Review

Artificial Intelligence for Forecasting the Prevalence of COVID-19 Pandemic: An Overview

Ammar H. Elsheikh [1,*], **Amal I. Saba** [2], **Hitesh Panchal** [3], **Sengottaiyan Shanmugan** [4], **Naser A. Alsaleh** [5] **and Mahmoud Ahmadein** [5]

1 Faculty of Engineering, Tanta University, Tanta 31527, Egypt
2 Faculty of Medicine, Tanta University, Tanta 31527, Egypt; amal.saba@med.tanta.edu.eg
3 Department of Mechanical Engineering, Government Engineering College, Patan 384265, Gujarat, India; engineerhitesh2000@gmail.com
4 Research Centre for Solar Energy, Department of Physics, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, Andhra Pradesh, India; s.shanmugam1982@gmail.com
5 Mechanical Engineering Department, Imam Mohammad Ibn Saud Islamic University, Riyadh 11564, Saudi Arabia; naalsaleh@imamu.edu.sa (N.A.A.); maahmadein@imamu.edu.sa (M.A.)
* Correspondence: ammar_elsheikh@f-eng.tanta.edu.eg

Abstract: Since the discovery of COVID-19 at the end of 2019, a significant surge in forecasting publications has been recorded. Both statistical and artificial intelligence (AI) approaches have been reported; however, the AI approaches showed a better accuracy compared with the statistical approaches. This study presents a review on the applications of different AI approaches used in forecasting the spread of this pandemic. The fundamentals of the commonly used AI approaches in this context are briefly explained. Evaluation of the forecasting accuracy using different statistical measures is introduced. This review may assist researchers, experts and policy makers involved in managing the COVID-19 pandemic to develop more accurate forecasting models and enhanced strategies to control the spread of this pandemic. Additionally, this review study is highly significant as it provides more important information of AI applications in forecasting the prevalence of this pandemic.

Keywords: artificial intelligence; review; COVID-19; forecasting

1. Introduction

COVID-19 is a global pandemic that has been rapidly spreading worldwide [1]. It has affected more than 100 million people and caused about two deaths per million as of the end of January 2021. Originating in Wuhan, China, it has spread to about 200 countries. As there is no treatment to this disease, the main way to control this pandemic is by applying strict quarantines and lockdown procedures [2]. Forecasting of the COVID-19 spread has attracted the attention of many scientists, researchers and policy makers, as it helps in developing suitable plans and applying protective measures to combat the spread of the pandemic [3].

Longer-term forecasting of COVID-19 prevalence plays a critical role in making pivotal decisions such as the potential to open universities, schools, houses of worship and workplaces. It may also help to define the precautions and restrictions for citizens to reduce disease transmission. These restrictions should be carefully investigated to minimize their impact on different economic outcomes, such as poverty and unemployment [4]. Moreover, forecasting of COVID-19 prevalence may help in establishing vaccine and therapy plans and selecting hot spots that that should receive more attention from policy makers. Nonetheless, numerous forecasting approaches have been reported in the last two years to predict the spread of the COVID-19 pandemic. Mathematical and statistical based approaches have been widely used to predict the spread of COVID-19 in many countries and regions [5,6]. Apart from statistical and computational forecasting models used to

model the spread of COVID-19 pandemic, AI models have recently excelled [7,8]. The outperformance of the AI models over other conventional models has been reported in several investigations [9–13]. This motivated us to prepare this literature review to shed light on the application of AI tools to predict the spread of this pandemic. In this paper, the commonly used models will be briefly explained. Then, the evaluation criteria of the forecasted results will be presented. Finally, a comparative study and discussion of different published articles as well as the future prospects will be discussed.

2. Artificial Intelligence Models

The commonly used AI models utilized in modeling the COVID-19 pandemic are introduced in this section. These models are nonlinear autoregressive ANN (NARANN), adaptive neuro-fuzzy inference system (ANFIS), hybrid fractal-fuzzy approach (HFFA), Bayesian neural network (BNN), long short-term memory network (LSTM), variational autoencoder (VAE) and singular spectrum analysis (SSA). The fundamental concepts and mathematical modeling of each approach are briefly explained.

2.1. Nonlinear Autoregressive ANN

Nonlinear autoregressive ANN (NARANN) is a special of ANN family that mimics the behavior of biological nervous system. It has been widely used in forecasting of one-dimensional time series for different engineering systems such as industrial production [14], fluctuation of groundwater level [15] and solar radiation [16]. Due to its nonlinear nature, it succeeded to model time series data that has transient nature and high variability [17]. The mathematical formulation of NARNN can be expressed as follows:

$$\xi(t) = \beta(x(t-1), \xi(t-2), \ldots, \xi(t-e)) + \delta(t) \tag{1}$$

where $\xi(t)$ is the response of time series at time t, and e denotes the delay of the network. $\xi(t-1), \xi(t-2), \ldots, \xi(t-e)$ represent the historical responses. $\beta(.)$ is a function with unknown value and it is computed during the training stage by obtaining the optimal bias and weights. $\delta(t)$ is a time compensation function employed to estimate the error with respect to time.

The typical structure of NARANN is presented in Figure 1. It contains an input layer followed by one or multiple hidden layers and an output layer. The number of processing units in the hidden layers and the synaptic weights between them could be adjusted to maximize the forecasting accuracy of the network by obtaining the optimal structure of NARANN. In Figure 1, w_1 and w_2 are the synaptic weights and y_t is the model out output. Increasing the processing units in the hidden layer requires much computational costs and may cause inappropriate over-fitting during a training process, while decreasing the processing units in the hidden layer may result in reducing the model accuracy as well as its generalization capabilities. As with any other ANN, NARANN is trained using real-world data using any training algorithm such as Levenberg-Marquardt backpropagation training algorithm. Once the network is trained, it may be used to predict the behavior of the time series in the future.

2.2. Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is a machine learning tool in which ANN is incorporated with fuzzy logic. It has been used to predict the response of different engineering systems such as drought forecasting [18], tropical cyclones [19] and oil consumption forecasting [20]. It figures out the knowledge between input and output process variables by employing If-Then fuzzy rules.

Each applied rule has antecedent part defined as fuzzy inputs and a consequence part depends on crisp input variables. Fuzzy inputs are combined with each other using

logistic mathematical operator (AND operator). The fuzzy rule F^r of a logic system that has n input variables is defined as

$$F^r : if \ x_1 \ is \ B_1^r, x_2 \ is \ B_2^r , \ldots, x_i \ is \ B_i^r, \ldots and \ x_n \ is \ B_n^r \rightarrow \qquad (2)$$

$$then \ y^r = b_0^r + b_1^r x_1 + \ldots + b_i^r x_i + \ldots + b_n^r x_n$$

where x_i and y^r denote the input variable number i and fuzzy output number r, respectively; B_i^r denotes the membership function, b_0^r and b_i^r denote the bias and the regression coefficient, respectively.
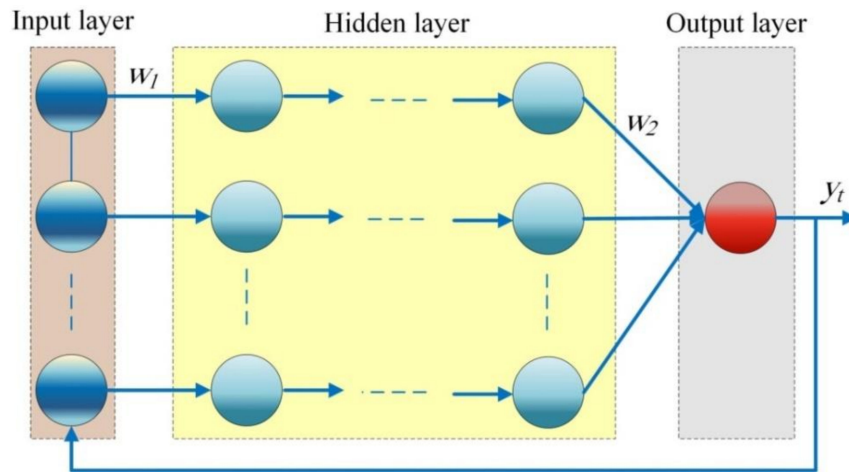
Figure 1. The typical structure of NARANN.

In a typical ANFIS structure, there are five consequence layers as shown in Figure 2, the tasks of each layer are explained as follows.
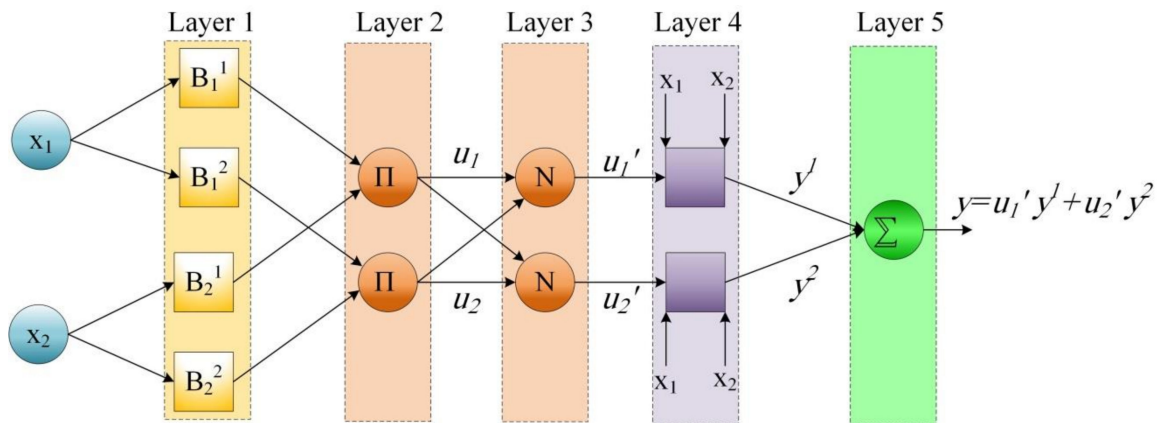
Figure 2. A typical structure of ANFIS.

In the first layer, membership degrees for each input variable are computed, thus this layer acts as fuzzification layer. Before calculating the membership degrees, the regression coefficients, membership functions and the fuzzy rules must be defined. After defining the primary fuzzy rules, the training process of the ANFIS model is executed. During the training process, the prediction error of the model is minimized via optimizing regression coefficients and membership functions. Two optimization algorithms are utilized to accomplish the optimization process; namely, least-squares algorithm and gradient descend algorithm. The first algorithm is used to optimize the regression coefficients, while the used second algorithm is employed to optimize the internal parameters of the membership functions.

In the subsequent layer, after computing the membership degrees, the antecedent part of fuzzy rules is subjected to the following firing function:

$$u_r = \prod_{i=1}^{n} B_i^r(x_i) \tag{3}$$

In the subsequent layer, the normalized weights are computed by

$$u_r' = \frac{u_r}{\sum_r u_r} \tag{4}$$

In the fourth layer, the consequence value of the applied rules is computed based on the defined regression coefficients as follows:

$$y^r = b_0^r + \sum_{i=1}^{n} b_i^r x_i \tag{5}$$

In the fifth layer, the final output of ANFIS is computed as follows:

$$y = \sum_r u_r' y^r \tag{6}$$

2.3. Hybrid Fractal-Fuzzy Approach

Hybrid fractal-fuzzy approach (HFFA) us a hybrid approach in which fuzzy logic is integrated with fractal theory. The fuzzy logic based system is used to express the time series knowledge. Fractal theory has shown promising applications in modeling complex engineering problems [21,22]. The main concepts of fuzzy logic and fractal dimension will be discussed in this section.

The fractal dimension of a certain object could be expressed as

$$b = \lim_{a \to 0} [lnE(a)] / [ln(1/a)] \tag{7}$$

where $E(a)$ is the number of boxes employed to shield an object and a denotes a measure parameter of the box size. The fractal dimension is computed as a function of the number of boxes used to shield the object boundary for different a sizes and then b is obtained using logarithmic regression. The box dimensions could be estimated using the following equation:

$$lnE(a) = ln\beta - blna \tag{8}$$

where b denotes the fractal dimension, which could be estimated using a least squares method.

Fuzzy rules could be used as a forecasting tool. First, the input space or time series should be partitioned so that the geometrical objects are distinguished based on their characteristics. Considering the geometrical objects as time series patterns, fuzzy clustering techniques could be applied to cluster the data, then fuzzy rules are constructed. These fuzzy rules will act as a forecasting tool for a certain time series application.

For m objects $OB_1, OB_2, \ldots OB_m$, clustering using fuzzy technique can be used to obtain pairs (u_i, v_i) for all m. Then a fuzzy rule is defined by

$$If\ U\ is\ u_1\ and\ V\ is\ v_1\ then\ Object\ is\ OB_1$$
$$If\ U\ is\ u_2\ and\ V\ is\ v_2\ then\ Object\ is\ OB_2$$
$$\ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots$$

$$If\ U\ is\ u_m\ and\ V\ is\ v_m\ then\ Object\ is\ OB_m \tag{9}$$

These fuzzy rules can be used in time series prediction.

In case of executing time series forecasting, data analysis should be carried out to explore the periodicities and trends within the data. Then, the time series is clustered into m objects $OB_1, OB_2, \ldots OB_m$. Objects $OB_1, OB_2, \ldots OB_m$ is classified based on their geometrical complexity into L or N for linear and non-linear, respectively. Assume the time

series has the form $v_1, v_2, \dots v_m$. Then, the fuzzy rules are expressed in a general form for time series forecasting as follows:

$$If \ L \ is \ u_1 \ and \ N \ is \ v_1 \ then \ Object \ is \ OB_1$$
$$If \ L \ is \ u_2 \ and \ N \ is \ v_2 \ then \ Object \ is \ OB_2$$
$$\dots \dots \dots \dots \dots$$

$$If \ L \ is \ u_m and \ N \ is \ v_m \ then \ Object \ is \ OB_m \tag{10}$$

Consequently, the membership functions are defined for the two different fractal dimensions, namely, linear and non-linear.

2.4. Long Short-Term Memory Network (LSTM)

LSTM is a well-known recurrent neural network (RNN). Like any other RNN, LSTM has many identical modules connected with each other to form a chain architecture, which is utilized as an artificial memory to process and store data [23]. LSTM is widely used in forecasting time series because of its deep learning capabilities that enable it to memorize and learn sequential evidence. It has superior advantages over other RNN models as it overcomes the main problems of RNN, such as vanishing/exploding gradient and instability. Moreover, LSTM has a good ability to detect patterns in time series, such as trends, autocorrelations, seasonality, and noise. It succeeded to predict different engineering time series such as explosibility of underground mines [24], electricity load [25], soil behavior [26], water yield of solar stills [27] and air pollution [28].

LSTMs can learn dependencies found in data and can remember huge amounts of information. A typical LSTM model consists of multiple modules connected in a chain manner. LSTM network has a combined short-term and long-term memory and has the ability to preserve cell state. The key parameter of the LSTM network is the memory cell. A typical LSTM cell, which acts as a fundamental network unit is shown Figure 3. The existence of input, output and forgot gates enhances the learning capabilities of LSTM.
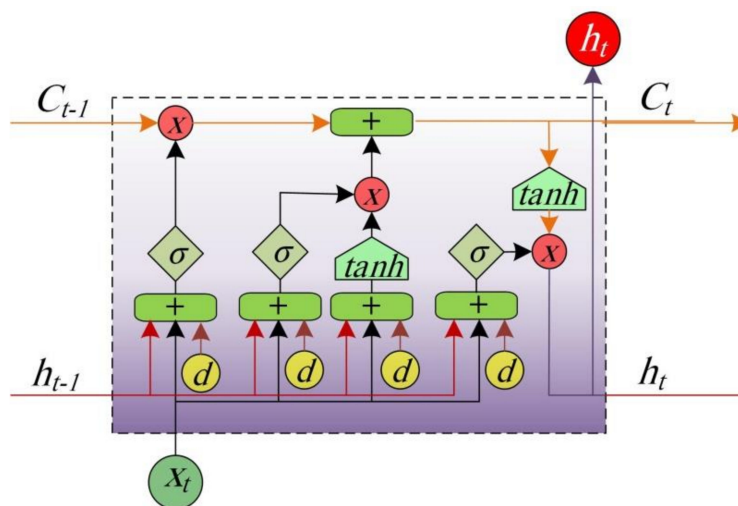

Figure 3. A typical LSTM cell.

LSTM network identifies the trivial evidence that may be omitted from the processed data. The importance of certain information within the data is identified via passing the data through a sigmoid activation function, which acts as a forget gate as it removes the trivial part of the processed data. This function has an output f_t defined as follows:

$$f_t = \sigma\left(U_f[h_{t-1}, X_t] + d_f\right) \tag{11}$$

where σ is the sigmoid function, d_f is the applied bias at the gate and U_f is the weight matrix of the gate.

The input Y_t of the following module is passed through two activation functions: tanh and sigmoid. These functions make a decision on saving or omitting information from the data. The sigmoid function makes a decision regarding excluding or updating the new input (0 or 1). The *tanh* function defines the updated weight of the input and assigns it a significance index with numerical values between -1 and 1. The computed output of the two functions produces the new cell state C_t by adding it to the old memory to C_{t-1}. These calculations are performed as follows:

$$S_t = \sigma(U_i[h_{t-1}, X_t] + d_i) \tag{12}$$

$$T_t = tanh(U_n[h_{t-1}, X_t] + d_n) \tag{13}$$

$$C_t = C_{t-1}f_t + T_tS_t \tag{14}$$

where, C_{t-1} and C_t denote the cell states at time $t-1$ and t, respectively, d_n denotes the applied bias at the gate, and U_n is the weight matrix of the gate.

The final output O_t is given by;

$$O_t = \sigma(U_o[h_{t-1}, X_t] + d_o) \tag{15}$$

where, d_o is the applied bias at the gate, and U_o is the weight matrix of the gate.

2.5. Bayesian Neural Network (BNN)

BNN is an advanced ANN integrated with an inference system, which is used in many applications such as electricity load forecasting [29], wind power forecasting [30] and hydrologic forecasting [31]. It overcomes the problems related to conventional ANN, such as over-fitting caused by the uncertainty of synaptic weights estimation. The Bayesian-based training algorithm may help in solving these problems. It applies a probability density function on the network weight space. This function is used to express the belief degree of the weight vector. It is initialized by equating it with a prior distribution. Then, the data is processed via the application of Bayes' theorem. Finally, the function is converted into a new posterior distribution. Therefore, instead of computing the best weights by conventional approaches via executing a minimization process of a defined error function, the Bayesian approach estimates a whole distribution for the network parameters. The determined posterior distribution is utilized as an inference prediction tool to forecast the new responses of the new input variables.

The training process of BNN is presented in Figure 4. The key target of the training process is to reduce the error between the predicted values and the real world values. This minimization process is defined by applying some rules: $B(x)$, $U_O(x)$ and $U_H(x)$ until the stopping criteria E_r is fulfilled. This process could be represented mathematically un the following form:

$$L(V_I) = L(V_I) + min(f_O(f_H(V_I.U_H + b_H).U_O + b_O) - L(V_I)) \tag{16}$$

where $L(V_I)$, U_H, U_O, b_H and b_O denote the real value, the matrix of hidden layer weights, the matrix of output layer weights, the vector of bias of the hidden layer and the vector of bias of the output layer, respectively. $f_H(.)$ and $f_O(.)$ are the hidden and output functions, respectively, which are given by

$$f_H(x) = \frac{2}{1 + \exp(-2x)} - 1 \tag{17}$$
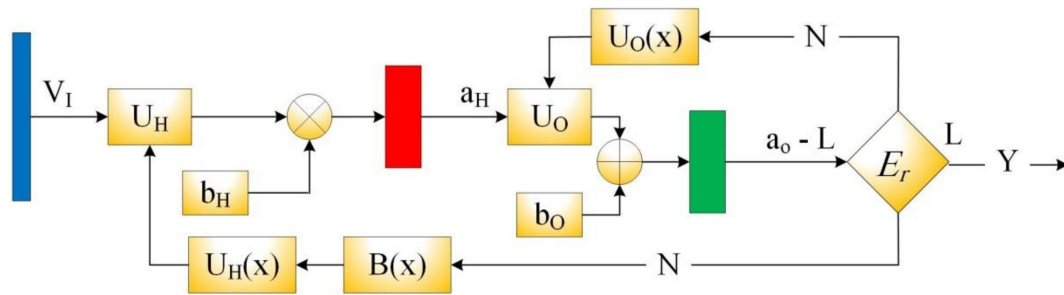
$$f_O(x) = x \tag{18}$$

Figure 4. The training of BNN.

Once the forecasting model is established, it may be used to forecast the future behavior of time series using the following equation:

$$L(IV) = f_O(f_H(IV.U_H + b_H).U_O + b_O) \tag{19}$$

where IV denotes the latest inputs values in a vector form and $L(IV)$ denotes the forecasted vector.

2.6. Variational AutoEncoder (VAE)

VAE is a machine learning technique that exploits the advantage of learned inference and could be established using gradient-based methods. The term variational represents the variational inference technique, which belongs to statistics-based methods. The term auto-encoder represents the unsupervised ANN model used to learn data coding. VAE is an auto-encoder model with enhanced training algorithm, which overcomes over-fitting problem by using a special regularization technique. It also creates a latent space with appropriate characteristics that permit the executive of efficient generative process. The basic architecture VAE is shown in Figure 5. VAE has a decoder (shown in green in the figure) and an encoder (shown in blue in the figure) like any other conventional auto-encoder. During the training of the model, the reconstruction error between the input data and the reconstructed encoded–decoded data is minimized. The input data is passed through the encoder $E(y|x)$ via the latent space. At that time, $y \sim E(y|x)$, is created from the code distribution. After that, the coded point is decoded by passing through the decoder $D(x|y)$. Finally, the reconstruction error is computed and the network is updated via a back-propagation process.
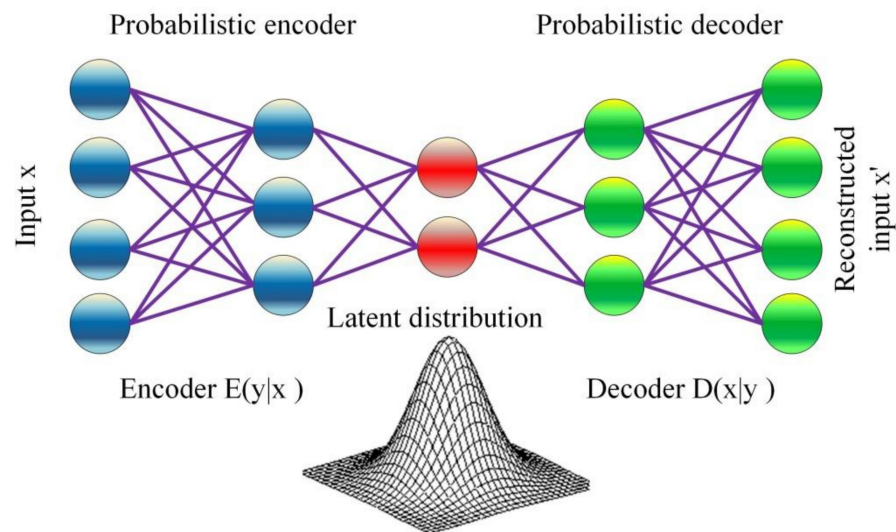

Figure 5. The architecture of VAE.

The training process is accomplished via minimizing a loss function $f(e)$ that consists of a reconstruction term and a regularization term. The first term is used to enhance the encoding–decoding process, while the second term is used to regularize the latent space structure. The latent distribution is chosen to match a standard distribution, such as Gaussian distribution. The loss function is given by:

$$f(e) = G_{y \sim E(yx)}(\log(xy)) - KL(E(y \mid x) \mid\mid D(y)) \tag{20}$$

The reconstruction term $G_{y \sim E(yx)}(\log(xy))$ enables the decoder to learn the reconstruction of the training data. The regularization term is expressed as a function of the Kulback–Leibler $KL(.)$ divergence between the distribution of the encoder $E(y \mid x)$ the latent sample y, $|D(y)$. The $KL(.)$ measures the loss when E is used to represent D.

2.7. Singular Spectrum Analysis (SSA)

SSA is an advanced time series analysis tool, which is classified as a nonparametric spectral estimation technique [32]. It has been used in forecasting different engineering time series such as annual precipitation and hourly water temperature [33], mutual investment funds [34], wind speed [35] and electricity consumption [36]. In a typical SSA model, a conventional time series analysis is integrated with signal processing, multivariate geometry, multivariate statistics and dynamical systems. It has two main phases. The first is the decomposition phase in which the time series is broken down into several components, such as cyclical, seasonal and trend components. This phase facilities the extraction of signals and a reduction of related noise. The second is the reconstruction phase in which data forecasting is accomplished via reconstructing the decomposed components.

The decomposition phase is accomplished via executing two steps: embedding and singular value decomposition. During the embedding step, the time series $\{y_1 \ y_2 \ldots y_M\}$ is converted into a trajectory matrix Y with dimension $N \times L$. Each column in this matrix has the form of $Y_1 = (y_i \ y_{i+1} \ \ldots \ y_{i+N-1})$, where $L = M - N + 1$. During the decomposition, the established trajectory matrix Y is decomposed into $Y = \Psi\Omega\Phi^T$ where Ψ and Φ are two orthogonal matrices and Ω is a diagonal matrix. The diagonal elements of Ω are $\tau_1 \ \tau_2 \ldots \tau_N$, which represent the singular values of Y. The decomposed trajectory matrix is given by

$$Y = Y_1 + Y_2 + \ldots + Y_d \tag{21}$$

where d is the maximum i such that $\tau_i \geq 0$ and $Y_i = \tau_i + \Psi_i\Phi_i^T$.

The reconstruction phase is accomplished via executing two stages: grouping and diagonal averaging. In the grouping stage, the elementary matrices Y_i is divided into many groups and the generated matrices within each new group are summed. For a certain group I, a new matrix Y_I is formed and defined as $Y_I = Y_{i1} + Y_{i2} + \ldots + Y_{ip}$. By applying singular value decomposition Y, the following decomposition is obtained:

$$Y = Y_{I1} + Y_{I2} + \ldots + Y_{IM} \tag{22}$$

In the diagonal average stage, each decomposed matric Y_{I_j} is transformed into a Hankel matrix. The transformed matrix can be transformed into a time series form with a length M. To obtain the following expansion $Y = \widetilde{Y}_{I_1} + \widetilde{Y}_{I_2} + \ldots. + \widetilde{Y}_{I_M}$ with $\widetilde{Y}_{I_i} = HY_{I_i}$ with new elements $\widetilde{y}_s$ over the anti-diagonals computed by $\widetilde{y}_s = \Sigma_{(l,n) \in Y_s} \frac{y_{ln}}{|Y_c|}$, where $Y_s = \{(l, n) : l + n = c + 1, 1 \leq l \leq L, 1 \leq n \leq N\}$ and $|Y_c|$ is the number of elements in Y_c.

The forecasting process could be accomplished with two main processes, namely, recurrent forecasting and vector forecasting.

The recurrent forecasting is executed according to the following procedures:

1. The new time series $X_{M+h} = \{x_1, x_2, \ldots., x_{M+h}\}$ is defined as follows:

$$x_i = \begin{cases} \widetilde{Y}_i, \ for \ i = 1, 2, \ldots., M \\ \sum_{j=1}^{N-1} e_j x_{i-j}, \ for \ i = M+1, M+2, \ldots., M+h \end{cases} \tag{23}$$

where the coefficients vector $V = (e_{N-1}, e_{N-2}, \ldots, e_{1,})^T$ is computed as follows:

$$V = \frac{1}{1 - \varphi^2} \Sigma_{i \in I} \pi_i \Psi_i \tag{24}$$

2. The numbers $x_{L+1}, x_{L+2}, \ldots, x_{L+h}$ represent the recurrent forecast with an h step.

The vector forecasting is executed according to the following procedures:

1. Define a new time series vector U as follows;

$$U_i = \begin{cases} \check{Y}_i, \ for \ i = 1, 2, \ldots, \ L \\ YU_{i-1}, \ for \ i = L+1, L+2, \ldots, \ L+h+M-1 \end{cases} \tag{25}$$

where $\check{Y}_i = \Sigma_{j \in I} \Psi_j \Psi_j^T Y_i$, $YU_{i-1} = \begin{pmatrix} \Gamma \overline{U}_{i-1} \\ V^T \overline{U}_{i-1} \end{pmatrix}$, $\Gamma = \Phi \Phi^T + (1 - \varphi^2) YY^T$, and $\overline{U}_{i-1}$ represent a vector of U_{i-1} last components.

2. The time series $\{x_1, x_2, \ldots, x_{M+h+N-1}\}$ is obtained by averaging the diagonal of constructed matrix $U = [U_1 : \ldots : U_{L+h+M-1}]$.
3. The numbers $x_{L+1}, x_{L+2}, \ldots, x_{L+h}$ represent the vector forecast with an h step.

3. Evaluation Criteria

The statistical metrics are used as indicators to evaluate the performance of AI models. The most common used statistical metrics are root mean square error (RMSE), normalized root-mean-square error (NRMSE), coefficient of variation (COV), coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), overall index (OI), efficiency coefficient (EC), mean relative error (MRE), and residual mass coefficient (RMC) [37].

As the numerical value of RMSE, NRMSE, MRE, MAE, MAPE, RMC and COV being a small, the discrepancy is small and the model accuracy increases. They can be calculated as

$$RMSE = \sqrt{\frac{1}{n_s} \Sigma_{i=1}^{n_s} (d_i - y_i)^2} \tag{26}$$

$$NRMSE = \frac{\sqrt{n_s \Sigma_{i=1}^{n_s} (d_i - y_i)^2}}{\Sigma_{i=1}^{n_s} y_i} \tag{27}$$

$$MRE = \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{d_i - y_i}{d_i} \tag{28}$$

$$MAE = \frac{1}{n_s} \sum_{i=1}^{n_s} |d_i - y_i| \tag{29}$$

$$MAPE = \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{|d_i - y_i|}{|d_i|} \tag{30}$$

$$RMC = \frac{\Sigma_{i=1}^{n_s} y_i - \Sigma_{i=1}^{n_s} d_i}{\Sigma_{i=1}^{n_s} d_i} \tag{31}$$

$$COV = \left(\left(\frac{RMSE}{\overline{y}} \right) \times 100 \right) \tag{32}$$

As the value of R^2, OI and EC increases (approaches the unity), the model has high accuracy. They are calculated by

$$R^2 = \frac{\left(\sum_{i=1}^{n_s}\left(d_i - \overline{d}\right)(y_i - \overline{y})\right)^2}{\sum_{i=1}^{n_s}\left(d_i - \overline{d}\right)^2 \times \sum_{i=1}^{n_s}(y_i - \overline{y})^2} \tag{33}$$

$$EC = 1 - \frac{\sum_{i=1}^{n_s}(d_i - y_i)^2}{\sum_{i=1}^{n_s}\left(d_i - \overline{d}\right)^2} \tag{34}$$

$$OI = \frac{1}{2}\left(1 - \left(\frac{RMSE}{d_{max} - d_{min}}\right) + EC\right) \tag{35}$$

where, n_s, d, and y denote the number of iterations, actual and the forecasted data, respectively. d_{min}, d_{max}, and $\overline{d}$ represent the minimum, maximum and mean values of the target data, while $\overline{y}$ represents the mean value of the predicted dataset.

4. Comparative Study and Discussion

In this section, the published studies on the applications of AI techniques for forecasting COVID-19 spread are discussed. The reported comparisons between different AI techniques as well as conventional statistical techniques are also presented.

4.1. Statistical Models

Many forecasting studies have been carried out by employing statistical and mathematical approaches to predict the prevalence of COVID-19 pandemic. An auto-regressive integrated moving average (ARIMA) model was employed to predict COVID-19 spread for one month in Pakistan [38]. The increase rate of confirmed cases, recoveries and deaths has an exponential trend. The same approach was applied to predict the COVID-19 cases in South Korea and Japan [39]. A hybrid ARIMA model incorporated with wavelet-based approach was developed to predict the daily cases in India, South Korea, Canada, the United Kingdom and France on a real-time basis [40]. A high correlation between the mortality rate and four control variables, namely, total number of cases, age, lockdown period and total number of hospitals have been reported. Another hybrid model composed of discrete wavelet and ARIMA model was developed to predict the pandemic spread in Spain, Italy, France, the United Kingdom, and United States [41]. A modified version of ARIMA incorporated with α-Sutte indicator and called SutteARIMA has been developed to predict the total cases of COVID-19 in Spain [42] and another modified version incorporated with Kalman filter has been developed to predict the total active cases, deaths and recoveries in Pakistan [43]. A comparison investigation has been carried out between ARIMA model, support vector regression (SVR), ridge regression, random forest, cubist regression and stacking-ensemble learning as statistical-based models to assess their capabilities to predict total confirmed cases in Brazil [44]. Support vector regression and stacking-ensemble learning showed the best forecasting accuracy among the other investigated models. Gaussian spreading hypothesis have been employed to predict the epidemiological behavior in eight countries, namely, Greece, Germany, Netherlands, Spain, Italy, France, the United Kingdom and the United States based on the historical data reported by the Chinese government [45]. It has been reported that applying strict measures has a significant role in reducing the disease spreading. Reduced-space Gaussian regression modeling was applied to predict the spreading of the pandemic in United States via applying a hybrid incorporation between Bayesian system and geographic information system [46]. Another mathematical-based epidemiological model called susceptible/infectious/recovered model was employed to predict the transmission dynamics of the COVID-19 outbreak in China, Italy, Spain, Russia, Germany, France, South Korea and the United States [47]. This model is used to predict

the infected cases as a time series, disease spread, pandemic duration, and the disease reproductive number.

The dynamic epidemiological behavior of the COVID-19 outbreak in three Chinese cities, namely, Beijing, Shenzhen and Guangzhou has been investigated using the ecological niche models [48]. Nine socioeconomic factors have been considered, namely, population density, pandemic data of infected communities, number of subway stations, number of bus stops, total length of roads, number of supermarkets, number of shopping malls, number of hospitals, number of appointed hospitals and average prices of rental houses. It was recommended that local hygienic authorities should adjust their interventions in the COVID-19 hotspots with high population density.

Grey prediction models have also been employed to predict the spread of COVID-19 pandemic [49]. A comparison investigation has been carried out between three grey models, namely, grey model, fractional nonlinear grey Bernoulli model, and nonlinear grey Bernoulli model, to evaluate their accuracy to predict the total number of cases in Italy, the United Kingdom and the United States. Fractional nonlinear grey Bernoulli model showed the best prediction accuracy compared to the others.

COVID-19 infected cases was predicted by fitting asymptotic distributions to the reported data [50]. The proposed approach was employed to compare between the pandemic behavior of the outbreak in Italy and China. Death rate due to the pandemic in Italy was predicted using patient information based algorithm [51]. A new iterative method was proposed to predict daily growth rate of confirmed cases in four countries (United States, Germany, Slovenia, and Iran) based on the reported cases, recoveries and deaths [52]. Exponential approach was proposed to predict the total cases outside China; a heuristic computing approach was employed to determine exponential curve parameters.

The total reported cases in Nigeria was estimated using different statistical models, namely, simple linear, quadratic, quartic, cubic, logarithm linear, logarithm quadratic, logarithm quartic, logarithm cubic and inverse linear regression models [53]. The quartic regression model was reported as the best accurate model, among others. The model parameters were calculated using four different approaches, namely, least absolute deviation, Cochrane Orcutt, Hildreth–Lu, ordinary least squares, and Prais-Winsten. The least absolute deviation approach was reported as the best one among others in terms of accuracy, robustness and speed.

Two mathematical models were developed to determine the variations in the daily reported cases [54]. In the first model, the confirmed cases were computed based on the total number of active cases, the number of infected persons people were in close contact with, the rate of performed tests and the mortality rate. In the second model, the Fourier decomposition technique was employed to decompose the COVID-19 time-series in terms of cosine and sine functions. The models were applied to predict the total cases and its peak value in United States, Italy, and India. The SIDARTHE dynamical model was employed to model COVID-19 pandemic [55]. The population of infection is divided into five population subcategories, namely, infected, diagnosed, ailing, recognized and threatened. The effects of the employed control strategies, such as vaccination and different treatment strategies on the infection population subcategories, were investigated.

However, there are many statistical approaches reported in the literature to predict COVID-19 spread. Most of these approaches have been reported to have lower accuracy compared with AI approaches based on different statistical indicators. Moreover, most of approaches are valid only for short-term forecasting. On the other hand, AI approaches succeeded to forecast the COVID-19 spread for both short and long terms.

4.2. Artificial Intelligence Models

AI models have been rapidly developed to predict COVID-19 spread. The commonly used AI models utilized in modeling the COVID-19 pandemic are introduced in Section 2. These models are NARANN, ANFIS, HFFA, LSTM, BNN, VAE and SSA. ANFIS has been applied to predict the future number of COVID-19 cases in the United Kingdom [56] and

Malaysia [57]. In the first study, a parametric comparison has been carried out to obtain the optimal ANFIS model based on model accuracy considering different model parameters. To enhance the forecasting accuracy of ANFIS, it has been integrated with a virus optimization algorithm to forecast COVID-19 cases in the United States [58]. Virus algorithm gas been used to optimize the membership functions and the regression coefficients instead of conventional optimization techniques, which have tendency to be trapped into local optima. The effects of population density and climatology factors on the pandemic spread were investigated. Population density was reported as a major factor that affects the infection rate. In another study, ANFIS has been integrated with chaotic marine predators algorithm to obtain its optimal parameters that maximize the forecasting accuracy [59]. The flow chart of the hybrid ANFIS/chaotic marine predators algorithm is shown in Figure 6. The results of the proposed algorithm have been compared with that of standalone ANFIS as well as the optimized ANFIS using particle swarm optimization. The proposed algorithm had better forecasting accuracy; however, it had the highest computational cost among other investigated algorithms.
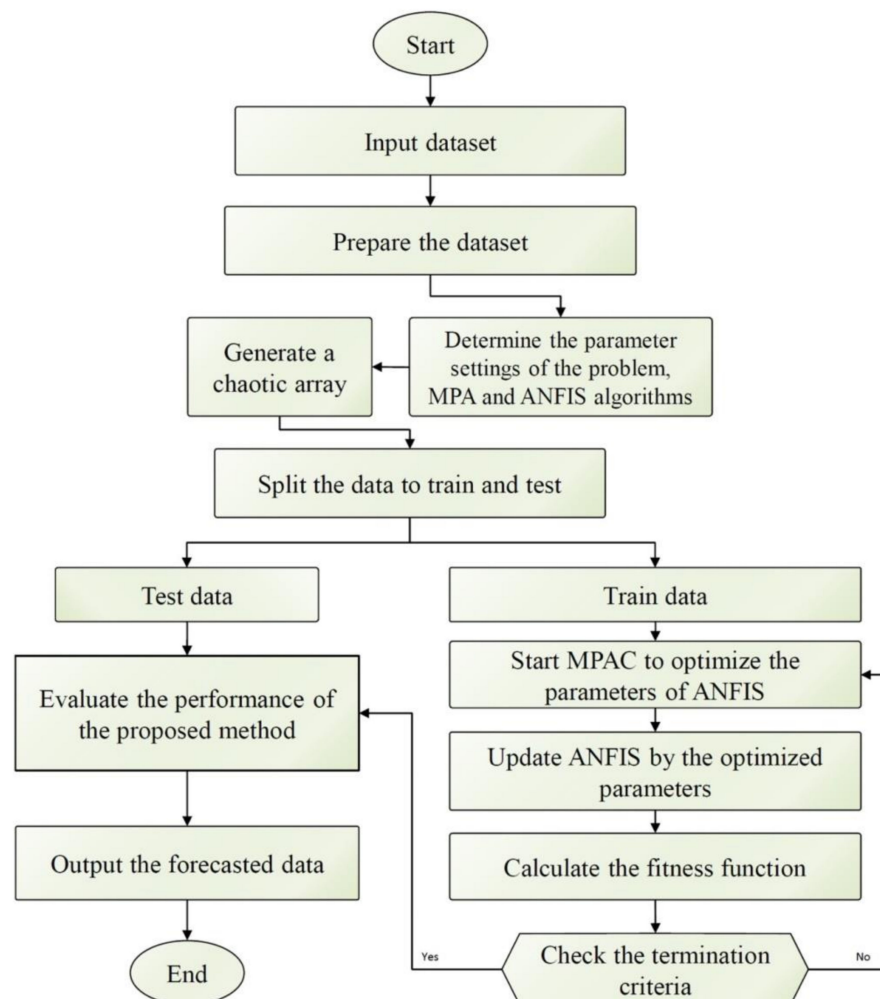
Figure 6. The flow chart of the hybrid ANFIS/chaotic marine predators algorithm used as an optimized forecasting tool [59].

HFFA has been employed for forecasting the cases and deaths of COVID-19 considering four main inputs, namely, confirmed cases with linear fractal dimension, confirmed cases with nonlinear fractal dimension, deaths with linear fractal dimension and deaths with nonlinear fractal dimension [60]. The fractal dimension was utilized to determine the complexity of the time series, while fuzzy logic was utilized to characterize the fore-

casting uncertainty. The proposed approach was used to predict the COVID-19 spread in 10 countries in Europe, Asia and America for a long forecasting period (one month).

SSA was employed to predict the confirmed case in the top 10 affected countries (United States, Brazil, India, France, Russia, Spain, United Kingdom, Mexico, Colombia and Argentina) considering both vector and recurrent forecasting algorithms [9]. The obtained results by SSA were compared by those obtained by ARIMA, ETS, NNAR, ARFIMA and TBATS. The evidence from that study indicated that there is not a single algorithm that may be considered as best for all investigated countries, which have different pandemic trends.

Recurrent SSA has been also used in forecasting COVID–19 cases in Malaysia [61]. The obtained results indicated that the proposed model is able to predict the spread of the outbreak with reasonable accuracy as it has over-forecasted by 0.36% with high correlation between predicted and confirmed cases. However, it cannot capture the sudden variation in the COVID-19 spread caused by the applied motion control strategies.

BNN was employed to predict the accumulated confirmed cases in five American states and five Brazilian states [62]. Moreover, it was reported that including the metrological conditions, such as precipitation and temperature, enhances the accuracy in the forecasting models. BNN had a higher forecasting accuracy compared with other investigated standalone models, such as support vector regression, cubist regression, quantile random forest, and k-nearest neighbors.

A hybrid approach composed of an RVFL model and a discrete wavelet transform was developed to predict the spread of COVID-19 in Russia, Brazil, Peru, India, and the United States [63]. The forecasting accuracy of the proposed approach is compared with the SVR model and the standalone RVFL model. The approach succeeded to obtain accurate forecasting for long period (60 days) and showed better performance compared with other models.
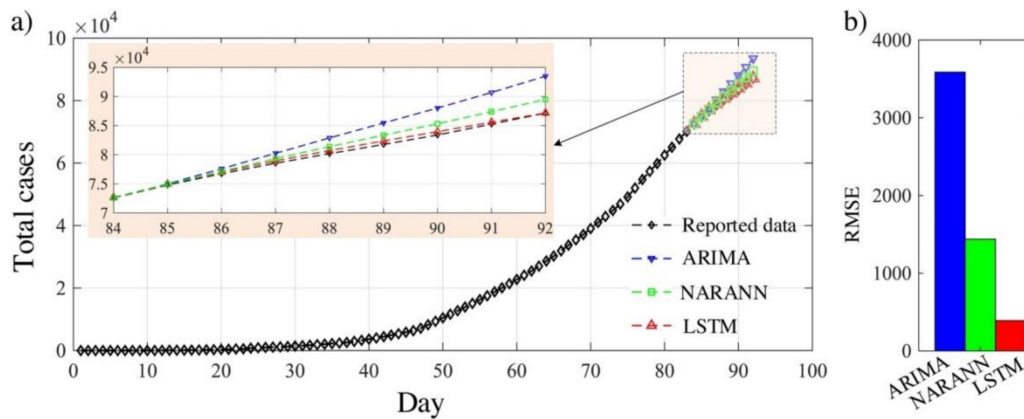
LSTM has been applied to predict COVID-19 cases in Canada [64], India [65], India, the United States [66], and Peru, Russia and Iran [67]. It succeeded to predict the trend of the pandemic despite of the nature of the transmission rate trend (linear, cubic, or exponential growth).

Three LSTM architectures, namely, stacked LSTM, convolutional LSTM and bi-directional LSTM have been investigated and compared with each other to predict the cumulative infected cases in India and the United States for one month ahead [66]. The convolutional LSTM model had the best accuracy followed by bi-directional, while the stacked LSTM model had the worst accuracy.

A comparative study between ARIMA, LSTM and NARNN has been carried out to predict COVID-19 cases in eight European countries, namely, Germany, the United Kingdom, Denmark, France, Belgium, Finland, Turkey and Switzerland [12]. Despite of the differences in human behavior, applied measures and available data for each country, LSTM had much higher forecasting accuracy compared to ARIMA and NARNN for all countries investigated within that study. In another study conducted on Saudi Arabia data, the three approaches were compared [11]. ARIMA and NARANN had a higher forecasting error compared with LSTM as their RMSE are higher than that of LSTM by 909% and 357%, respectively. Thus, LSTM has outperformance over ARIMA and NARNN in forecasting COVID-19 cases, as shown in Figure 7.

Transfer learning has been applied in LSTM networks to predict the spread trend of COVID-19 cases and deaths in France, Germany, India, Brazil and Nepal based on the data of early infected countries such as the United States and Italy [68]. The network was trained using data of early infected countries. LSTM succeeded to capture the complex patterns of the pandemic spread in the United States and Italy and in utilizing them in the forecasting process. LSTM has been also used to predict the number of Chinese tourists that travel to the United States and Australia during the COVID-19 pandemic based on the tourism data collected after the 2003 SARS outbreak [69]. It was reported that the delay of vaccine discovery would significantly decelerate the economic recovery. In addition, the application of preventive measures could affect the prediction accuracy of the proposed approach.

Figure 7. (**a**) Time series plot of the reported COVID-19 cases and forecasted data by NARANN, ARIMA and LSTM; (**b**) the RMSE for three approaches [11].

An optimized version of LSTM has been proposed by [70] to predict the spread of the COVID-19 pandemic. The Bayesian optimization technique was used to select the optimal LSTM parameters, such as activation function, learning rate and number of neurons. The hybrid LSTM/Bayesian optimization technique is shown in Figure 8. The proposed method has been used for short and long horizon forecasting. The optimized LSTM outperforms the standalone LSTM as well as the optimized version of CNN.
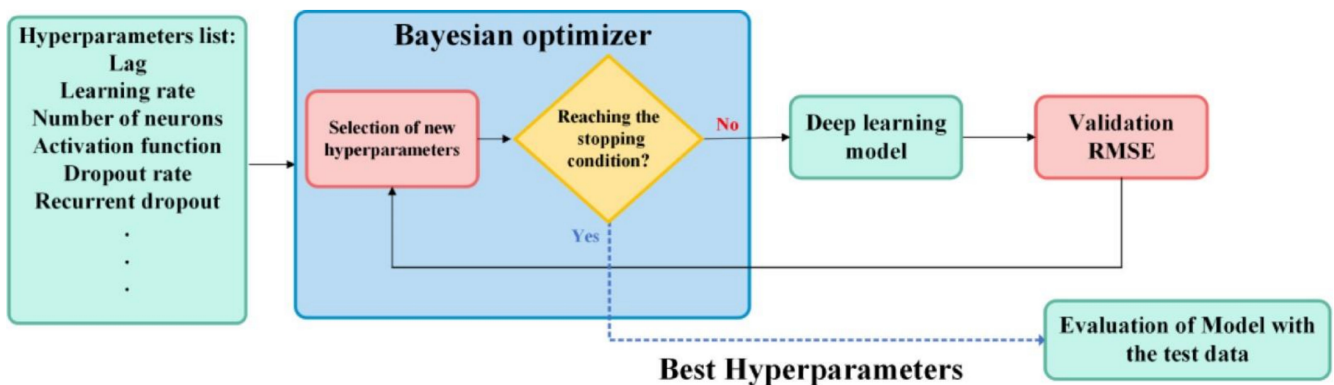
Figure 8. The hybrid LSTM/Bayesian optimization technique for forecasting COVID-19 data [70].

The number of daily positive cases, recovered cases and deceased cases in India was forecasted using LSTM and curve fitting approach [71]. The effect of the applied preventing measures, such as hotspots lockdown and social distancing on the outbreak spreading, have been investigated. The spread of the outbreak can be significantly reduced by applying strict preventive measures.

A comparison study between ANN and two mathematical approaches, namely, logistic and Gompertz, to predict the total cases in Mexico was carried out by [72]. A good fit between the reported data and those predicted by ANN, logistic, and Gompertz models with correlation coefficient of 0.9999, 0.9996, and 0.9998, respectively. It was declared that the forecasted results may significantly differ from the reported ones, if social distancing policies and testing strategies would be changing in the coming days. Recurrent ANN model was developed to predict the daily mutation rate of COVID-19 [73]. The developed model was applied to the United States, Australia, China and the world. A comparison investigation was carried out to evaluate the forecasting accuracy of three commonly used models: ARIMA, nonlinear autoregression neural network and LSTM ANN [12]. The results of the three investigated models were compared with actual reported data and the forecasting accuracy was evaluated using different statistical criteria. LSTM ANN showed the best accuracy among all investigated models to predict the total cases in eight

countries, namely, Switzerland, Finland, Germany, Denmark, France, Turkey, Belgium, and the United Kingdom. In another study, LSTM ANN has been employed to predict the infected cases in Canada [64]. The spread of the pandemic in Canada is lower than that of the United States and Italy, which reveals the effective control policies employed by the Canadian government. The transmission of the disease in Canada has a linear trend in contrary to that of the Unites States and Italy, which has an exponential trend. Based on the obtained results, the COVID-19 pandemic was expected to end by December 2020. The number of confirmed total cases in Egypt was forecasted using nonlinear autoregressive artificial neural networks, which showed a better forecasting accuracy compared with conventional autoregressive integrated moving average approach [10].

A novel hybrid approach consists of LSTM optimized by grey wolf optimizer for forecasting the future cumulated infected cases in the United States, India and the United Kingdom was proposed by [74]. The proposed approach utilized Google Trends and official reported data of the pandemic spread in training the network. The hybrid approach outperformed other conventional statistical approaches, such as ARIMA.

Five artificial intelligence approaches, namely, RNN, LSTM, bi-directional LSTM, GRUs and VAE have been developed to forecast recovered and confirmed cases in six different countries (United States, China, Italy, France, Spain, and Australia) based on short-period collected data [75]. VAE had better forecasting accuracy compared with other proposed approaches for all countries.

Metaheuristic optimization approaches have been also used to predict the prevalence of this pandemic [76]. A genetic-based programming approach has been developed to predict the number of total cases and total deaths in three Indian states, namely, Gujarat, Maharashtra and Delhi [77]. An enhanced ANFIS incorporated with two metaheuristic optimization algorithms, flower pollination and salp swarm algorithms, has been developed to predict the prevalence of COVID-19 in China [78].

Apart from the utilization of AI models in predicting studies, they have been also reported as powerful tools to diagnose the infected cases by processing X-ray imaging of the chest to avoid human errors [79,80]. The most common reported AI approaches to detect COVID-19 infections based on X-ray or CT images are CNN models [81–84]. Aswathy et al. [85] used two CNN models, namely ResNet-50 and DenseNet-201, to identify and assess the COVID-19 infection from CT images as well as the severity condition of the patient. They succeeded in developing a single architecture of the model that can be used to achieve both targets. Li et al. [86] utilized three other CNN models, namely VGG-16, LeNet-5 and ResNet-18, to detect COVID-19 infection using X-ray images. EfficientNet CNN was used to classify the diagnosed cases into normal, pneumonia andCOVID-19 cases based on their X-ray images [87]. CoroNet CNN was used to classify the diagnosed cases into normal, pneumonia-bacterial, pneumonia-viral and COVID–19 cases based on their X-ray images [88]. ResNet50 and VGG16 CCN were utilized to detect COVID-19 infections based on CT images [89]. The proposed models exhibited excellent accuracy when they were used as binary classifiers (normal and COVID-19), but their accuracy degraded when they were used as multiclass classifiers (normal, pneumonia, COVID-19). The same models (ResNet-50, VGG-16) along with auto-encoder CNN and machine learning techniques, such as nearest neighbor, logical regression, support vector machine, random forest and stochastic gradient descent, were proposed to classify CT images of COVID-19 [90]. A hybrid CNN, Sobel filter and support vector machine model was developed to detect COVID-19 based on X-ray images [91]. First, X-ray images are filtered using Sobel filter to detect the edges and contours of the images. Then, the filtered images are fed into to CNN model and classified using a support vector machine. Saha et al. [92] proposed a hybrid AIT to detect COVID-19 infection from X-ray images. First, a CNN model was utilized to extract the main features of X-ray images. Then, machine learning approaches such as support vector machine, random forest, and decision tree were utilized to detect COVID-19 infection based on the extracted features. Another hybrid CNN and deep neural network was developed to examine CT images to detect COVID-19 infections [93]. Shorfuzzaman

and Hossain [94] developed a Siamese neural network model to detect COVID-19 infections using X-ray images. A fine-tuned VGG16 CNN was used as an encoder to detect unbiased feature in examined images. The diagnosis of the COVID-19 from the images was formulated as N-way/K-shot classification problem where N and K denote the number of data samples and class labels used to train the model. The model succeeded in obtaining reasonable accuracy using a limited number of samples and examples during the training process. The integration between AI models and the metaheuristic optimizers [95,96] has been also reported as an efficient tool to diagnose COVID-19 infections.

More investigations should be carried out to develop hybrid AI models to predict and forecast the prevalence of this pandemic as well as detect the infected cases via image processing approaches. These hybrid models are composed of an AI model, such as ANN, ANFIS and RVFL integrated with metaheuristic optimizers [96] such as pigeon optimizer [97], gradient-based optimizer [98], parasitism-predation optimizer [99], ecosystem-based optimization [100], Hunger games algorithm [101], flower pollination algorithm [102] and political optimizer [103].

5. Conclusions

This study presents a review on the applications of AI techniques for forecasting the prevalence of the COVID-19 pandemic. The basics and the mathematical formulation of different AI approaches used in this context are presented, including nonlinear autoregressive neural network, adaptive neuro-fuzzy inference system, hybrid fractal-fuzzy approach, long short-term memory network, Bayesian neural network, variational auto-encoder and singular spectrum analysis. Different statistical measures used to evaluate the forecasting accuracy of the AI approaches are discussed. Despite of the differences in human behavior, applied measures and available data for each country, AI approaches had much higher forecasting accuracy compared to conventional statistical approaches. The evidence from the current review indicates that there is not a single approach that may be considered as the best one for all investigated countries, which have different pandemic trends. Moreover, the integration between AI forecasting approaches and advanced optimization methods could significantly enhance the accuracy of the whole model with acceptable increase in the computational cost. Thus, it is recommended to use hybrid models instead of standalone models. A substantial concern should be paid towards the utilization of AI approaches to capture the sudden variation in the Covid-19 spread caused by the applied motion control strategies. Future work may focus on the integration between AI approaches and metaheuristic optimization methods to enhance their forecasting accuracy. Forecasting the trend of the economic recovery after this pandemic deserves the researchers' attention.

Author Contributions: A.H.E. and A.I.S. made a substantial contribution to the concept and design of the article. H.P., S.S., N.A.A., M.A. and A.I.S. collected the data. A.H.E., N.A.A., M.A. and A.I.S. analyzed the data. S.S., H.P. and A.H.E. revised the article. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. He, Y.; Li, M.; Zhong, Q.; Li, Q.; Yang, R.; Lin, J.; Zhang, X. The Chinese Government's response to the pandemic: Measures, dynamic changes, and main patterns. *Healthcare* **2021**, *9*, 1020. [CrossRef]
2. Xu, W.; Wu, J.; Cao, L. COVID-19 pandemic in China: Context, experience and lessons. *Health Policy Technol.* **2020**, *9*, 639–648. [CrossRef]

3. Briz-Redón, Á. The impact of modelling choices on modelling outcomes: A spatio-temporal study of the association between COVID-19 spread and environmental conditions in Catalonia (Spain). *Stoch. Environ. Res. Risk Assess.* **2021**, *35*, 1701–1713. [CrossRef] [PubMed]

4. Rakha, A.; Hettiarachchi, H.; Rady, D.; Gaber, M.M.; Rakha, E.; Abdelsamea, M.M. Predicting the economic impact of the COVID-19 pandemic in the United Kingdom using time-series mining. *Economies* **2021**, *9*, 137. [CrossRef]

5. Alzahrani, S.I.; Aljamaan, I.A.; Al-Fakih, E.A. 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*, 914–919. [CrossRef] [PubMed]

6. Li, X.-P.; Wang, Y.; Khan, M.A.; Alshahrani, M.Y.; Muhammad, T. A dynamical study of SARS-COV-2: A study of third wave. *Results Phys.* **2021**, *29*, 104705. [CrossRef] [PubMed]

7. Helwan, A.; Ma'Aitah, M.K.S.; Hamdan, H.; Ozsahin, D.U.; Tuncyurek, O. Radiologists versus deep convolutional neural networks: A comparative study for diagnosing COVID-19. *Comput. Math. Methods Med.* **2021**, *2021*, 5527271. [CrossRef] [PubMed]

8. Uddin, M.I.; Shah, S.A.A.; Al-Khasawneh, M.A. A novel deep convolutional neural network model to monitor people following guidelines to avoid COVID-19. *J. Sens.* **2020**, *2020*, 8856801. [CrossRef]

9. Kalantari, M. Forecasting COVID-19 pandemic using optimal singular spectrum analysis. *Chaos Solitons Fractals* **2021**, *142*, 110547. [CrossRef]

10. Saba, A.I.; Elsheikh, A.H. Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks. *Process. Saf. Environ. Prot.* **2020**, *141*, 1–8. [CrossRef]

11. Elsheikh, A.H.; Saba, A.I.; Elaziz, M.A.; Lu, S.; Shanmugan, S.; Muthuramalingam, T.; Kumar, R.; Mosleh, A.O.; Essa, F.A.; Shehabeldeen, T.A. Deep learning-based forecasting model for COVID-19 outbreak in Saudi Arabia. *Process. Saf. Environ. Prot.* **2021**, *149*, 223–233. [CrossRef]

12. Kırbaş, İ.; Sözen, A.; Tuncer, A.D.; Kazancıoğlu, F.Ş. Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. *Chaos Solitons Fractals* **2020**, *138*, 110015. [CrossRef] [PubMed]

13. Anam, S.; Maulana, M.H.A.A.; Hidayat, N.; Yanti, I.; Fitriah, Z.; Mahanani, D.M. Predicting the number of COVID-19 sufferers in Malang city using the backpropagation neural network with the Fletcher-Reeves method. *Appl. Comput. Intell. Soft Comput.* **2021**, *2021*, 6658552. [CrossRef]

14. Potdar, K.; Kinnerkar, R. A non-linear autoregressive neural network model for forecasting Indian index of industrial production. In Proceedings of the 2017 IEEE Region 10 Symposium (TENSYMP), Cochin, India, 14–16 July 2017; pp. 1–5.

15. Alsumaiei, A.A. A nonlinear autoregressive modeling approach for forecasting groundwater level fluctuation in urban aquifers. *Water* **2020**, *12*, 820. [CrossRef]

16. Boussaada, Z.; Curea, O.; Remaci, A.; Camblong, H.; Mrabet Bellaaj, N. A nonlinear autoregressive exogenous (NARX) neural network model for the prediction of the daily direct solar radiation. *Energies* **2018**, *11*, 620. [CrossRef]

17. Gautam, S.S. A novel moving average forecasting approach using fuzzy time series data set. *J. Control Autom. Electr. Syst.* **2019**, *30*, 532–544. [CrossRef]

18. Bacanli, U.G.; Firat, M.; Dikbas, F. Adaptive neuro-fuzzy inference system for drought forecasting. *Stoch. Environ. Res. Risk Assess.* **2009**, *23*, 1143–1154. [CrossRef]

19. Duong, T.H.; Nguyen, D.C.; Nguyen, S.D.; Hoang, M.H. An adaptive neuro-fuzzy inference system for seasonal forecasting of tropical cyclones making landfall along the Vietnam Coast. In *Advanced Computational Methods for Knowledge Engineering*; Springer: Heidelberg, Germany, 2013; pp. 225–236.

20. Al-qaness, M.A.; Abd Elaziz, M.; Ewees, A.A.; Cui, X. A modified adaptive neuro-fuzzy inference system using multi-verse optimizer algorithm for oil consumption forecasting. *Electronics* **2019**, *8*, 1071. [CrossRef]

21. Yang, L.; Zhang, D.; Zhang, X.; Tian, A. Surface profile topography of ionic polymer metal composite based on fractal theory. *Surf. Interfaces* **2021**, *22*, 100834. [CrossRef]

22. Karaca, Y.; Zhang, Y.-D.; Muhammad, K. Characterizing complexity and self-similarity based on fractal and entropy analyses for stock market forecast modelling. *Expert Syst. Appl.* **2020**, *144*, 113098. [CrossRef]

23. Sherstinsky, A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [CrossRef]

24. Kumari, K.; Dey, P.; Kumar, C.; Pandit, D.; Mishra, S.S.; Kisku, V.; Chaulya, S.K.; Ray, S.K.; Prasad, G.M. UMAP and LSTM based fire status and explosibility prediction for sealed-off area in underground coal mine. *Process. Saf. Environ. Prot.* **2021**, *146*, 837–852. [CrossRef]

25. Memarzadeh, G.; Keynia, F. Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm. *Electr. Power Syst. Res.* **2021**, *192*, 106995. [CrossRef]

26. Zhang, N.; Shen, S.-L.; Zhou, A.; Jin, Y.-F. Application of LSTM approach for modelling stress–strain behaviour of soil. *Appl. Soft Comput.* **2021**, *100*, 106959. [CrossRef]

27. Elsheikh, A.H.; Katekar, V.P.; Muskens, O.L.; Deshmukh, S.S.; Elaziz, M.A.; Dabour, S.M. Utilization of LSTM neural network for water production forecasting of a stepped solar still with a corrugated absorber plate. *Process. Saf. Environ. Prot.* **2021**, *148*, 273–282. [CrossRef]

28. Chang, Y.-S.; Chiao, H.-T.; Abimannan, S.; Huang, Y.-P.; Tsai, Y.-T.; Lin, K.-M. An LSTM-based aggregated model for air pollution forecasting. *Atmos. Pollut. Res.* **2020**, *11*, 1451–1463. [CrossRef]

29. Zheng, S.; Zhong, Q.; Peng, L.; Chai, X. A simple method of residential electricity load forecasting by improved Bayesian neural networks. *Math. Probl. Eng.* **2018**, *2018*, 225–236. [CrossRef]

30. Mbuvha, R.; Jonsson, M.; Ehn, N.; Herman, P. Bayesian neural networks for one-hour ahead wind power forecasting. In Proceedings of the 2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA), San Diego, CA, USA, 5–8 November 2017; pp. 591–596.

31. Zhang, X.; Liang, F.; Yu, B.; Zong, Z. Explicitly integrating parameter, input, and structure uncertainties into Bayesian Neural Networks for probabilistic hydrologic forecasting. *J. Hydrol.* **2011**, *409*, 696–709. [CrossRef]

32. Leles, M.C.R.; Sansão, J.P.H.; Mozelli, L.A.; Guimarães, H.N. A new algorithm in singular spectrum analysis framework: The Overlap-SSA (ov-SSA). *SoftwareX* **2018**, *8*, 26–32. [CrossRef]

33. Marques, C.A.F.; Ferreira, J.A.; Rocha, A.; Castanheira, J.M.; Melo-Gonçalves, P.; Vaz, N.; Dias, J.M. Singular spectrum analysis and forecasting of hydrological time series. *Phys. Chem. Earth Parts A/B/C* **2006**, *31*, 1172–1179. [CrossRef]

34. Rodrigues, P.C.; Pimentel, J.; Messala, P.; Kazemi, M. The decomposition and forecasting of mutual investment funds using singular spectrum analysis. *Entropy* **2020**, *22*, 83. [CrossRef] [PubMed]

35. Moreno, S.R.; dos Santos Coelho, L. Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. *Renew. Energy* **2018**, *126*, 736–754. [CrossRef]

36. Lima de Menezes, M.; Castro Souza, R.; Moreira Pessanha, J.F. Electricity consumption forecasting using singular spectrum analysis. *Dyna* **2015**, *82*, 138–146. [CrossRef]

37. Elsheikh, A.H.; Sharshir, S.W.; Abd Elaziz, M.; Kabeel, A.; Guilan, W.; Haiou, Z. Modeling of solar energy systems using artificial neural network: A comprehensive review. *Sol. Energy* **2019**, *180*, 622–639. [CrossRef]

38. Yousaf, M.; Zahir, S.; Riaz, M.; Hussain, S.M.; Shah, K. Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan. *Chaos Solitons Fractals* **2020**, *138*, 109926. [CrossRef] [PubMed]

39. Duan, X.; Zhang, X. ARIMA modelling and forecasting of irregularly patterned COVID-19 outbreaks using Japanese and South Korean data. *Data Brief* **2020**, *31*, 105779. [CrossRef] [PubMed]

40. Chakraborty, T.; Ghosh, I. Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis. *Chaos Solitons Fractals* **2020**, *135*, 109850. [CrossRef] [PubMed]

41. Singh, S.; Parmar, K.S.; Kumar, J.; Makkhan, S.J.S. Development of new hybrid model of discrete wavelet decomposition and Autoregressive Integrated Moving Average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. *Chaos Solitons Fractals* **2020**, *135*, 109866. [CrossRef] [PubMed]

42. Ahmar, A.S.; del Val, E.B. SutteARIMA: Short-term forecasting method, a case: Covid-19 and stock market in Spain. *Sci. Total. Environ.* **2020**, *729*, 138883. [CrossRef]

43. Aslam, M. Using the Kalman filter with ARIMA for the COVID-19 pandemic dataset of Pakistan. *Data Brief* **2020**, *31*, 105854. [CrossRef] [PubMed]

44. Ribeiro, M.H.D.M.; da Silva, R.G.; Mariani, V.C.; Coelho, L.D.S. Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos Solitons Fractals* **2020**, *135*, 109853. [CrossRef] [PubMed]

45. Barmparis, G.D.; Tsironis, G.P. Estimating the infection horizon of COVID-19 in eight countries with a data-driven approach. *Chaos Solitons Fractals* **2020**, *135*, 109842. [CrossRef] [PubMed]

46. Arias Velásquez, R.M.; Mejía Lara, J.V. Forecast and evaluation of COVID-19 spreading in USA with reduced-space Gaussian process regression. *Chaos Solitons Fractals* **2020**, *136*, 109924. [CrossRef] [PubMed]

47. Postnikov, E.B. Estimation of COVID-19 dynamics "on a back-of-envelope": Does the simplest SIR model provide quantitative parameters and predictions? *Chaos Solitons Fractals* **2020**, *135*, 109841. [CrossRef] [PubMed]

48. Ren, H.; Zhao, L.; Zhang, A.; Song, L.; Liao, Y.; Lu, W.; Cui, C. Early forecasting of the potential risk zones of COVID-19 in China's megacities. *Sci. Total Environ.* **2020**, *729*, 138995. [CrossRef]

49. Şahin, U.; Şahin, T. Forecasting the cumulative number of confirmed cases of COVID-19 in Italy, UK and USA using fractional nonlinear grey Bernoulli model. *Chaos Solitons Fractals* **2020**, *138*, 109948. [CrossRef] [PubMed]

50. Alberti, T.; Faranda, D. On the uncertainty of real-time predictions of epidemic growths: A COVID-19 case study for China and Italy. *Commun. Nonlinear Sci. Numer. Simul.* **2020**, *90*, 105372. [CrossRef]

51. Wang, L.; Li, J.; Guo, S.; Xie, N.; Yao, L.; Cao, Y.; Day, S.W.; Howard, S.C.; Graff, J.C.; Gu, T.; et al. Real-time estimation and prediction of mortality caused by COVID-19 with patient information based algorithm. *Sci. Total Environ.* **2020**, *727*, 138394. [CrossRef] [PubMed]

52. Perc, M.; Gorišek Miksić, N.; Slavinec, M.; Stožer, A. Forecasting COVID-19. *Front. Phys.* **2020**, *8*, 127. [CrossRef]

53. Ayinde, K.; Lukman, A.F.; Rauf, R.I.; Alabi, O.O.; Okon, C.E.; Ayinde, O.E. Modeling Nigerian covid-19 cases: A comparative analysis of models and estimators. *Chaos Solitons Fractals* **2020**, *138*, 109911. [CrossRef] [PubMed]

54. Singhal, A.; Singh, P.; Lall, B.; Joshi, S.D. Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. *Chaos Solitons Fractals* **2020**, *138*, 110023. [CrossRef] [PubMed]

55. Higazy, M. Novel fractional order SIDARTHE mathematical model of the COVID-19 pandemic. *Chaos Solitons Fractals* **2020**, *138*, 110007. [CrossRef] [PubMed]

56. Ly, K.T. A COVID-19 forecasting system using adaptive neuro-fuzzy inference. *Financ. Res. Lett.* **2020**, *41*, 101844. [CrossRef] [PubMed]

57. Alsayed, A.; Sadir, H.; Kamil, R.; Sari, H. Prediction of epidemic peak and infected cases for COVID-19 disease in Malaysia, 2020. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4076. [CrossRef] [PubMed]

58. Behnood, A.; Mohammadi Golafshani, E.; Hosseini, S.M. Determinants of the infection rate of the COVID-19 in the U.S. using ANFIS and Virus Optimization Algorithm (VOA). *Chaos Solitons Fractals* **2020**, *139*, 110051. [CrossRef]

59. Al-qaness, M.A.A.; Saba, A.I.; Elsheikh, A.H.; Elaziz, M.A.; Ibrahim, R.A.; Lu, S.; Hemedan, A.A.; Shanmugan, S.; Ewees, A.A. Efficient artificial intelligence forecasting models for COVID-19 outbreak in Russia and Brazil. *Process. Saf. Environ. Prot.* **2021**, *149*, 399–409. [CrossRef] [PubMed]

60. Castillo, O.; Melin, P. Forecasting of COVID-19 time series for countries in the world based on a hybrid approach combining the fractal dimension and fuzzy logic. *Chaos Solitons Fractals* **2020**, *140*, 110242. [CrossRef] [PubMed]

61. Shaharudin, S.M.; Ismail, S.; Tan, M.L.; Mohamed, N.S.; AininaFilzaSulaiman, N. Predictive modelling of covid-19 cases in Malaysia based on recurrent forecasting-singular spectrum analysis approach. *Int. J. Adv. Trends Comput. Sci. Eng.* **2020**, *9*.

62. Da Silva, R.G.; Ribeiro, M.H.D.M.; Mariani, V.C.; Coelho, L.D.S. Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos Solitons Fractals* **2020**, *139*, 110027. [CrossRef] [PubMed]

63. Hazarika, B.B.; Gupta, D. Modelling and forecasting of COVID-19 spread using wavelet-coupled random vector functional link networks. *Appl. Soft Comput.* **2020**, *96*, 106626. [CrossRef]

64. Chimmula, V.K.R.; Zhang, L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos Solitons Fractals* **2020**, *135*, 109864. [CrossRef] [PubMed]

65. Shastri, S.; Singh, K.; Kumar, S.; Kour, P.; Mansotra, V. Deep-LSTM ensemble framework to forecast Covid-19: An insight to the global pandemic. *Int. J. Inf. Technol.* **2021**, *13*, 1291–1301. [CrossRef] [PubMed]

66. Shastri, S.; Singh, K.; Kumar, S.; Kour, P.; Mansotra, V. Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. *Chaos Solitons Fractals* **2020**, *140*, 110227. [CrossRef] [PubMed]

67. Wang, P.; Zheng, X.; Ai, G.; Liu, D.; Zhu, B. Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos Solitons Fractals* **2020**, *140*, 110214. [CrossRef] [PubMed]

68. Gautam, Y. Transfer Learning for COVID-19 cases and deaths forecast using LSTM network. *ISA Trans.* **2021**. [CrossRef]

69. Polyzos, S.; Samitas, A.; Spyridou, A.E. Tourism demand and the COVID-19 pandemic: An LSTM approach. *Tour. Recreat. Res.* **2020**, *46*, 175–187. [CrossRef]

70. Abbasimehr, H.; Paki, R. Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. *Chaos Solitons Fractals* **2020**, *142*, 110511. [CrossRef] [PubMed]

71. Tomar, A.; Gupta, N. Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Sci. Total Environ.* **2020**, *728*, 138762. [CrossRef]

72. Torrealba-Rodriguez, O.; Conde-Gutiérrez, R.A.; Hernández-Javier, A.L. Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos Solitons Fractals* **2020**, *138*, 109946. [CrossRef]

73. Pathan, R.K.; Biswas, M.; Khandaker, M.U. Time series prediction of COVID-19 by mutation rate analysis using recurrent neural network-based LSTM model. *Chaos Solitons Fractals* **2020**, *138*, 110018. [CrossRef]

74. Prasanth, S.; Singh, U.; Kumar, A.; Tikkiwal, V.A.; Chong, P.H.J. Forecasting spread of COVID-19 using google trends: A hybrid GWO-deep learning approach. *Chaos Solitons Fractals* **2020**, *142*, 110336. [CrossRef]

75. 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. [CrossRef] [PubMed]

76. Issa, M.; Helmi, A.M.; Elsheikh, A.H.; Abd Elaziz, M. A biological sub-sequences detection using integrated BA-PSO based on infection propagation mechanism: Case study COVID-19. *Expert Syst. Appl.* **2022**, *189*, 116063. [CrossRef]

77. Salgotra, R.; Gandomi, M.; Gandomi, A.H. Time Series Analysis and Forecast of the COVID-19 Pandemic in India using Genetic Programming. *Chaos Solitons Fractals* **2020**, *138*, 109945. [CrossRef] [PubMed]

78. Al-Qaness, M.A.; Ewees, A.A.; Fan, H.; Abd El Aziz, M. Optimization method for forecasting confirmed cases of COVID-19 in China. *J. Clin. Med.* **2020**, *9*, 674. [CrossRef]

79. Sitaula, C.; Hossain, M.B. Attention-based VGG-16 model for COVID-19 chest X-ray image classification. *Appl. Intell.* **2021**, *51*, 2850–2863. [CrossRef]

80. Abd Elaziz, M.; Dahou, A.; Alsaleh, N.A.; Elsheikh, A.H.; Saba, A.I.; Ahmadein, M. Boosting COVID-19 Image Classification Using MobileNetV3 and Aquila Optimizer Algorithm. *Entropy* **2021**, *23*, 1383. [CrossRef]

81. Wang, L.; Lin, Z.Q.; Wong, A. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci. Rep.* **2020**, *10*, 19549. [CrossRef]

82. Zhao, W.; Jiang, W.; Qiu, X. Deep learning for COVID-19 detection based on CT images. *Sci. Rep.* **2021**, *11*, 14353. [CrossRef] [PubMed]

83. Das, A.K.; Ghosh, S.; Thunder, S.; Dutta, R.; Agarwal, S.; Chakrabarti, A. Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. *Pattern Anal. Appl.* **2021**, *24*, 1111–1124. [CrossRef]

84. Rahimzadeh, M.; Attar, A. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Inform. Med. Unlocked* **2020**, *19*, 100360. [CrossRef]

85. Aswathy, A.L.; Hareendran, A.; Vinod Chandra, S.S. COVID-19 diagnosis and severity detection from CT-images using transfer learning and back propagation neural network. *J. Infect. Public Health* **2021**, *14*, 1435–1445. [CrossRef]

86. Li, X.; Zhai, M.; Sun, J. DDCNNC: Dilated and depthwise separable convolutional neural Network for diagnosis COVID-19 via chest X-ray images. *Int. J. Cogn. Comput. Eng.* **2021**, *2*, 71–82. [CrossRef]

87. Marques, G.; Agarwal, D.; de la Torre Díez, I. Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. *Appl. Soft Comput.* **2020**, *96*, 106691. [CrossRef]

88. Khan, A.I.; Shah, J.L.; Bhat, M.M. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Comput. Methods Programs Biomed.* **2020**, *196*, 105581. [CrossRef] [PubMed]

89. Mishra, N.K.; Singh, P.; Joshi, S.D. Automated detection of COVID-19 from CT scan using convolutional neural network. *Biomed. Signal. Process. Control* **2021**, *41*, 572–588. [CrossRef] [PubMed]

90. Fouladi, S.; Ebadi, M.J.; Safaei, A.A.; Bajuri, M.Y.; Ahmadian, A. Efficient deep neural networks for classification of COVID-19 based on CT images: Virtualization via software defined radio. *Comput. Commun.* **2021**, *176*, 234–248. [CrossRef] [PubMed]

91. Sharifrazi, D.; Alizadehsani, R.; Roshanzamir, M.; Joloudari, J.H.; Shoeibi, A.; Jafari, M.; Hussain, S.; Sani, Z.A.; Hasanzadeh, F.; Khozeimeh, F.; et al. Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of COVID-19 patients using X-ray images. *Biomed. Signal. Process. Control* **2021**, *68*, 102622. [CrossRef] [PubMed]

92. Saha, P.; Sadi, M.S.; Islam, M.M. EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers. *Inform. Med. Unlocked* **2021**, *22*, 100505. [CrossRef] [PubMed]

93. Hassantabar, S.; Ahmadi, M.; Sharifi, A. Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches. *Chaos Solitons Fractals* **2020**, *140*, 110170. [CrossRef]

94. Shorfuzzaman, M.; Hossain, M.S. MetaCOVID: A Siamese neural network framework with contrastive loss for n-shot diagnosis of COVID-19 patients. *Pattern Recognit.* **2021**, *113*, 107700. [CrossRef] [PubMed]

95. Oliva, D.; Elaziz, M.A.; Elsheikh, A.H.; Ewees, A.A. A review on meta-heuristics methods for estimating parameters of solar cells. *J. Power Sources* **2019**, *435*, 126683. [CrossRef]

96. Abd Elaziz, M.; Elsheikh, A.H.; Oliva, D.; Abualigah, L.; Lu, S.; Ewees, A.A. Advanced Metaheuristic Techniques for Mechanical Design Problems: Review. *Arch. Comput. Methods Eng.* **2021**. [CrossRef]

97. Elsheikh, A.H.; Muthuramalingam, T.; Shanmugan, S.; Mahmoud Ibrahim, A.M.; Ramesh, B.; Khoshaim, A.B.; Moustafa, E.B.; Bedairi, B.; Panchal, H.; Sathyamurthy, R. Fine-tuned artificial intelligence model using pigeon optimizer for prediction of residual stresses during turning of Inconel 718. *J. Mater. Res. Technol.* **2021**, *15*, 3622–3634. [CrossRef]

98. Elsheikh, A.H.; Abd Elaziz, M.; Vendan, A. Modeling ultrasonic welding of polymers using an optimized artificial intelligence model using a gradient-based optimizer. *Weld. World* **2021**. [CrossRef]

99. Elsheikh, A.H.; Abd Elaziz, M.; Ramesh, B.; Egiza, M.; Al-qaness, M.A.A. Modeling of drilling process of GFRP composite using a hybrid random vector functional link network/parasitism-predation algorithm. *J. Mater. Res. Technol.* **2021**, *14*, 298–311. [CrossRef]

100. Essa, F.A.; Abd Elaziz, M.; Elsheikh, A.H. Prediction of power consumption and water productivity of seawater greenhouse system using random vector functional link network integrated with artificial ecosystem-based optimization. *Process. Saf. Environ. Prot.* **2020**, *144*, 322–329. [CrossRef]

101. AbuShanab, W.S.; Abd Elaziz, M.; Ghandourah, E.I.; Moustafa, E.B.; Elsheikh, A.H. A new fine-tuned random vector functional link model using Hunger games search optimizer for modeling friction stir welding process of polymeric materials. *J. Mater. Res. Technol.* **2021**, *14*, 1482–1493. [CrossRef]

102. Khoshaim, A.B.; Elsheikh, A.H.; Moustafa, E.B.; Basha, M.; Mosleh, A.O. Prediction of residual stresses in turning of pure iron using artificial intelligence-based methods. *J. Mater. Res. Technol.* **2021**, *11*, 2181–2194. [CrossRef]

103. Elsheikh, A.H.; Elaziz, M.A.; Das, S.R.; Muthuramlddalingam, T.; Lu, S. A new optimized predictive model based on political optimizer for eco-friendly MQL-turning of AISI 4340 alloy with nano-lubricants. *J. Manuf. Process.* **2021**, *67*, 562–578. [CrossRef]