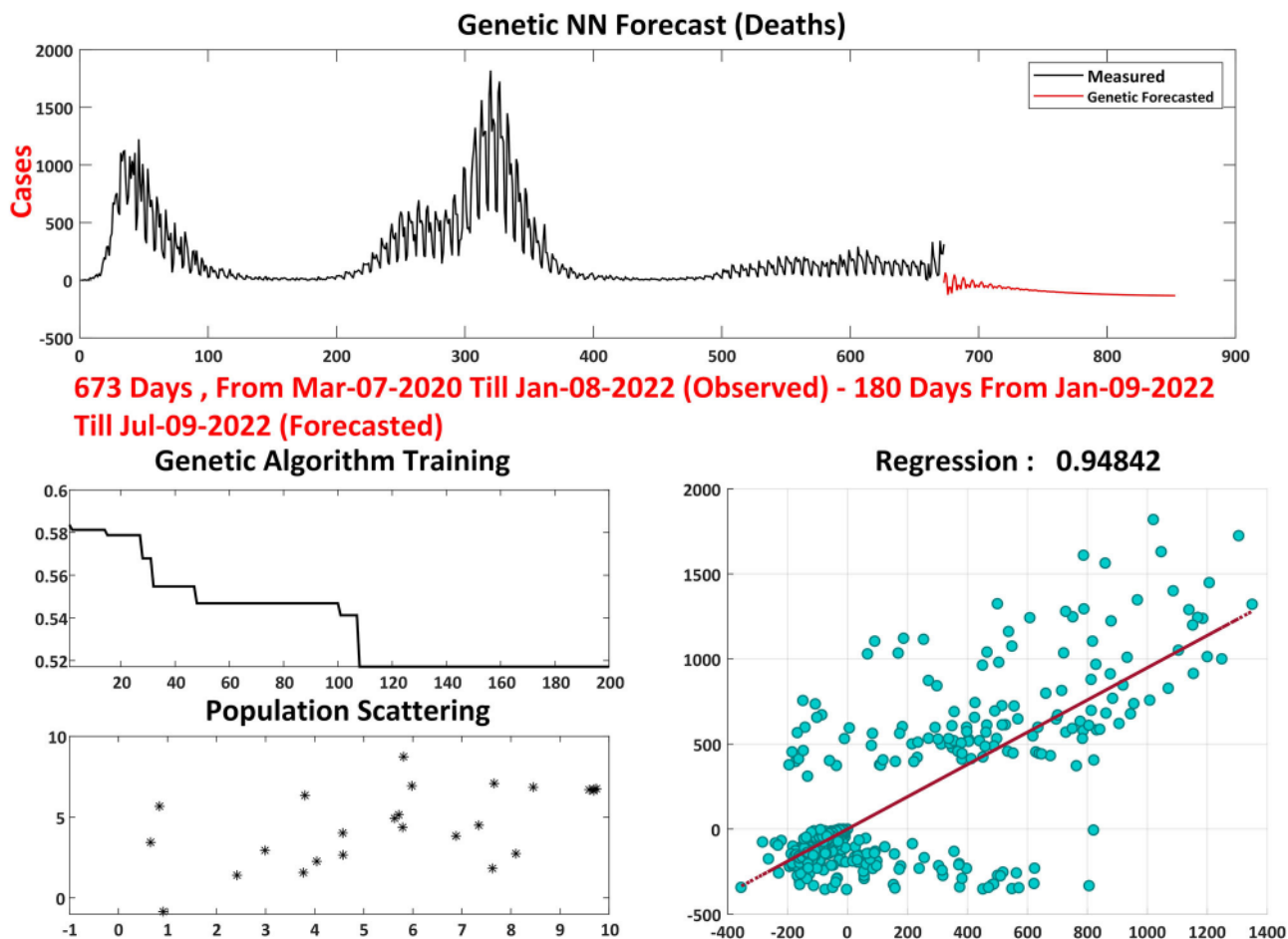


Figure 10. Forecasted result for “Death cases” using the GA + NN algorithm for the UK (next 180 days).

its assigned data. Figures 5 and 6 are of high practical relevance. By using the previous 690 days’ results of confirmed cases obtained using the GMDH algorithm, the next 180 days are shown to be very risky, particularly in the case of the UK. The orange lines show the next 180 days in these figures. At the end of the 180th day, it is shown that there are 100 new confirmed cases per day for Iran. In the case of the UK, however, the number is equal to the exorbitant value of 241,000 new confirmed cases per day. Figures 7 and 8 show the train, forecast, and error results for the next 180 days for the number of recovered cases, in both Iran and the UK. Fortunately, the ascending values for this variable represent good news and according to the achieved results, each day more patients recover from COVID-19; however, the results for the UK show a drop in recovered cases, although it finally picks up again. Lastly, in the case of death cases, as Figures 9 and 10 indicate, the values are presenting more optimistic horizons. The red line in Figures 9 and 10 shows the forecasted death cases for the next 180 days, in Iran and the UK, covering the time period from 09 Jan 2022 until 09 Jul 2022; and it can be observed that the numbers remain relatively low. It is to be noted that the average prediction interval, which represents the likelihood of a

new observation falling within a specific range for all cases, is 97% for the GMDH method, 98% for the Bi-LSTM method, and 97% for the GA + NN method. This indicates a high level of accuracy for all three methods.

Our study holds significant managerial relevance for various types of decisions in the context of COVID-19 and pandemics, more broadly. Unlike two- or three-month ahead forecasts, which enable short-term planning and tactical decision-making (e.g., anticipating demand for hospital beds and medical supplies, implementing targeted vaccination campaigns, and enhancing testing strategies in areas projected to experience a surge in cases), six-month ahead forecasts support strategic decision-making and planning, long-term resource allocation, and risk assessment and scenario planning. In this sense, six-month ahead forecasts enable healthcare systems to utilise this information to assess and enhance their resilience by developing contingency plans, adjusting staffing levels, and ensuring the availability of critical resources over an extended period. Governments, on the other hand, can rely on these forecasts to formulate policies related to travel restrictions, border control measures, and economic recovery plans, taking into account the projected trends in COVID-19 future infections, recovered

patients, and death cases. As mentioned, longer-term forecasts aid in risk assessment and scenario planning, too. Decision-makers can evaluate the impact of various scenarios and identify potential challenges and opportunities associated with the evolving pandemic by using six-month forecasts. These forecasts serve as the foundation for developing comprehensive risk management strategies, facilitating proactive decision-making, and reducing uncertainty in the face of a situation that is inherently unpredictable.

Naturally, the future is unpredictable, which means that the present forecasted results must be viewed with a critical eye. Nonetheless, more accurate forecasting of the number of confirmed, recovered, and death cases is crucial for optimising the available resources and slowing down or curbing the progression of the pandemic (Hansun et al., 2023). Moreover, forecasts can be used to motivate the general public to consider and follow the measures imposed by local and national authorities to slow down the spread of the pandemic. In this sense, we hope that the present paper can assist the range of stakeholders in their decision-making processes, helping to implement suitable actions that contain the spreading of COVID-19.

6. Conclusions and future research

Since its emergence, the COVID-19 pandemic has been exponentially spreading around the world, placing increased pressure on the healthcare systems in every country, but primarily in the most affected countries, among which we count Iran and the UK. Accurately forecasting the number of confirmed, recovered, and death cases provides useful information which can inform the measures that governments, policymakers, and decision-makers need to take to slow down or curb the spread of the virus. As Nikolopoulos et al. (2021, p. 109) elegantly stated, “COVID-19 forecasts provide indications and quantify the needs that appear in an emergency, and thus more research should be directed towards identifying the best forecasting models for all geographical contexts and temporal frequencies”. We echo these calls for further research into COVID-19 forecasting models.

Using NN-based algorithm forecasting systems has been shown to provide the most highly precise results possible, especially when a high amount of training data is available. In this research, NN-based algorithms, including GMDH, Bi-LSTM, and GA + NN, have been applied to the real-time data of the daily confirmed, recovered, and death COVID-19 cases, with an application to the cases of Iran and the UK. This choice was motivated by the

extended capacity of deep learning models in capturing process nonlinearity and their flexibility in modelling time-dependent data (Zeroual et al., 2020). The performance of each model has been verified in terms of CC, R-Squared value, MSE, RMSE, MAE, MAPE, EV, and RMSLE. Furthermore, as forecasting is about something that has not yet happened, a comparison with other methods is also of high importance. In this research, a comparison was provided with their immediate variants, *i.e.*, hybrid-GMDH, LSTM, and PSO-NN algorithms. In addition, two traditional and well-known algorithms, Exponential Smoothing and ARIMA were included in the comparisons to provide a better basis for understanding the advantages of using intelligent forecasting methods.

It is to be noted that the models used in this manuscript are all well-known ML models, which can understand the stochasticity of the data when various interventions (such as lockdowns, reopening, vaccination, etc.) are taking place. Hence, their prediction capacity is not affected, and we do not need to intervene in the models to indicate the interventions information. In other words, the models understand the dynamics/behaviour of the data. The employed methods returned acceptable results in all the tests, with a slight exception in the case of the PSO + NN algorithm, which showed a slightly better performance versus the counterpart method (GA + NN) for all the three categories of data. However, the difference is insignificant.

All in all, the main takeaway for policymakers is that the results show that the forecasted numbers of COVID-19 confirmed cases over the next 180 days in Iran are encouraging (Figure 5), while the ones for the UK are concerning (Figure 6). Our forecasts cover the time period up until 09 July 2022. Moreover, our results are corroborated by recent data. For Iran, we can see from Figure 5 that confirmed cases first seem to increase, but then decrease quite sharply until the end of the forecasted period. Actual data over the past five months show that, indeed, the country officially entered the sixth wave on 28 January 2022, with sharp increases in cases over the following month, after which the number of cases started to decrease, indicating the exit from the wave. For the UK, on the other hand, recent data show a dramatic surge in COVID-19 cases over Spring 2022, placing the country on the brink of a sixth wave. What this seems to indicate is that at the time of writing (one month before 09 July 2022), we are about one month away from a full-fledged new wave. And this poses substantial relevance for practice, as it would allow policymakers and healthcare providers to prepare and know where to deploy resources.

Another important takeaway, this time for data scientists and modelers of the pandemic, is that as highlighted in the introductory section of this paper, the countries of our choice (Iran and the UK) have found themselves in different waves at different times, and have applied different interventions; hence, the dynamics in the data varied between the two countries. Despite these different dynamics, however, our analysis shows that the robustness of the models used holds. For future research avenues, it is suggested to use additional nature-inspired algorithms, such as Bat algorithm or Differential Evolution algorithm to forecast the COVID-19 pandemic; this is because this virus is a nature-inspired phenomenon in itself. Furthermore, adding different types of data and even combining more datasets are also suggested for more precise forecasting. Another interesting avenue for future work on the topic would be to forecast “re-infection” or “re-confirmed” COVID-19 cases. Lastly, another future direction could be to adopt auto-ML procedures to evaluate all the potential algorithms available in the marketplace for their superiority and robustness for the given data, an exercise that was beyond the scope of the present paper.

Acknowledgements

The authors would like to thank the Editor, Associate Editor, and the anonymous reviewers for their very insightful comments and suggestions made on the previous versions of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Abdulaal, A., Patel, A., Charani, E., Denny, S., Mughal, N., & Moore, L. (2020). Prognostic modeling of COVID-19 using artificial intelligence in the United Kingdom: Model development and validation. *Journal of Medical Internet Research*, 22(8), e20259. <https://doi.org/10.2196/20259>
- Acter, T., Uddin, N., Das, J., Akhter, A., Choudhury, T. R., & Kim, S. (2020). Evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as coronavirus disease 2019 (COVID-19) pandemic: A global health emergency. *Science of the Total Environment*, 730, 138996. <https://doi.org/10.1016/j.scitotenv.2020.138996>
- Agbehadji, I. E., Awuzie, B. O., Ngowi, A. B., & Millham, R. C. (2020). Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing. *International Journal of Environmental Research and Public Health*, 17, 5330. <https://doi.org/10.3390/ijerph17155330>
- Al-Qaness, M. A. A., Ewees, A. A., Fan, H., & El Aziz, M. A. (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. *Journal of Clinical Medicine*, 9(3), 1–15. <https://doi.org/10.3390/jcm9030674>
- Asuero, A. G., Sayago, A., & González, A. G. (2006). The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry*, 36(1), 41–59. <https://doi.org/10.1080/10408340500526766>
- Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, 4(6), 588–596. <https://doi.org/10.1038/s41562-020-0898-6>
- Chang, J. T., & Kaplan, E. H. (2023). Modeling local coronavirus outbreaks. *European Journal of Operational Research*, 304(1), 57–68. <https://doi.org/10.1016/j.ejor.2021.07.049>
- Charles, V., Emrouznejad, A., Gherman, T., & Cochran, J. (2022). Why data analytics is an art. *Significance*, 19(6), 42–45. <https://doi.org/10.1111/1740-9713.01707>
- Chumachenko, O. I., Kot, A. T., & Mandrenko, A. E. (2020). Algorithm of hybrid GMDH-network construction for time series forecast. *Electronics and Control Systems*, 2(64), 24–31. <https://doi.org/10.18372/1990-5548.64.14852>
- De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496), 1513–1527. <https://doi.org/10.1198/jasa.2011.tm09771>
- Dezfoulian, M. H., MiriNezhad, Y., Mousavi, S. M. H., Mosleh, M. S., & Shalchi, M. M. (2016). Optimization of the Ho-Kashyap classification algorithm using appropriate learning samples [Paper presentation]. 2016 Eighth International Conference on Information and Knowledge Technology (IKT) (pp. 167–169). <https://doi.org/10.1109/IKT.2016.7777760>
- Dhama, K., Sharun, K., Tiwari, R., Dadar, M., Malik, Y. S., Singh, K. P., & Chaicumpa, W. (2020). COVID-19, an emerging coronavirus infection: Advances and prospects in designing and developing vaccines, immunotherapeutics, and therapeutics. *Human Vaccines & Immunotherapeutics*, 16(6), 1232–1238. <https://doi.org/10.1080/21645515.2020.1735227>
- Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet. Infectious Diseases*, 20(5), 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. Wiley.
- Fanelli, D., & Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons, and Fractals*, 134, 109761. <https://doi.org/10.1016/j.chaos.2020.109761>
- Farlow, S. J. (1984). *Self-organizing methods in modeling: GMDH-type algorithms*. M. Dekker.
- Ghanbari, B. (2020). On forecasting the spread of the COVID-19 in Iran: The second wave. *Chaos, Solitons, and Fractals*, 140, 110176. <https://doi.org/10.1016/j.chaos.2020.110176>
- Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks: The Official Journal of the International Neural Network*

- Society*, 18(5-6), 602–610. <https://doi.org/10.1016/j.neu-net.2005.06.042>
- Gu, J., Zhu, M., & Jiang, L. (2011). Housing price forecasting based on genetic algorithm and support vector machine. *Expert Systems with Applications*, 38(4), 3383–3386. <https://doi.org/10.1016/j.eswa.2010.08.123>
- Hansun, S., Charles, V., & Gherman, T. (2023). The role of the mass vaccination programme in combating the COVID-19 pandemic: An LSTM-based analysis of COVID-19 confirmed cases. *Heliyon*, 9(3), e14397. <https://doi.org/10.1016/j.heliyon.2023.e14397>
- Hansun, S., Charles, V., Gherman, T., Subanar, N. A., & Indrati, C. R. (2021). A tuned Holt-Winters white-box model for COVID-19 prediction. *International Journal of Management and Decision Making*, 20(3), 241–262. <https://doi.org/10.1504/IJMDM.2021.116018>
- Hansun, S., Charles, V., Gherman, T., & Vijayakumar, V. (2022). Hull-WEMA: A novel zero-lag approach in the moving average family, with an application to COVID-19. *International Journal of Management and Decision Making*, 21(1), 92–112. <https://doi.org/10.1504/IJMDM.2022.119582>
- Ho, S. L., & Xie, M. (1998). The use of ARIMA models for reliability forecasting and analysis. *Computers & Industrial Engineering*, 35(1-2), 213–216. [https://doi.org/10.1016/S0360-8352\(98\)00066-7](https://doi.org/10.1016/S0360-8352(98)00066-7)
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press.
- Ivakhnenko, A. G. (1966). Group method of data handling—A rival of the method of stochastic approximation. *Soviet Automatic Control*, 1(3), 43–55.
- JHU CSSE. (2022). COVID-19 data repository by the center for systems science and engineering (CSSE) at Johns Hopkins University. https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series
- Kalgotra, P., Gupta, A., & Sharda, R. (2021). Pandemic information support lifecycle: Evidence from the evolution of mobile apps during COVID-19. *Journal of Business Research*, 134, 540–559. <https://doi.org/10.1016/j.jbusres.2021.06.002>
- Kapoor, A., Ben, X., Liu, L., Perozzi, B., Barnes, M., Blais, M., & O'Banion, S. (2020). Examining COVID-19 forecasting using spatio-temporal graph neural networks. *arXiv preprint arXiv:2007.03113*.
- Katris, C. (2021). A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. *Expert Systems with Applications*, 166, 114077. <https://doi.org/10.1016/j.eswa.2020.114077>
- Kuepper, J. (2020). *Using genetic algorithms to forecast financial markets*. <https://www.investopedia.com/articles/financial-theory/11/using-genetic-algorithms-forecast-financial-markets.asp>
- Liu, C., Zhou, Q., Li, Y., Garner, L. V., Watkins, S. P., Carter, L. J., Smoot, J., Gregg, A. C., Daniels, A. D., Jerve, S., & Albaiu, D. (2020). Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. *ACS Central Science*, 6(3), 315–331. <https://doi.org/10.1021/acscentsci.0c00272>
- Mandal, M., Jana, S., Nandi, S. K., Khatua, A., Adak, S., & Kar, T. K. (2020). A model based study on the dynamics of COVID-19: Prediction and control. *Chaos, Solitons & Fractals*, 136, 109889. <https://doi.org/10.1016/j.chaos.2020.109889>
- Meng, X. (2012). Weather forecast based on improved genetic algorithm and neural network. In Z. Zhong (Ed.), *Proceedings of the International Conference on Information Engineering and Applications (IEA) 2012* (Vol. 219, pp. 833–838). *Lecture notes in electrical engineering*. Springer.
- Mousavi, S. M. H., Charles, V., & Gherman, T. (2020). An evolutionary Pentagon Support Vector finder method. *Expert Systems with Applications*, 150, 113284. <https://doi.org/10.1016/j.eswa.2020.113284>
- Mousavi, S. M. H., & Lyashenko, V. (2017, November). *Extracting old Persian cuneiform font out of noisy images (handwritten or inscription)* [Paper presentation]. 2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP) (pp. 241–246). IEEE. <https://doi.org/10.1109/IranianMVIP.2017.8342358>
- Mousavi, S. M. H., MiriNezhad, S. Y., & Mirmoini, A. (2017). *A new support vector finder method, based on triangular calculations and K-means clustering* [Paper presentation]. 9th International Conference on Information and Knowledge Technology (IKT 2017). Institute of Electrical and Electronics Engineers, Amirkabir University of Technology.
- Nielsen, M. A. (2015). *Neural networks and deep learning* (Vol. 2018). Determination Press.
- Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., & Vasilakis, C. (2021). Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *European Journal of Operational Research*, 290(1), 99–115. <https://doi.org/10.1016/j.ejor.2020.08.001>
- Petropoulos, F., & Makridakis, S. (2020). Forecasting the novel coronavirus COVID-19. *PloS One*, 15(3), e0231236. <https://doi.org/10.1371/journal.pone.0231236>
- Petropoulos, F., Makridakis, S., & Stylianou, N. (2022). COVID-19: Forecasting conformed cases and deaths with a simple time-series model. *International Journal of Forecasting*, 38(2), 439–452. <https://doi.org/10.1016/j.ijforecast.2020.11.010>
- Rizk-Allah, R. M., & Hassanien, A. E. (2020). COVID-19 forecasting based on an improved interior search algorithm and multi-layer feed forward neural network. *arXiv preprint arXiv:2004.05960*.
- Roosa, K., Lee, Y., Luo, R., Rothenberg, R., Hyman, J. M., Yan, P., & Chowell, G. K., A. (2020). Short-term forecasts of the COVID-19 epidemic in Guangdong and Zhejiang, China: February 13–23, 2020. *Journal of Clinical Medicine*, 9(2), 596. <https://doi.org/10.3390/jcm9020596>
- Salgotra, R., Gandomi, M., & Gandomi, A. H. (2020). Evolutionary modelling of the COVID-19 pandemic in fifteen most affected countries. *Chaos, Solitons, and Fractals*, 140, 110118. <https://doi.org/10.1016/j.chaos.2020.110118>
- Sarkar, K., Khajanchi, S., & Nieto, J. J. (2020). Modeling and forecasting the COVID-19 pandemic in India. *Chaos, Solitons, and Fractals*, 139, 110049. <https://doi.org/10.1016/j.chaos.2020.110049>
- Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11), 2673–2681. <https://doi.org/10.1109/78.650093>
- Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017). *Stock price*

- prediction using LSTM, RNN and CNN-sliding window model. Proceedings of the [Paper presentation]. 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1643–1647). <https://doi.org/10.1109/ICACCI.2017.8126078>*
- Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons, and Fractals*, 140, 110212. <https://doi.org/10.1016/j.chaos.2020.110212>
- Sokol, E. (2020, July 17). COVID-19 accelerating interoperability, data exchange, analytics. *EHR Intelligence*. <https://ehrintelligence.com/features/covid-19-acceleratinginteroperability-data-exchange-analy>
- Vaishnav, V., & Vajpai, J. (2020). Assessment of impact of relaxation in lockdown and forecast of preparation for combating COVID-19 pandemic in India using Group Method of Data Handling. *Chaos, Solitons, and Fractals*, 140, 110191. <https://doi.org/10.1016/j.chaos.2020.110191>
- WHO (2022). *WHO coronavirus disease (COVID-19) dashboard*. World Health Organization. <https://covid19.who.int/>
- Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons, and Fractals*, 140, 110121. <https://doi.org/10.1016/j.chaos.2020.110121>
- Zhang, G., & Liu, X. (2021). Prediction and control of COVID-19 spreading based on a hybrid intelligent model. *PloS One*, 16(2), e0246360. <https://doi.org/10.1371/journal.pone.0246360>