Article

COVID-19 Pandemic Prediction for Hungary; A Hybrid Machine Learning Approach

Gergo Pinter [1], **Imre Felde** [1], **Amir Mosavi** [2,3,4,5,*] (ID), **Pedram Ghamisi** [6] (ID) **and Richard Gloaguen** [7] (ID)

[1] John von Neumann Faculty of Informatics, Obuda University, 1034 Budapest, Hungary; pinter.gergo@nik.uni-obuda.hu (G.P.); felde@uni-obuda.hu (I.F.)

[2] Faculty of Civil Engineering, Technische Universität Dresden, 01069 Dresden, Germany

[3] Kalman Kando Faculty of Electrical Engineering, Obuda University, 1034 Budapest, Hungary

[4] Thuringian Institute of Sustainability and Climate Protection, 07743 Jena, Germany

[5] Department of Mathematics, J. Selye University, 94501 Komarno, Slovakia

[6] Machine Learning Group, Helmholtz-Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resource Technology, Chemnitzer Straße 40, 09599 Freiberg, Germany; p.ghamisi@hzdr.de

[7] Helmholtz-Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resource Technology, Chemnitzer Straße 40, 09599 Freiberg, Germany; r.gloaguen@hzdr.de

* Correspondence: amir.mosavi@mailbox.tu-dresden.de

Abstract: Several epidemiological models are being used around the world to project the number of infected individuals and the mortality rates of the COVID-19 outbreak. Advancing accurate prediction models is of utmost importance to take proper actions. Due to the lack of essential data and uncertainty, the epidemiological models have been challenged regarding the delivery of higher accuracy for long-term prediction. As an alternative to the susceptible-infected-resistant (SIR)-based models, this study proposes a hybrid machine learning approach to predict the COVID-19, and we exemplify its potential using data from Hungary. The hybrid machine learning methods of adaptive network-based fuzzy inference system (ANFIS) and multi-layered perceptron-imperialist competitive algorithm (MLP-ICA) are proposed to predict time series of infected individuals and mortality rate. The models predict that by late May, the outbreak and the total morality will drop substantially. The validation is performed for 9 days with promising results, which confirms the model accuracy. It is expected that the model maintains its accuracy as long as no significant interruption occurs. This paper provides an initial benchmarking to demonstrate the potential of machine learning for future research.

1. Introduction

Severe acute respiratory syndrome coronavirus 2, also known as SARS-CoV-2, is reported as a virus strain causing the respiratory disease of COVID-19 [1]. The World Health Organization (WHO) and the global nations confirmed the coronavirus disease to be extremely contagious [2,3]. The COVID-19 pandemic has been widely recognized as a public health emergency of international concern [4]. To estimate the outbreak, identify the peak ahead of time, and also predict the mortality rate, epidemiological models have been widely used by officials and media. Outbreak prediction models have shown to be fundamental to provide insights into the damages caused by COVID-19. Furthermore, the prediction models are used as a reference to make new policies and to evaluate the conditions of curfew [5]. The COVID-19 pandemic has been reported to be extremely aggressive to spread [6]. Due to the uncertainty and complexity of the COVID-19 outbreak and its irregularity in

different countries, the standard epidemiological models, i.e., susceptible-infected-resistant (SIR)-based models, have been challenged for delivering higher performance in individual nations. Furthermore, as the COVID-19 outbreak showed significant differences with the recent epidemics, e.g., swine fever, H1N1 influenza, Ebola, Cholera, dengue fever, and Zika, several advanced epidemiological models have emerged to provide higher accuracy [7]. Nevertheless, due to the involvement of numerous unknown factors, the outbreaks, various curfew strategies in different countries, and existing double standards for evaluation metrics, the model uncertainty for the COVID-19 is generally reported extensive [8–10]. As a result, SIR-based models are fundamentally challenged to provide reliable insight into the progress of COVID-19.

The general strategy behind the SIR-based models for predicting the COVID-19 outbreak, similar to other epidemics, is formed around the assumption of transmitting the contagious disease through social contacts. The SIR-based models assume that infection spreads through several groups, e.g., susceptible, exposed, infected, recovered, deceased, and immune [11]. For instance, the standard SIR-based models assume that an epidemic includes susceptible to infection (class S), infected (class I), and the removed population (class R). Such principal groups build the foundation of epidemiological modeling. Note that the definition of various classes of the outbreak may vary. For instance, R is often referred to as those that have recovered, developed immunity, been isolated, or passed away. However, in various countries, R may or may not be susceptible to infection again, and there exist uncertainties in allocating R a value. Advancing SIR-based models require several assumptions. Thus, modeling with SIR-based models may include several contradicting assumptions.

Nevertheless, in most epidemiological models, it is generally agreed that class I has a high probability of infecting class S. The transmission probability is proportional to the total social contacts. The transmission can be estimated through the implementation of basic differential equations as follows [12–14].

$$\frac{dS}{dt} = -\alpha SI \tag{1}$$

where I, S, and α represent the infected population, the susceptible population, and the daily reproduction rate, respectively. The time-series of S, which is calculated by the differential equation, decreases gradually. However, it is observed that at the beginning of the pandemics, where the increment $\frac{dI}{dt}$ is linear, $S \approx 1$. Eventually, class I is estimated as follows.

$$\frac{dI}{dt} = \alpha SI - \beta I \tag{2}$$

where β represents the parameter of the daily rate of spread. Furthermore, the susceptible individuals excluded from the SIR-based models are calculated. Furthermore, the class R, representing individuals excluded from the spread of infection, is computed as follows:

$$\frac{dR}{dt} = \beta I \tag{3}$$

Considering the above-mentioned assumptions, and under the unconstrained conditions of the excluded group, Equation (3), the outbreak modeling with SIR is finally stated as:

$$I(t) \approx I_0 \, exp\left\{(\alpha - \beta)\right\}. \tag{4}$$

Furthermore, to evaluate the accuracy of the SIR-based models, the outbreak model's median success function is used, which is represented as:

$$f = \frac{Prediction}{True \; value}. \tag{5}$$

Several analytical solutions to the SIR models have been provided in the literature [15,16]. As the different nations take different actions toward slowing down the outbreak, the SIR-based model must be adapted according to the local assumptions [17]. The inaccuracy of many SIR-based models in predicting the outbreak and mortality rate has been evidenced during the COVID-19 in many nations. The critical success of a SIR-based model relies on choosing the right model according to the context and the relevant assumptions. SIS (susceptible-infectious-susceptible), SIRD (susceptible-infected-recovered-deceased-model), MSIR (Maternally-derived-immunity-susceptible-infected-recovered), SEIR (Susceptible-exposed-infected-recovered), SEIS (Susceptible-exposed-infected-susceptible), MSEIR (Maternally-derived-immunity-susceptible-exposed-infected-recovered), and MSEIRS (Maternally-derived-immunity-susceptible-exposed-infected-recovered-susceptible) models are among the popular models used for predicting COVID-19 outbreaks worldwide. The more advanced variation of SIR-d models carefully considers the vital dynamics and constant population [16]. For instance, at the presence of the long-lasting immunity assumption when the immunity is not realized upon recovery from infection, the susceptible-infectious-susceptible (SIS) model was suggested [18]. In contrast, the susceptible-infected-recovered-deceased-model (SIRD) is used when immunity is assumed [19]. In the case of COVID-19, different nations took different approaches in this regard.

SEIR models have been reported among the most popular tools to predict the notable outbreaks. To advance an SEIR model, often, the incubation period of an infected person is carefully estimated to achieve more accurate predictions. In the case of Varicella and Zika outbreaks, the SEIR models showed increased model accuracy [20,21]. To do so, SEIR models might assign the incubation period to a random variable. Furthermore, similar to the standard SIR-based models, SEIR models work on the ideology of the disease-free-equilibrium [22,23]. However, it should be noted that SEIR models can not fit well where the contact network is non-stationary through time [24]. Social mixing as a critical factor of non-stationarity determines the reproductive number R_0, which is the number of susceptible individuals for infection. The value of R_0 for COVID-19 was estimated to be 4, which greatly triggered the pandemic [1]. The lockdown measures aimed at reducing the R_0 value down to 1. Nevertheless, the SEIR models are reported to be difficult to fit in the case of COVID-19 due to the non-stationarity of mixing caused by nudging intervention measures. Therefore, to develop more accurate SIR-based models, in-depth information about the social movement and the quality of lockdown measures would be essential. Another drawback of SIR-based models is the short lead-time. For long-term prediction, the accuracy of the most SIR-based models declines. For the COVID-19 outbreak case of Italy, for instance, the accuracy of the model drops significantly. The performance of $f = 1$ for the lead-time of 120 h reduces to $f = 0.86$ for 144 h lead time [17]. Overall, the SIR-based models would be accurate if firstly the status of social interactions is stable. Secondly, class R can be computed precisely. To better estimate class R, several data sources can be integrated with SIR-based models, e.g., CCTVs, social media, mobile apps, and call data records. Nevertheless, using such systems are reported to still associate with significant complexity and uncertainties [25–32]. Due to the high level of uncertainties involved in the advancement of SIR-based models, the generalization ability is yet to be improved to achieve a scalable model with high performance [33].

Due to the presence of uncertainties and a high degree of complexity in advancing epidemiological models of the outbreak, machine learning has increasingly been seen as a potential technology. ML has already shown promising results in a contribution to developing better SIR-based models with higher performance with generalization ability and robustness [34–38]. Machine learning has already been recognized as a computing technique with great potential in outbreak prediction. The notable machine learning algorithms include, e.g., random forest for swine fever [39,40], neural networks for H1N1 flu, dengue fever, and Oyster norovirus [11,41,42], genetic programming for Oyster norovirus [43], classification and regression tree (CART) for Dengue [44], Bayesian Network for Dengue and Aedes [45], LogitBoost for Dengue [46], multi-regression and Naïve Bayes for Dengue outbreak prediction [47]. Machine learning has often been used as a complementary computation tool to enhance SIR-based models [11,39–49]. Nevertheless, there is a gap in using machine learning in the case of

COVID-19. However, several recent research works point out the enormous potential of machine learning for the fight against COVID-19 [49–52]. Machine learning delivered promising results in several aspects for mitigation and prevention and have been endorsed in the scientific community for, e.g., case identifications [53], classification of novel pathogens [54], modification of SIR-based models [55], diagnosis [56,57], survival prediction [58], and ICU demand prediction [59]. Furthermore, the non-peer reviewed sources suggest novel applications for fighting COVID-19 [60]. Among the applications of machine learning improvement of the existing models of prediction, identifying vulnerable groups, early diagnosis, the advancement of drugs delivery, evaluation of the probability of next pandemic, the advancement of the integrated systems for Spatio-temporal prediction, evaluating the risk of infection, advancing reliable biomedical knowledge graphs, and data mining the social networks are being noted.

As stated in our recent paper, machine learning can be used for data preprocessing. Improving the quality of data can particularly improve the quality of the SIR-based model. For instance, the number of cases reported by Worldometer is not precisely the number of infected cases (E in the SEIR model), or calculating the number of infectious people (I in SEIR) cannot be easily determined, as many people who might be infectious may not turn up for testing although the number of people who are admitted to hospital and deceased will not support R as most COVID-19 positive cases recover without entering the hospital. Considering this data problem, it is challenging to fit SEIR models satisfactorily. Considering such challenges, for future research, the ability of machine learning for estimation of the missing information on the number of exposed E or infected individuals can be evaluated. Along with the prediction of the outbreak, prediction of the total mortality rate (n(deaths)/n(infecteds)) is also essential to accurately estimate the number of potential patients in the critical situation and the required beds in intensive care units. Although the research is in the very early stage, the trend in outbreak prediction with machine learning can be classified in two directions. Firstly, improvement of the SIR-based models, e.g., [55,61], and secondly time-series prediction [62,63]. Consequently, the state-of-the-art machine learning methods for outbreak modeling suggest two major research gaps for machine learning to address. Firstly, the improvement of SIR-based models and secondly advancement in outbreak time series. Considering the drawbacks of the SIR-based models, machine learning should be able to contribute. This paper contributes to the advancement of time-series modeling and prediction of COVID-19. Although ML has been currently established in predicting several scientific phenomena [64–68], its advancement for pandemic modeling remains at the early stages [69]. More sophisticated ML methods are yet to be explored. A recent paper by Ardabili et al., [51] explored the potential of MLP and ANFIS in time series prediction of COVID-19 in several countries. The contribution of the present paper is to improve the quality of prediction by proposing a hybrid machine learning and compare the results with ANFIS. In the present paper, the time series of the total mortality is also included. This article continues as follows. Section 2 describes the methods and materials. The results are given in Section 3. Section 4 presents the conclusions.

2. Materials and Methods

2.1. Data

Dataset is related to the statistical reports of COVID-19 cases and mortality rate of Hungary, which is available online [70]. Figures 1 and 2 represent the total and daily reports of COVID-19 statistics, respectively, from 4 March to 19 April. The actual dataset includes the daily cases and the number of deaths from 4 March to 28 April. The data of 4 March to 19 April has been used for training, and the data of 20–28 April has been solely used for validation.
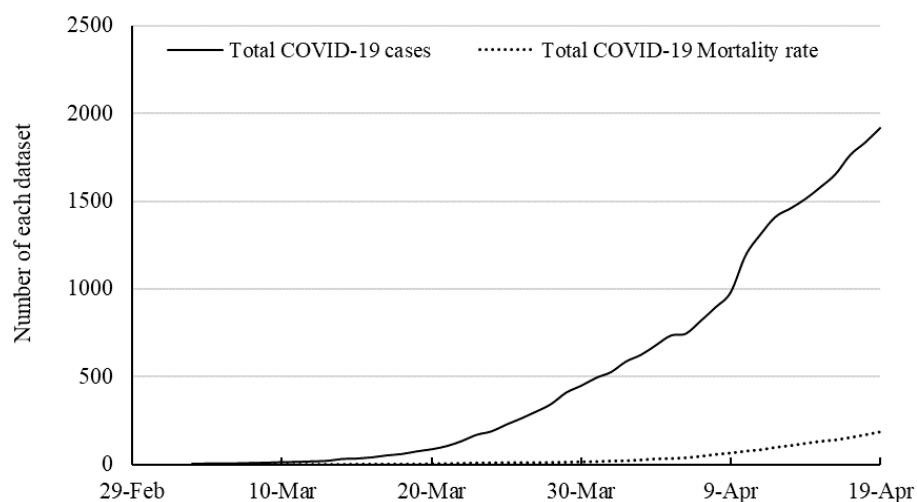
Figure 1. Total statistics for the number of cases and mortality rate of COVID-19.
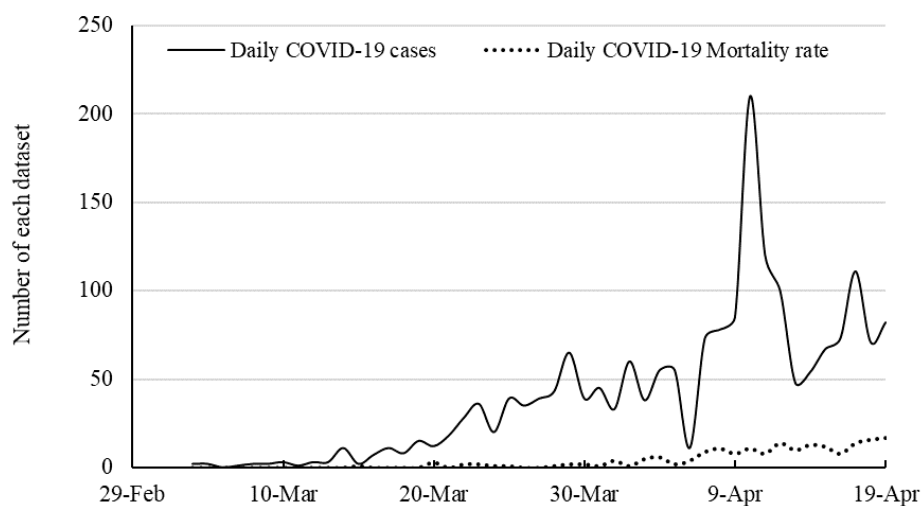

Figure 2. Daily statistics for the number of cases and mortality rate of COVID-19.

2.2. Methods and Modeling Strategy

In the present study, modeling is performed by machine learning methods. Training is the basis of these methods, as well as many artificial intelligence (AI) methods [71]. According to psychological and social literiture, creatures interact with their surroundings through the trial and error approach to achieve the optimum performance [72]. Based on this theory and using the ability of computers to repeat a set of instructions, these conditions can be provided for computer programs to interact with the environment by updating values and optimizing functions, according to the results of interaction with the environment to solve a problem or achieve a specific goal. How to update values and parameters in successive repetitions by a computer is called a training algorithm [72–74]. One of these methods is Neural Networks (NN), according to which the modeling of the connection of neurons in the human brain, software programs were designed to solve various problems. To solve these problems, operational NN, such as classification, clustering, or function approximation is performed using appropriate learning methods [75,76]. The training of the algorithm is the initial and important step for developing a model [77,78]. Developing a predictive AI model requires a dataset categorized into two sections i.e., input(s) (as independent variable(s)) and output(s) (as dependent variable(s)) [79].

In the present study, time-series data have been considered as the independent variables for the prediction of COVID-19 cases and mortality rate (as dependent variables). The Time-series dataset was prepared based on two scenarios, as described in Table 1. The first scenario categorizes the time-series

data into four inputs for the last four consequently odd days' cases or mortality rate for the prediction of x_t as the next day's case or mortality rate, and the second scenario categorizes the time-series data into four inputs for the last four consequently even days' cases or mortality rate for the prediction of x_t as the next day's case or mortality rate.

Table 1. Two proposed scenarios for time-series prediction of COVID-19 in Hungary.

	Inputs	Outputs
Scenario 1	x_{t-1}, x_{t-3}, x_{t-5} and x_{t-7}	x_t
Scenario 2	x_{t-2}, x_{t-4}, x_{t-6} and x_{t-8}	x_t

In the present study, the two robust hybrid methods of ANN algorithm, i.e., MLP-ICA and ANFIS, have been employed for developing the required models. The dataset is devised into two parts. One part is devoted to training and the second part, i.e., 20–28 April, is devoted to model validation.

2.2.1. Hybrid Multi-Layered Perceptron-Imperialist Competitive Algorithm (MLP-ICA)

Multi-layered-perceptron (MLP) is the frequently used ANN method for prediction and modeling purposes. This technique, as a single method, provides acceptable accuracy for prediction tasks in the simple and semi-complex datasets. However, in the case of doing modeling tasks in complex datasets, there is a need for more robust techniques [80,81]. For this reason, hybrid methods have been growing [80,81]. Hybrid methods contain a predictor and one or more optimizers [80,81]. The present study develops a hybrid MLP-ICA approach as a robust hybrid algorithm for generating a platform for predicting the COVID-19 cases and mortality rate in Hungary. The ICA is a method in the field of evolutionary calculations that seeks to find the optimal answer to various optimization problems. This algorithm, by mathematical modeling, provides a socio-political evolutionary algorithm for solving numerical optimization problems [82]. Like all algorithms in this category, the ICA constitutes an initial set of possible answers. These answers are known as countries in the ICA. The ICA gradually improves the initial responses and ultimately provides the appropriate solution to the optimization problem [82,83].

The algorithms are based on the policy of assimilation, imperialist competition, and revolution. This algorithm, by imitating the process of social, economic, and political development of countries and by mathematical modeling of parts of this process, provides operators in the form of a regular algorithm that can help to solve complex optimization problems [82,84]. In fact, this algorithm looks at the answers to the optimization problem in the form of countries and tries to gradually improve these answers during a repetitive process and eventually reach the optimal answer to the problem [82,84]. In the *Nvar* dimension optimization problem, a country is an array of *Nvar* × *1* length. This array is defined as follows:

$$country = [p1, \ p2, \ \ldots, \ pNvar] \tag{6}$$

To start the algorithm, the country's initial number is gererated ($N_{country}$). To select the N_{imp} as the best members of the population (countries with the lowest amount of cost function) as imperialists. The rest of the N_{col} countries form colonies, belonging to an empire. To divide the early colonies between the imperialists, imperialist owns a number of colonies, the number of which is proportional to its power [82,84]. The following figure symbolically shows how the colonies are divided among the colonial powers.

The integration of this model with the neural network causes the error in the network to be defined as a cost function, and with the changes in weights and biases, the output of the network improves, and the resulting error decreases [85]. The most important factors in training an ANN-ICA method contain the number of countries, imperialists, and decades and the number of neurons in the hidden layer, which can be defined in different ways. One of the common ways for defining them is trial and error.

In the present study, inputs and outputs of each scenario are imported into the model. The number of countries, imperialists, and decades have been defined by trial and error. For developing an MLP model, three architectures including 4–10-1, 4–14-1, and 4–18-1 (with 10, 14, and 18 neurons in the hidden layer, respectively) are used.

2.2.2. ANFIS

ANFIS falls as a member of the artificial neural networks (ANNs). The architecture of ANFIS is improved through the integration of the Takagi-Sugeno fuzzy system, i.e., a variation of fuzzy logic [86,87]. The ANFIS was coined in the early 90s and gained a widespread popularity within the research community for scientific modeling and estimation. This method was developed in the early 1990s. As ANFIS hybridizes the ANNs and Takagi-Sugeno fuzzy system, it benefits from the high performance of the both computing and learning technique for dealing with nonlinear functions [86,87]. Figure 3 presents the architecture of the developed ANFIS model.
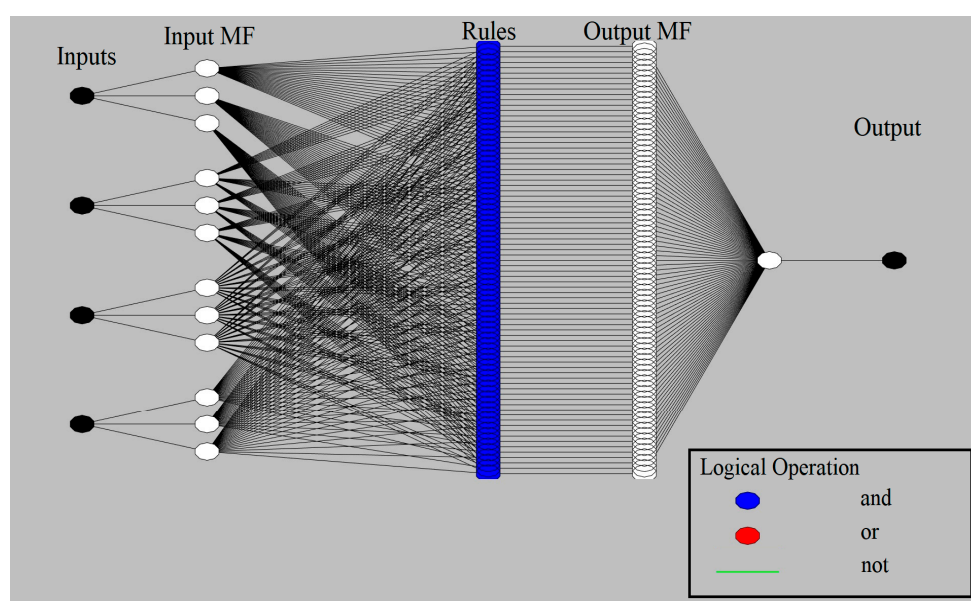

Figure 3. Architecture of the developed ANFIS.

As illustrated in Figure 3, ANFIS has five main layers [88–90]. The first layer is the inputs layer, which takes the parameters and imports them to the model. This layer is also called the input layer of the fuzzy system. The outputs of the first layer import to the second layer and carry prior values of membership functions (MFs). Fuzzy rules are concluded from the nodes on the second layer devoted to the degree of activity. Furthermore, the ANFIS's third layer multiplies the degree of activity of any rules. The fourth layer nomalizes the functions and nodes and further produces the outputs [91] and sends them to the output layer. The important factors for determining the accuracy of ANFIS are the number and type of MFs, the optimum method, and the output MF type [92–96].

Model of ANFIS is implemented through MATLAB's ANFIS toolbox. To initiliaze the model, the input parameters are set as the independent variables of each scenario, and the output variable was the number of cases or mortality rate. The ANFIS model was trained with three triangular, trapoizidal, and gaussian MFs. This step was performed in order to select the best MF. The output membership function type selected linear type because of its ability to further reduce errors. Training of FIS was done with optimum backpropagation (BP) method and 0 value of error tolerance.

2.2.3. Evaluation Criteria

Evaluations were conducted through calculating the standard values for determination coefficient, mean absolute percentage error root mean square error. These factors evaluate the target values and further estimate the model performance through calculating an index score. Such indexes are used to estimate the model accuracy [97,98]. Table 2 presents the evaluation criteria equations used in this study.

Table 2. Evaluation metrics used in this study.

Accuracy and Performance Index

$$\text{Determination coefficient} = \sqrt{\frac{N \sum (XY) - \sum (Y) \sum (Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum XY)^2]}}}$$

$$MAPE = \frac{1}{N}\left|\frac{X-Y}{X}\right| \times 100$$

$$RMSE = \sqrt{\frac{1}{N} \sum (A - P)^2}$$

In the evaluation metrics, *N* represents the number of data. In addition, *X* and *Y* are the predicted and desired values, respectively.

3. Results

The performance of the proposed algorithm is evaluated using both training and validation data. The training data are used to train the algorithm and define the best set of parameters to be used in ANFIS and MLP-ICA. After that, the best setup for each algorithm is used to predict outbreaks on the validation samples. It is worth mentioning that due to the lack of adequate sample data to avoid overfitting, the training is used to evaluate the model with higher performance.

3.1. Training Results

The training step for ANFIS is performed by employing three MF types, i.e., Triangular, Trapezoidal, and Gaussian as shown in Tables 3 and 4. Table 3 presents the training results of ANFIS for COVID-19 cases, and Table 4 presents the training results of ANFIS for the COVID-19 mortality rate in Hungary. The results have been compared using evaluation metrics of RMSE and MAPE. According to Table 3, Gaussian MF type with MF number 3, BP optimum method, and linear output MF provided the highest performance by the lowest value of RMSE. On the other hand, as is precise, scenario 1 provided the lowest RMSE in comparison with the scenario 2 for the selected MF. Therefore, it can be concluded that scenario 1 is suitable for modeling COVID-19 cases in contrast with scenario 2.

According to Table 4, it can be claimed that Gaussian MF provided the lowest error and highest accuracy compared with other MF types for the prediction of mortality rate. Also it can be claimed that, for the selected MF type, scenario 2 provides the highest performance compared with scenario 1 for the prediction of mortality rate.

Table 3. ANFIS training results (cases).

	MF Type	No. of MFs	Optimum Method	Output MF	Epoch No.	RMSE	MAPE
	Triangular	3	BP	Linear	233	248.06	38.44
Scenario 1	Trapezoidal	3	BP	Linear	253	158.59	35.67
	Gaussian	**3**	**BP**	**Linear**	**388**	**119.42**	**39.95**
	Triangular	3	BP	Linear	266	250.26	37.75
Scenario 2	Trapezoidal	3	BP	Linear	206	211.98	33.39
	Gaussian	**3**	**BP**	**Linear**	**285**	**182.76**	**34.6**

Table 4. ANFIS training results (mortality rate).

	MF Type	No. of MFs	Optimum Method	Output MF	Epoch No.	RMSE	MAPE
	Triangular	3	Back-propagation	Linear	204	31.14	47.12
Scenario 1	Trapezoidal	3	Back-propagation	Linear	135	31.09	44.36
	Gaussian	*3*	*Back-propagation*	*Linear*	*198*	*26.42*	*39.86*
	Triangular	3	Back-propagation	Linear	245	32.69	45.56
Scenario 2	Trapezoidal	3	Back-propagation	Linear	155	31.49	41.53
	Gaussian	*3*	*Back-propagation*	*Linear*	*260*	*23.95*	*33.95*

Tables 5 and 6 present results for the prediction of COVID-19 cases and mortality rate, respectively by MLP-ICA. According to Table 5, MLP architecture with 10 neurons in the hidden layer provided the highest accuracy for the prediction of the COVID-19 cases in the presence of both scenarios. However, Scenario 2 provided higher accuracy than scenario 1, according to the lowest RMSE value.

Table 5. MLP-ICA training results (cases).

	No. of Countries	No. of Decades	No. of Initial Imperialists	No. of Neurons	RMSE	MAPE (%)
	300	*40*	*50*	*10*	*37.30*	*23.15*
Scenario 1	300	40	50	14	37.39	23.82
	300	40	50	18	37.48	23.57
	300	*40*	*50*	*10*	*37.24*	*23.47*
Scenario 2	300	40	50	14	37.43	23.63
	300	40	50	18	38.29	22.15

Table 6. MLP-ICA training results (mortality rate).

	No. of Countries	No. of Decades	No. of Initial Imperialists	No. of Neurons	RMSE	MAPE
	250	55	70	10	1.902	37.1
Scenario 1	*250*	*55*	*70*	*14*	*1.9*	*36.99*
	250	55	70	18	1.901	36.99
	250	55	70	10	1.902	37.16
Scenario 2	250	55	70	14	1.917	37.18
	250	*55*	*70*	*18*	*1.902*	*37.09*

According to Table 6, neuron number 14 for scenario 1 and neuron number 18 for scenario 2 provided the highest accuracy in integrating by ICA method in comparison with other architectures. By comparing the evaluation criteria values, it can be concluded that scenario 1 is more suitable than scenario 2 for the prediction of mortality rate using MLP-ICA.

Figure 4 presents the plot diagram for the selected models according to Tables 3–6. In these figures, the vertical axes represent the "target values". In fact, the actual values and the horizontal axis is the "predicted values" or in another word the output value of the model. In these plots, the dash-line is the 1:1 line. The distance of each point from the dash-line discusses the accuracy of the prediction. Such that, points on the dash-line have the minimum error and increase the determination coefficient, and points that have distance from the dash line increase the error and reduce the accuracy in accordance with its distance and reduce the determination coefficient.

As is clear from Figure 5, ANN provides a higher determination coefficient (0.9963, 0.9963, 0.9987, and 0.9987, respectively for the prediction of COVID-19 cases and COVID-19 mortality rate in the presence of scenario 1 and scenario (see the Figure 5)) than ANFIS. Also, the ability of ANN in the prediction of the COVID-19 mortality rate is higher than that for the prediction of COVID-19 cases. However, ANFIS provides a little different behavior than ANN. Such that, in ANFIS, scenario 1 provides the highest determination coefficient for the prediction of COVID-19 cases but scenario 2

provides the highest correlation coefficient for the prediction of mortality rate (0.9689, 0.934, 0.8909 and 0.9427, respectively for the prediction of COVID-19 cases and COVID-19 mortality rate in the presence of scenario 1 and scenario (see the Figure 5)).
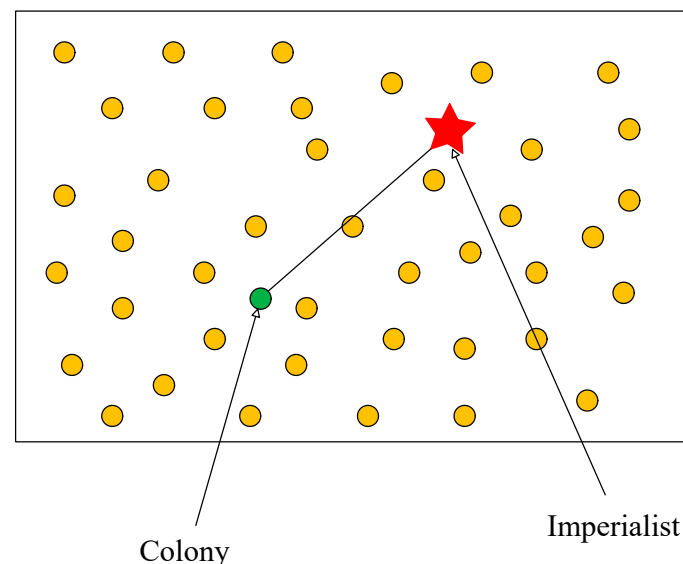

Figure 4. The initial empires generation [82].

From considering Scen. 1 for mortality model using MLP-ICA method and Scen. 2 for mortality model using MLP-ICA method, the differences are negligible. It is, therefore, worth mentioning that using different scenarios for data sampling has a minimum effect on performance. Thus, using either odd or even days for data sampling gives relatively acceptable results. Furthermore, Figure 6 represents the deviation from target values for the COVID-19 cases from 4 March to 19 April. According to Figure 6, MLP-ICA scenario 2 provides a lower deviation from the target value followed by MLP-ICA in the presence of scenario 1 than the ANFIS in the presence of both scenarios.

Figure 7 presents the deviation from target values for the COVID-19 mortality rate from 4 March to 19 April. According to Figure 7, MLP-ICA in the presence of scenario 1 provides a lower deviation from the target value followed by MLP-ICA in the presence of scenario 2 than the ANFIS in the presence of both situations.
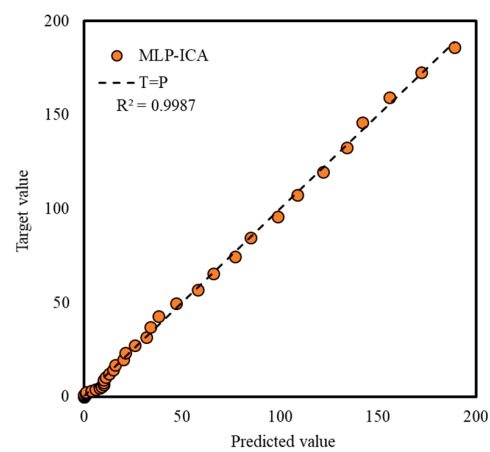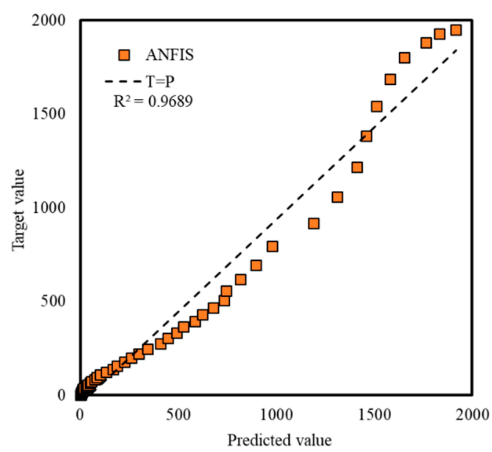

Scen. 1 (Cases, MLP-ICA)   Scen. 2 (Cases, MLP-ICA)
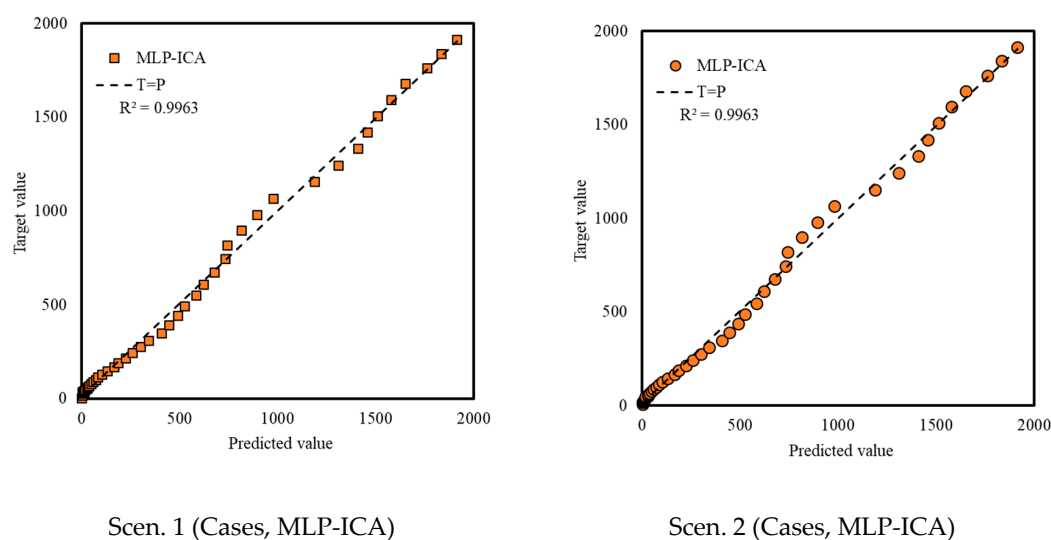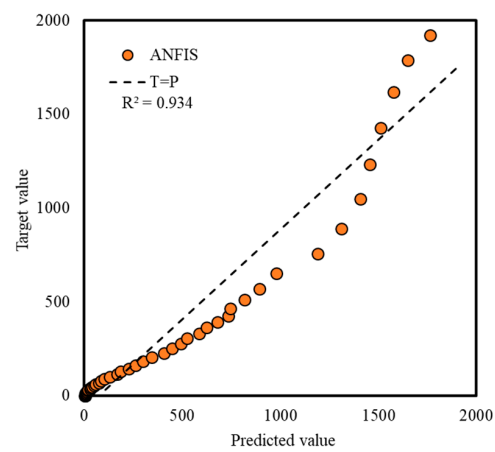
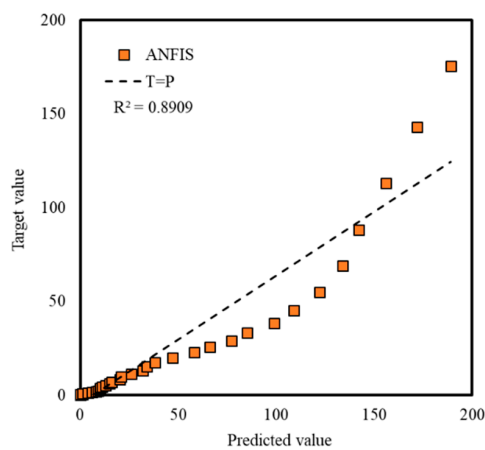**Figure 5.** *Cont.*
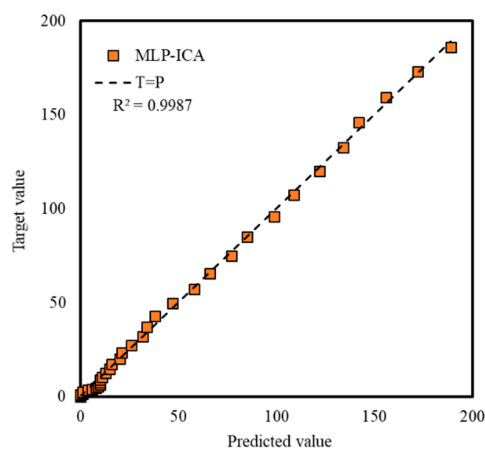
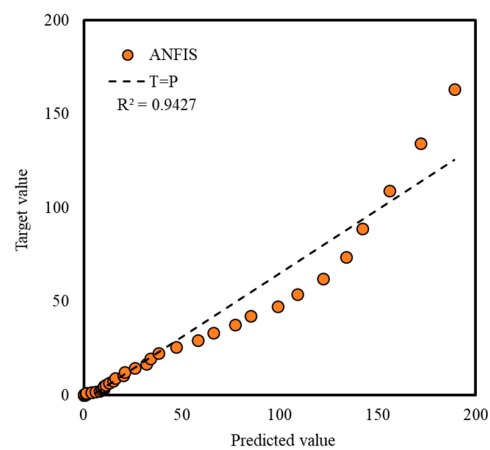Scen. 1 (Mortality, MLP-ICA)

Scen. 2 (Mortality, MLP-ICA)

Scen. 1 (Cases, ANFIS)

Scen. 2 (Cases, ANFIS)

Scen. 1 (Mortality, ANFIS)

Scen. 2 (Mortality, ANFIS)

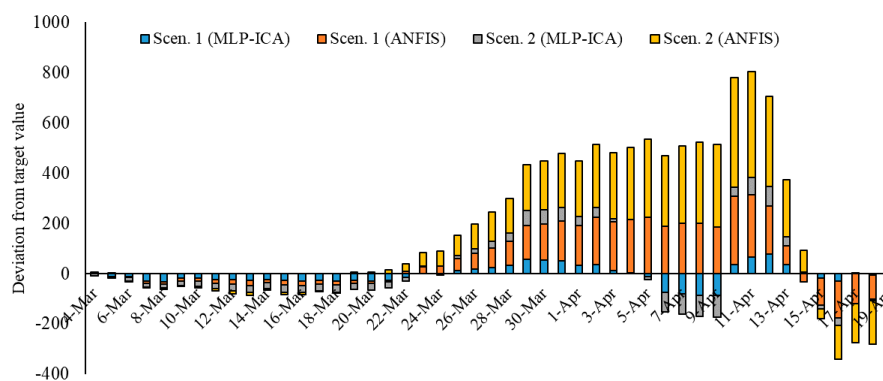**Figure 5.** Plot Diagrams for the prediction of COVID-19 cases and mortality rate.

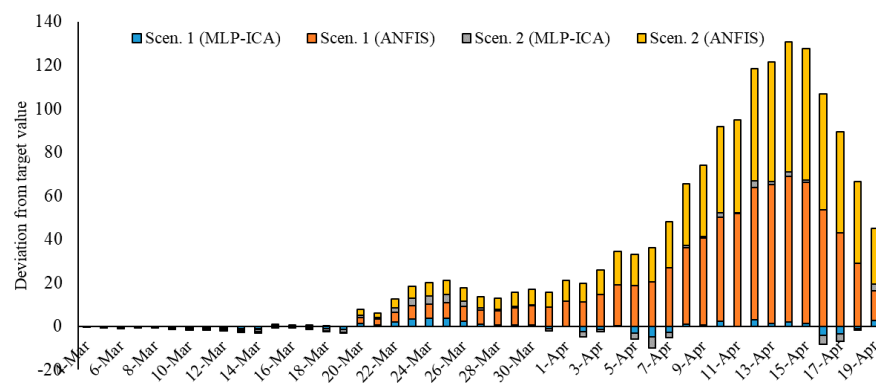**Figure 6.** Deviation from target value (Cases).

Figure 7. Deviation from target value (mortality rate).

By evaluating the modeling methods in the last sections, it was decided to select the MLP-ICA in the presence of scenario 2 for the prediction of COVID-19 outbreak and MLP-ICA in the presence of scenario 1 for the prediction of COVID-19 mortality rate in Hungary. Predictions were performed in two stages. The first stage for total prediction and the second one for daily prediction. Figures 8 and 9 present total cases and total mortality rate, respectively, and Figures 10 and 11 present the daily Prediction of the results from 20 April to 30 July. Each figure has two sections, including the reported statistic (from 24 March to 19 April) and the predicted by the selected model (from 20 April to 30 July).
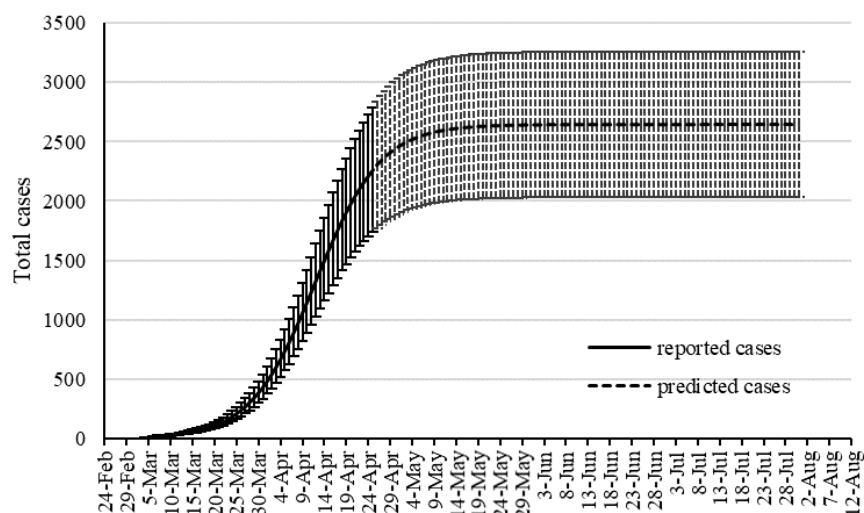

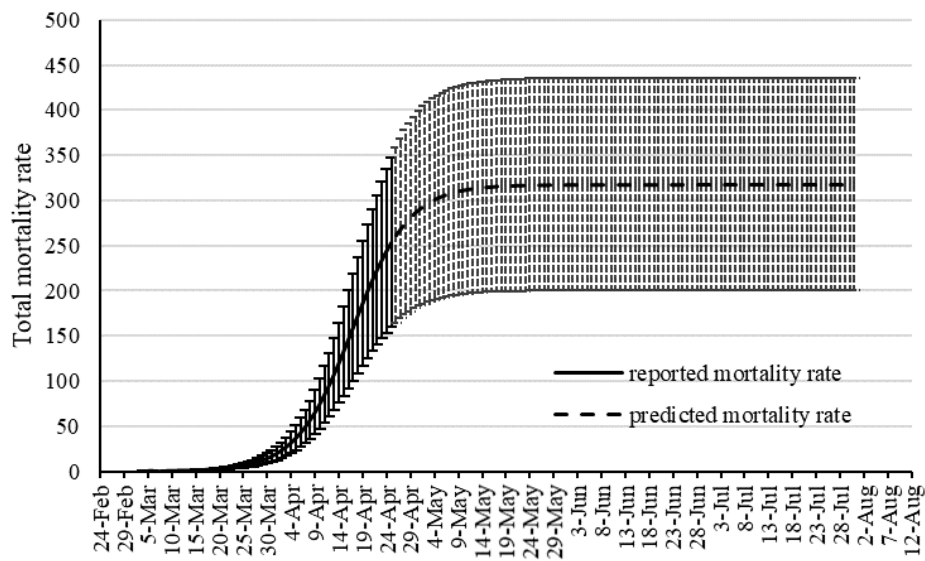**Figure 8.** Total outbreak prediction for MLP-ICA.

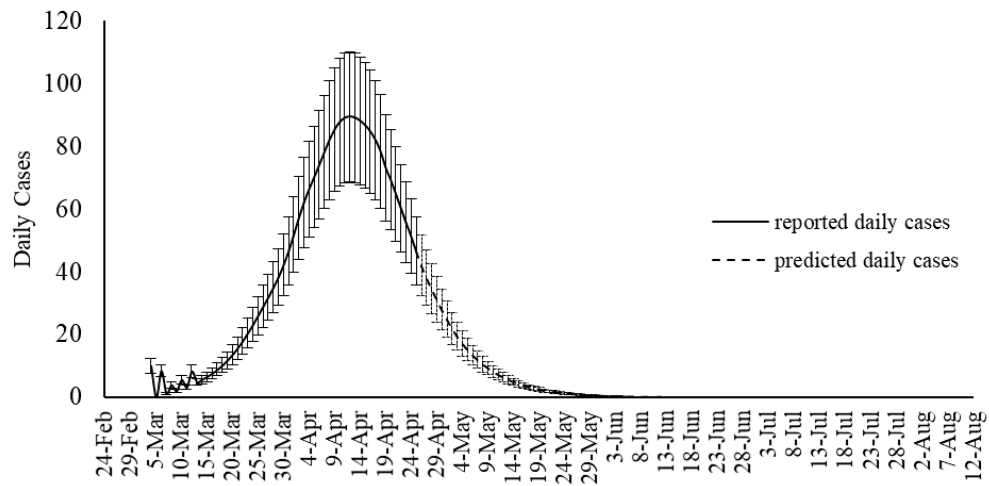**Figure 9.** Total mortality rate prediction for MLP-ICA.

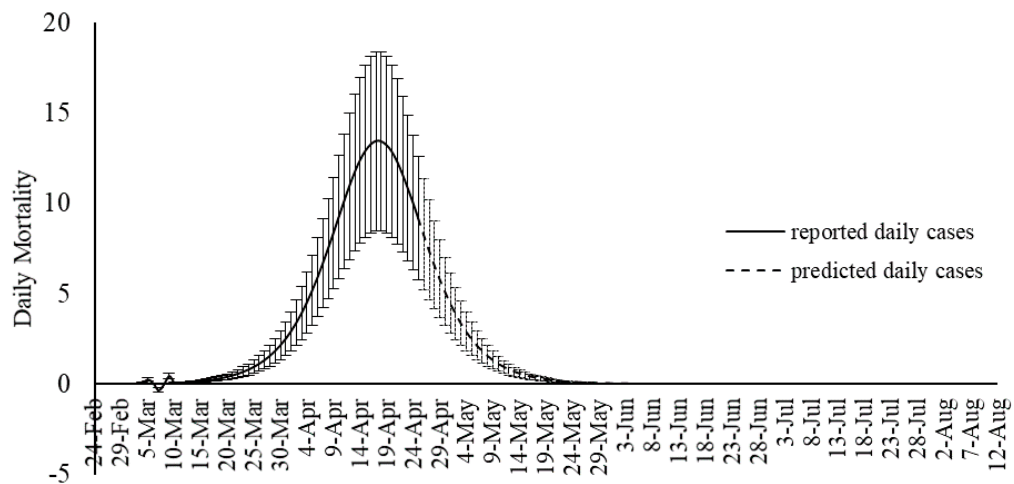**Figure 10.** Daily outbreak prediction for MLP-ICA.

Figure 11. Daily mortality rate prediction for MLP-ICA.

3.2. Validation

Table 7 represents the validation of the MPL-ICA and ANFIS models for the period of 20–28 April. The proposed model of MPL-ICA presented promising values for RMSE and determination coefficient for the prediction of both outbreak and total mortality.

Table 7. Validation of the models for 9 days.

	Total Cases		Total Mortality Rate	
For 20–28 April	MPL-ICA	ANFIS	MLP-ICA	ANFIS
RMSE	167.88	194.10	8.32	15.25
Determination coefficient	0.9971	0.9563	0.9986	0.7491

4. Discussion

Considering uncertainties in modeling with the SIR-based models, the proposed approach shows the potential to outperform the commonly used prediction tools in the case of Hungary. More work is required to confirm if this technique is adequate in all cases and for different population types and sizes. Nonetheless, the learning approach could overcome imperfect input data. Incomplete catalogs can occur because infected persons are asymptotic, not tested, or not listed in databases. Tests in a closed environment such as large aircraft carriers in France and the US reported that up to 60% of the infected personnel might be asymptomatic. Of course, military personals are not representative of large and mixed populations. Nonetheless, it shows that false negatives can be abundant. In emerging countries, access to laboratory equipment required for testing is extremely limited. This will introduce a bias in the counting. Finally, it is unclear if all the cases are registered. In the UK, for example, it took public pressure for the government to make the casualties in retirement hospices known. In addition, there is still growing suspicion in the community that a few countries might have reported false data for political reasons. At the same time, national governments and local administrations implemented containment measures such as confinement and social distancing. These actions have a huge impact on transmissions and, thus, on cases and casualties. Access to modern medical facilities is also a parameter that will mainly influence the number of casualties. All these aspects will influence traditional estimation procedures, whereas learning algorithms might be able to adapt, especially if multiple datasets are available for a given region. Not only can our approach outperform the commonly used SIR-based models, but it requires fewer input data to estimate the trends. While we provide promising results for Hungarian data we need to further test such novel approaches on other databases. Nonetheless, the presented results are showing signs of success and should incite the community to implement these new tools rapidly.

5. Conclusions

Although SIR-based models been widely used for modeling the COVID-19 outbreak, they include some degree of uncertainties. Several advancements are emerging to improve the quality of SIR-based models suitable to the COVID-19 outbreak. As an alternative to the SIR-based models, this study proposed machine learning as a new trend in advancing outbreak models. The machine learning approach makes no assumption on the pandemic and spread of the infection. Instead, it predicts the time series of the infected cases as well as total mortality cases. In this study, the hybrid machine learning model of MLP-ICA and ANFIS is used to predict the COVID-19 outbreak in Hungary. The models predict that by late May, the outbreak and the total morality will drop substantially. Based on the promising results reported in this study, and due to the complex phenomenon of COVID-19 outbreak, this study, as an alternative modeling strategy, suggests machine learning as a potential technology to be considered to model the outbreak. However, further research would be essential to validate the results and improve the quality of prediction.

In this study, two scenarios were proposed for sampling the data. Scenario 1 considered sampling the odd days, and Scenario 2 used even days for training the data. Training the two machine learning models of ANFIS and MLP-ICA, were considered through using two scenarios. It is concluded that using different scenarios for data sampling has a minimum effect on the model performance. A detailed investigation was carried out to explore the most suitable number of neurons. Furthermore, the performance of the proposed algorithm is evaluated using both training and validation data. The training data are used to train the algorithm and define the best set of parameters to be used in ANFIS and MLP-ICA. After that, the best setup for each algorithm is used to predict outbreaks on the validation samples. The validation is performed for 9 days with promising results, which confirms the model accuracy. In this study, due to the lack of adequate sample data to avoid overfitting, the training is used to choose and evaluate the model with higher performance. In future research, as the COVID-19 progresses in time and with the availability of more sample data, further testing and validation can be used to better evaluate the models.

Both models showed promising results in terms of predicting the time series without the assumptions that epidemiological models require. Both machine learning models, as an alternative to epidemiological models, showed potential in predicting COVID-19 outbreak as well as estimating total mortality. Yet, MLP-ICA outperformed ANFIS with delivering accurate results on validation samples. Considering the availability of a small amount of training data, further investigation would be essential to explore the true capability of the proposed hybrid model. It is expected that the model maintains its accuracy as long as no major interruption occurs. For instance, if other outbreaks would initiate in the other cities, or the prevention regime changes, naturally the model will not maintain its accuracy. For future studies, advancing deep learning and deep reinforcement learning models is strongly encouraged for comparative studies on various ML models for individual countries.

Nomenclatures

Adaptive network-based fuzzy inference system	ANFIS	Susceptible-infectious-susceptible	SIS
Multi-layered perceptron	MLP	Mean square error	MSE
Call data record	CDR	Artificial intelligence	AI
Classification and regression tree	CART	Artificial neural network	ANN
Evolutionary algorithms	EA	Susceptible-exposed-infected-recovered	SEIR
Imperialist competitive algorithm	ICA	Random Forest	RF
Maternally-derived-immunity-susceptible-exposed-infected-recovered-susceptible	MSEIRS	Maternally-derived-immunity-susceptible-exposed-infected-recovered	MSEIR
Maternally-derived-immunity-susceptible-infected-recovered	MSIR	Susceptible-exposed-infected- susceptible	SEIS
Membership function	MF	Machine learning	ML
Particle swarm optimization	PSO	Neural Networks	NN
Susceptible-infected-recovered	SIR	Root mean square error	RMSE
Susceptible-infected-recovered-deceased	SIRD	Backpropagation	BP

References

1. Gorbalenya, A.E.; Baker, S.C.; Baric, R.S.; de Groot Raoul, J.; Drosten, C.; Gulyaeva, A.A.; Haagmans, B.L.; Lauber, C.; Leontovich, A.M.; Neuman, B.W. The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536–544.

2. Chan, J.F.W.; Yuan, S.; Kok, K.H.; To, K.K.W.; Chu, H.; Yang, J.; Xing, F.; Liu, J.; Yip, C.C.Y.; Poon, R.W.S. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *Lancet* **2020**, *395*, 514–523. [CrossRef]

3. World Health Organization. *Novel Coronavirus (2019-nCoV): Situation Report-3*; WHO: Geneva, Switzerland, 2020; Available online: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200123-sitrep-3-2019-ncov.pdf (accessed on 28 April 2020).

4. World Health Organization. *Coronavirus Disease 2019 (COVID-19): Situation Report-72*; WHO: Geneva, Switzerland, 2020; Available online: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200401-sitrep-72-covid-19.pdf?sfvrsn=3dd8971b_2 (accessed on 28 April 2020).

5. Remuzzi, A.; Remuzzi, G. COVID-19 and Italy: What next? *Lancet* **2020**. [CrossRef]

6. Ivanov, D. Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. *Transp. Res. Part E Logist. Transp. Rev.* **2020**, *136*. [CrossRef] [PubMed]

7. Koolhof, I.S.; Gibney, K.B.; Bettiol, S.; Charleston, M.; Wiethoelter, A.; Arnold, A.L.; Campbell, P.T.; Neville, P.J.; Aung, P.; Shiga, T.; et al. The forecasting of dynamical Ross River virus outbreaks: Victoria, Australia. *Epidemics* **2020**, *30*. [CrossRef]

8. Rypdal, M.; Sugihara, G. Inter-outbreak stability reflects the size of the susceptible pool and forecasts magnitudes of seasonal epidemics. *Nat. Commun.* **2019**, *10*. [CrossRef]

9. Scarpino, S.V.; Petri, G. On the predictability of infectious disease outbreaks. *Nat. Commun.* **2019**, *10*. [CrossRef]

10. Zhan, Z.; Dong, W.; Lu, Y.; Yang, P.; Wang, Q.; Jia, P. Real-Time Forecasting of Hand-Foot-and-Mouth Disease Outbreaks using the Integrating Compartment Model and Assimilation Filtering. *Sci. Rep.* **2019**, *9*. [CrossRef]

11. Koike, F.; Morimoto, N. Supervised forecasting of the range expansion of novel non-indigenous organisms: Alien pest organisms and the 2009 H1N1 flu pandemic. *Glob. Ecol. Biogeogr.* **2018**, *27*, 991–1000. [CrossRef]

12. Dallas, T.A.; Carlson, C.J.; Poisot, T. Testing predictability of disease outbreaks with a simple model of pathogen biogeography. *R. Soc. Open Sci.* **2019**, *6*. [CrossRef]

13. De Groot, M.; Ogris, N. Short-term forecasting of bark beetle outbreaks on two economically important conifer tree species. *For. Ecol. Manag.* **2019**, *450*. [CrossRef]

14. Kelly, J.D.; Park, J.; Harrigan, R.J.; Hoff, N.A.; Lee, S.D.; Wannier, R.; Selo, B.; Mossoko, M.; Njoloko, B.; Okitolonda-Wemakoy, E.; et al. Real-time predictions of the 2018–2019 Ebola virus disease outbreak in the Democratic Republic of the Congo using Hawkes point process models. *Epidemics* **2019**, *28*. [CrossRef] [PubMed]

15. Miller, J.C. A note on the derivation of epidemic final sizes. *Bull. Math. Biol.* **2012**, *74*, 2125–2141. [CrossRef] [PubMed]

16. Miller, J.C. Mathematical models of SIR disease spread with combined non-sexual and sexual transmission routes. *Infect. Dis. Model.* **2017**, *2*, 35–55. [CrossRef] [PubMed]

17. Maier, B.F.; Brockmann, D. Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in Mainland China. *medRxiv* **2020**. [CrossRef]

18. Werkman, M.; Green, D.; Murray, A.G.; Turnbull, J. The effectiveness of fallowing strategies in disease control in salmon aquaculture assessed with an SIS model. *Prev. Vet. Med.* **2011**, *98*, 64–73. [CrossRef] [PubMed]

19. Osemwinyen, A.C.; Diakhaby, A. Mathematical modelling of the transmission dynamics of ebola virus. *Appl. Comput. Math.* **2015**, *4*, 313–320. [CrossRef]

20. Pan, J.R.; Huang, Z.Q.; Chen, K. Evaluation of the effect of varicella outbreak control measures through a discrete time delay SEIR model. *Chin. J. Prev. Med.* **2012**, *46*, 343–347.

21. Zha, W.T.; Pang, F.R.; Zhou, N.; Wu, B.; Liu, Y.; Du, Y.B.; Hong, X.Q.; Lv, Y. Research about the optimal strategies for prevention and control of varicella outbreak in a school in a central city of China: Based on an SEIR dynamic model. *Epidemiol. Infect.* **2020**. [CrossRef]

22. Dantas, E.; Tosin, M.; Cunha, A., Jr. Calibration of a SEIR–SEI epidemic model to describe the Zika virus outbreak in Brazil. *Appl. Math. Comput.* **2018**, *338*, 249–259. [CrossRef]

23. Leonenko, V.N.; Ivanov, S.V. Fitting the SEIR model of seasonal influenza outbreak to the incidence data for Russian cities. *Russ. J. Numer. Anal. Math. Model.* **2016**, *31*, 267–279. [CrossRef]

24. Imran, M.; Usman, M.; Dur-e-Ahmad, M.; Khan, A. Transmission Dynamics of Zika Fever: A SEIR Based Model. *Differ. Equ. Dyn. Syst.* **2020**. [CrossRef]

25. Miranda, G.H.B.; Baetens, J.M.; Bossuyt, N.; Bruno, O.M.; De Baets, B. Real-time prediction of influenza outbreaks in Belgium. *Epidemics* **2019**, *28*. [CrossRef]

26. Sinclair, D.R.; Grefenstette, J.J.; Krauland, M.G.; Galloway, D.D.; Frankeny, R.J.; Travis, C.; Burke, D.S.; Roberts, M.S. Forecasted Size of Measles Outbreaks Associated With Vaccination Exemptions for Schoolchildren. *JAMA Netw. Open* **2019**, *2*, e199768. [CrossRef] [PubMed]

27. Zhao, S.; Musa, S.S.; Fu, H.; He, D.; Qin, J. Simple framework for real-time forecast in a data-limited situation: The Zika virus (ZIKV) outbreaks in Brazil from 2015 to 2016 as an example. *Parasites Vectors* **2019**, *12*. [CrossRef]

28. Fast, S.M.; Kim, L.; Cohn, E.L.; Mekaru, S.R.; Brownstein, J.S.; Markuzon, N. Predicting social response to infectious disease outbreaks from internet-based news streams. *Ann. Oper. Res.* **2018**, *263*, 551–564. [CrossRef]

29. McCabe, C.M.; Nunn, C.L. Effective network size predicted from simulations of pathogen outbreaks through social networks provides a novel measure of structure-standardized group size. *Front. Vet. Sci.* **2018**, *5*. [CrossRef]

30. Bragazzi, N.L.; Mahroum, N. Google trends predicts present and future plague cases during the plague outbreak in Madagascar: Infodemiological study. *J. Med. Internet Res.* **2019**, *21*. [CrossRef]

31. Jain, R.; Sontisirikit, S.; Iamsirithaworn, S.; Prendinger, H. Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data. *BMC Infect. Dis.* **2019**, *19*. [CrossRef]

32. Kim, T.H.; Hong, K.J.; Shin, S.D.; Park, G.J.; Kim, S.; Hong, N. Forecasting respiratory infectious outbreaks using ED-based syndromic surveillance for febrile ED visits in a Metropolitan City. *Am. J. Emerg. Med.* **2019**, *37*, 183–188. [CrossRef]

33. Reis, J.; Yamana, T.; Kandula, S.; Shaman, J. Superensemble forecast of respiratory syncytial virus outbreaks at national, regional, and state levels in the United States. *Epidemics* **2019**, *26*, 1–8. [CrossRef] [PubMed]

34. Burke, R.M.; Shah, M.P.; Wikswo, M.E.; Barclay, L.; Kambhampati, A.; Marsh, Z.; Cannon, J.L.; Parashar, U.D.; Vinjé, J.; Hall, A.J. The Norovirus Epidemiologic Triad: Predictors of Severe Outcomes in US Norovirus Outbreaks, 2009–2016. *J. Infect. Dis.* **2019**, *219*, 1364–1372. [CrossRef] [PubMed]

35. Carlson, C.J.; Dougherty, E.; Boots, M.; Getz, W.; Ryan, S.J. Consensus and conflict among ecological forecasts of Zika virus outbreaks in the United States. *Sci. Rep.* **2018**, *8*. [CrossRef] [PubMed]

36. Kleiven, E.F.; Henden, J.A.; Ims, R.A.; Yoccoz, N.G. Seasonal difference in temporal transferability of an ecological model: Near-term predictions of lemming outbreak abundances. *Sci. Rep.* **2018**, *8*. [CrossRef] [PubMed]

37. Rivers-Moore, N.A.; Hill, T.R. A predictive management tool for blackfly outbreaks on the Orange River, South Africa. *River Res. Appl.* **2018**, *34*, 1197–1207. [CrossRef]

38. Yin, R.; Tran, V.H.; Zhou, X.; Zheng, J.; Kwoh, C.K. Predicting antigenic variants of H1N1 influenza virus based on epidemics and pandemics using a stacking model. *PLoS ONE* **2018**, *13*. [CrossRef]

39. Liang, R.; Lu, Y.; Qu, X.; Su, Q.; Li, C.; Xia, S.; Liu, Y.; Zhang, Q.; Cao, X.; Chen, Q.; et al. Prediction for global African swine fever outbreaks based on a combination of random forest algorithms and meteorological data. *Transbound. Emer. Dis.* **2020**, *67*, 935–946. [CrossRef]

40. Tapak, L.; Hamidi, O.; Fathian, M.; Karami, M. Comparative evaluation of time series models for predicting influenza outbreaks: Application of influenza-like illness data from sentinel sites of healthcare centers in Iran. *BMC Res. Notes* **2019**, *12*. [CrossRef]

41. Anno, S.; Hara, T.; Kai, H.; Lee, M.A.; Chang, Y.; Oyoshi, K.; Mizukami, Y.; Tadono, T. Spatiotemporal dengue fever hotspots associated with climatic factors in taiwan including outbreak predictions based on machine-learning. *Geospat. Health* **2019**, *14*, 183–194. [CrossRef]

42. Chenar, S.S.; Deng, Z. Development of artificial intelligence approach to forecasting oyster norovirus outbreaks along Gulf of Mexico coast. *Environ. Int.* **2018**, *111*, 212–223. [CrossRef]

43. Chenar, S.S.; Deng, Z. Development of genetic programming-based model for predicting oyster norovirus outbreak risks. *Water Res.* **2018**, *128*, 20–37. [CrossRef] [PubMed]

44. Titus Muurlink, O.; Stephenson, P.; Islam, M.Z.; Taylor-Robinson, A.W. Long-term predictors of dengue outbreaks in Bangladesh: A data mining approach. *Infect. Dis. Model.* **2018**, *3*, 322–330. [CrossRef] [PubMed]

45. Raja, D.B.; Mallol, R.; Ting, C.Y.; Kamaludin, F.; Ahmad, R.; Ismail, S.; Jayaraj, V.J.; Sundram, B.M. Artificial intelligence model as predictor for dengue outbreaks. *Malays. J. Public Health Med.* **2019**, *19*, 103–108.

46. Iqbal, N.; Islam, M. Machine learning for dengue outbreak prediction: A performance evaluation of different prominent classifiers. *Informatica* **2019**, *43*, 363–371. [CrossRef]

47. Agarwal, N.; Koti, S.R.; Saran, S.; Senthil Kumar, A. Data mining techniques for predicting dengue outbreak in geospatial domain using weather parameters for New Delhi, India. *Curr. Sci.* **2018**, *114*, 2281–2291. [CrossRef]

48. Mezzatesta, S.; Torino, C.; De Meo, P.; Fiumara, G.; Vilasi, A. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. *Comput. Methods Programs Biomed.* **2019**, *177*, 9–15. [CrossRef]

49. Alimadadi, A.; Aryal, S.; Manandhar, I.; Munroe, P.B.; Joe, B.; Cheng, X. *Artificial Intelligence and Machine Learning to Fight COVID-19*; American Physiological Society: Bethesda, MD, USA, 2020.

50. Alimadadi, A.; Aryal, S.; Manandhar, I.; Munroe, P.; Joe, B.; Cheng, X. Artificial Intelligence and Machine Learning to Fight COVID-19. *Physiol. Genom.* **2020**. [CrossRef]

51. Ardabili, S.F.; Mosavi, A.; Ghamisi, P.; Ferdinand, F.; Varkonyi-Koczy, A.R.; Reuter, U.; Rabczuk, T.; Atkinson, P.M. COVID-19 Outbreak Prediction with Machine Learning. *medRxiv* **2020**. [CrossRef]

52. Miralles-Pechuán, L.; Jiménez, F.; Ponce, H.; Martínez-Villaseñor, L. A Deep Q-learning/genetic Algorithms Based Novel Methodology For Optimizing Covid-19 Pandemic Government Actions. *arXiv* **2020**, arXiv:2005.07656.

53. Rao, A.S.S.; Vazquez, J.A. Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey in the populations when cities/towns are under quarantine. *Infect. Control Hosp. Epidemiol.* **2020**, 1–18. [CrossRef]

54. Randhawa, G.S.; Soltysiak, M.P.; El Roz, H.; de Souza, C.P.; Hill, K.A.; Kari, L. Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study. *bioRxiv* **2020**. [CrossRef] [PubMed]

55. Yang, Z.; Zeng, Z.; Wang, K.; Wong, S.S.; Liang, W.; Zanin, M.; Liu, P.; Cao, X.; Gao, Z.; Mai, Z. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J. Thorac. Dis.* **2020**, *12*, 165. [CrossRef] [PubMed]

56. Barstugan, M.; Ozkaya, U.; Ozturk, S. Coronavirus (covid-19) classification using ct images by machine learning methods. *arXiv* **2020**, arXiv:2003.09424.

57. Apostolopoulos, I.D.; Mpesiana, T.A. Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **2020**, *1*. [CrossRef]

58. Yan, L.; Zhang, H.T.; Xiao, Y.; Wang, M.; Sun, C.; Liang, J.; Li, S.; Zhang, M.; Guo, Y.; Xiao, Y. Prediction of survival for severe Covid-19 patients with three clinical features: Development of a machine learning-based prognostic model with clinical data in Wuhan. *medRxiv* **2020**. [CrossRef]

59. Grasselli, G.; Pesenti, A.; Cecconi, M. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: Early experience and forecast during an emergency response. *JAMA* **2020**. [CrossRef]

60. Santosh, K. AI-driven tools for coronavirus outbreak: Need of active learning and cross-population train/test models on multitudinal/multimodal data. *J. Med. Syst.* **2020**, *44*, 93. [CrossRef]

61. Pandey, G.; Chaudhary, P.; Gupta, R.; Pal, S. SEIR and Regression Model based COVID-19 outbreak predictions in India. *arXiv* **2020**, arXiv:2004.00958.

62. Liu, D.; Clemente, L.; Poirier, C.; Ding, X.; Chinazzi, M.; Davis, J.T.; Vespignani, A.; Santillana, M. A machine learning methodology for real-time forecasting of the 2019-2020 COVID-19 outbreak using Internet searches, news alerts, and estimates from mechanistic models. *arXiv* **2020**, arXiv:2004.04019.

63. Yan, L. Prediction of criticality in patients with severe Covid-19 infection using three clinical features: A machine learning-based prognostic model with clinical data in Wuhan. *medRxiv* **2020**. [CrossRef]

64. Nosratabadi, S.; Mosavi, A.; Duan, P.; Ghamisi, P. Data Science in Economics. *arXiv* **2020**, arXiv:2003.13422.

65. Mosavi, A.; Ozturk, P.; Chau, K.W. Flood prediction using machine learning models: Literature review. *Water* **2018**, *10*, 1536. [CrossRef]

66. Sheikh Khozani, Z.; Sheikhi, S.; Mohtar, W.H.M.W.; Mosavi, A. Forecasting shear stress parameters in rectangular channels using new soft computing methods. *PLoS ONE* **2020**, *15*, e0229731. [CrossRef] [PubMed]

67. Lorestani, Y.; Feiznia, S.; Mosavi, A.; Nádai, L. Hybrid Model of Morphometric Analysis and Statistical Correlation for Hydrological Units Prioritization. 2020. Available online: https://easychair.org/publications/preprint/lND9 (accessed on 5 May 2020).

68. Datta, A.; Si, S.; Biswas, S. Complete Statistical Analysis to Weather Forecasting. In *Computational Intelligence in Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 751–763.

69. Suzuki, Y.; Suzuki, A. Machine learning model estimating number of COVID-19 infection cases over coming 24 days in every province of South Korea (XGBoost and MultiOutputRegressor). *medRxiv* **2020**. [CrossRef]

70. Worldometer. Available online: https://www.worldometers.info/coronavirus/country/hungary/ (accessed on 28 April 2020).

71. Mojrian, S.; Pinter, G.; Joloudari, J.H.; Felde, I.; Nabipour, N.; Nádai, L.; Mosavi, A. Hybrid Machine Learning Model of Extreme Learning Machine Radial basis function for Breast Cancer Detection and Diagnosis; a Multilayer Fuzzy Expert System. *arXiv* **2019**, arXiv:1910.13574.

72. Mosavi, A.; Ardabili, S.; Várkonyi-Kóczy, A.R. List of Deep Learning Models. In *Engineering for Sustainable Future, Lecture Notes in Networks and Systems*; Springer: Cham, Switzerland, 2019.

73. Choubin, B.; Moradi, E.; Golshan, M.; Adamowski, J.; Sajedi-Hosseini, F.; Mosavi, A. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci. Total Environ.* **2019**, *651*, 2087–2096. [CrossRef]

74. Ardabili, S.; Mosavi, A.; Dehghani, M.; Várkonyi-Kóczy, A.R. Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review. In Proceedings of the International Conference on Global Research and Education, Balatonfüred, Hungary, 4–7 September 2019; pp. 52–62.

75. Samadianfard, S.; Hashemi, S.; Kargar, K.; Izadyar, M.; Mostafaeipour, A.; Mosavi, A.; Nabipour, N.; Shamshirband, S. Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm. *Energy Rep.* **2020**, *6*, 1147–1159. [CrossRef]

76. Ardabili, S.; Mosavi, A.; Varkonyi-Koczy, A. Systematic Review of Deep Learning and Machine Learning Models in Biofuels Research. In *Engineering for Sustainable Future, Lecture Notes in Networks and Systems*; Springer: Cham, Switzerland, 2019.

77. Nádai, L.; Imre, F.; Ardabili, S.; Gundoshmian, T.M.; Gergo, P.; Mosavi, A. Performance Analysis of Combine Harvester using Hybrid Model of Artificial Neural Networks Particle Swarm Optimization. *arXiv* **2020**, arXiv:2002.11041.

78. Nosratabadi, S.; Karoly, S.; Beszedes, B.; Felde, I.; Ardabili, S.; Mosavi, A. Comparative Analysis of ANN-ICA and ANN-GWO for Crop Yield Prediction. *Preprints* **2020**, 2020020353.

79. Sharabiani, V.; Kassar, F.; Gilandeh, Y.; Ardabili, S. Application of soft computing methods and spectral reflectance data for wheat growth monitoring. *Iraqi J. Agric. Sci.* **2019**, *50*, 1064–1076.

80. Gundoshmian, T.M.; Ardabili, S.; Mosavi, A.; Varkonyi-Koczy, A.R. Prediction of combine harvester performance using hybrid machine learning modeling and response surface methodology. In *Engineering for Sustainable Future, Lecture Notes in Networks and Systems*; Springer: Cham, Switzerland, 2019.

81. Nosratabadi, S.; Mosavi, A.; Keivani, R.; Ardabili, S.; Aram, F. State of the art survey of deep learning and machine learning models for smart cities and urban sustainability. In Proceedings of the International Conference on Global Research and Education, Balatonfüred, Hungary, 4–7 September 2019; pp. 228–238.

82. Atashpaz-Gargari, E.; Lucas, C. Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition. In Proceedings of the 2007 IEEE Congress on Evolutionary Computation, Singapore, 25–28 September 2007; pp. 4661–4667.

83. Lei, D.; Yuan, Y.; Cai, J.; Bai, D. An imperialist competitive algorithm with memory for distributed unrelated parallel machines scheduling. *Int. J. Prod. Res.* **2020**, *58*, 597–614. [CrossRef]

84. Khabbazi, A.; Atashpaz-Gargari, E.; Lucas, C. Imperialist competitive algorithm for minimum bit error rate beamforming. *Int. J. Bio-Inspired Comput.* **2009**, *1*, 125–133. [CrossRef]

85. Ahmadi, M.A.; Ebadi, M.; Shokrollahi, A.; Majidi, S.M.J. Evolving artificial neural network and imperialist competitive algorithm for prediction oil flow rate of the reservoir. *Appl. Soft Comput.* **2013**, *13*, 1085–1098. [CrossRef]

86. Chang, F.J.; Chang, Y.T. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Adv. Water Resour.* **2006**, *29*, 1–10. [CrossRef]

87. Polat, K.; Güneş, S. An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. *Digit. Signal Process.* **2007**, *17*, 702–710. [CrossRef]

88. Yilmaz, I.; Yuksek, G. Prediction of the strength and elasticity modulus of gypsum using multiple regression, ANN, and ANFIS models. *Int. J. Rock Mech. Min. Sci.* **2009**, *46*, 803–810. [CrossRef]

89. Wali, W.; Al-Shamma'a, A.; Hassan, K.H.; Cullen, J. Online genetic-ANFIS temperature control for advanced microwave biodiesel reactor. *J. Process Control* **2012**, *22*, 1256–1272. [CrossRef]

90. Mostafaei, M.; Javadikia, H.; Naderloo, L. Modeling the effects of ultrasound power and reactor dimension on the biodiesel production yield: Comparison of prediction abilities between response surface methodology (RSM) and adaptive neuro-fuzzy inference system (ANFIS). *Energy* **2016**, *115*, 626–636. [CrossRef]

91. Keshavarzi, A.; Sarmadian, F.; Shiri, J.; Iqbal, M.; Tirado-Corbalá, R.; Omran, E.S.E. Application of ANFIS-based subtractive clustering algorithm in soil Cation Exchange Capacity estimation using soil and remotely sensed data. *Measurement* **2017**, *95*, 173–180. [CrossRef]

92. Yadav, D.; Chhabra, D.; Gupta, R.K.; Phogat, A.; Ahlawat, A. Modeling and analysis of significant process parameters of FDM 3D printer using ANFIS. *Mater. Today Proc.* **2020**, *21*, 1592–1604. [CrossRef]

93. Naderloo, L.; Alimardani, R.; Omid, M.; Sarmadian, F.; Javadikia, P.; Torabi, M.Y.; Alimardani, F. Application of ANFIS to predict crop yield based on different energy inputs. *Measurement* **2012**, *45*, 1406–1413. [CrossRef]

94. Najah, A.A.; El-Shafie, A.; Karim, O.A.; Jaafar, O. Water quality prediction model utilizing integrated wavelet-ANFIS model with cross-validation. *Neural Comput. Appl.* **2012**, *21*, 833–841. [CrossRef]

95. Mohandes, M.; Rehman, S.; Rahman, S. Estimation of wind speed profile using adaptive neuro-fuzzy inference system (ANFIS). *Appl. Energy* **2011**, *88*, 4024–4032. [CrossRef]

96. Ekici, B.B.; Aksoy, U.T. Prediction of building energy needs in early stage of design by using ANFIS. *Expert Syst. Appl.* **2011**, *38*, 5352–5358. [CrossRef]

97. Ardabili, S.F.; Mahmoudi, A.; Gundoshmian, T.M. Modeling and simulation controlling system of HVAC using fuzzy and predictive (radial basis function, RBF) controllers. *J. Build. Eng.* **2016**, *6*, 301–308. [CrossRef]

98. Ardabili, S.; Mosavi, A.; Mahmoudi, A.; Gundoshmian, T.M.; Nosratabadi, S.; Varkonyi-Koczy, A.R. *Modelling Temperature Variation of Mushroom Growing Hall Using Artificial Neural Networks. Engineering for Sustainable Future, Lecture Notes in Networks and Systems*; Springer: Cham, Switzerland, 2019.