International Journal of Computer Applications (0975 – 8887)
Volume 184 – No.2, March 2022

COVID-19 Outbreak Prediction using Artificial Neural Network: A Review

Manu Tyagi
Department of CSE
MIT Moradabad, India

Peeyush Kumar
Department of CSE
MIT Moradabad, India

Mayank Agarwal
Department of CSE
MIT Moradabad, India

Unnati Gangwar
Department of CSE
MIT Moradabad, India

Jayati Bhardwaj
Assistant Professor, CSE
MIT Moradabad, India

Priyanshu Murari
Department of CSE
MIT Moradabad, India

ABSTRACT

The global spread of the COVID-19 outbreak has led to studies on a variety of topics, including predictions of predictable cases. Because it helps to identify the need to deal with epidemic situations. We used artificial neural networks (ANNs) in this study to predict the number of COVID-19 cases in Brazil, Mexico, India, and Italy in the coming days. The Prey Predator Method (PPA) is a type of meta-heuristic algorithm for training guessing models. The root function of the mean squared error (RMSE) and the correlation coefficient were used to evaluate the performance of the target ANN models (R). ANN models performed much better than other models in Brazil, Mexico, India, and Italy in terms of disease rates (active cases), recovery, and death. The simulation results for the ANN models predict the values accurately. Traditional monolithic neural networks have a much higher percentage of predictor errors than meta-heuristic algorithms. The report shows an estimated daily morbidity, recovery and mortality in Brazil, Mexico, India and Italy in early 2021.

Keywords

COVID 19, ANN, SIR, SEIR, PSO, MLPNN, Prey predator algorithm (PPA)

1. INTRODUCTION

The COVID-19 pandemic started in China and spread rapidly around the world, forcing the World Health Organization (WHO) to declare a global public health emergency. After Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome, COVID-19 is the third most prevalent corona virus respiratory disease (MERS) in the past 20 years. Since its first appearance in Wuhan, China, COVID-19 has claimed thousands of lives and severely damaged health systems around the world. In the fight against the COVID-19. The first case was reported in India on 27, 2020.

July 5, 2020, the number of COVID-19 infected people surged drastically, and by July 30, the virus had swept throughout India. On February The first case was reported in India on January 27. The largest number of cases was 1397 as of March 30, 2020. However, beginning on 28 April, 2020, the first cases of COVID-19 infection in Mexico were reported.

The behaviour of COVID-19 is evident in a number of scientists from a number of fields, including mathematics, chemistry, computer science, and health care, with a view to finding a solution. To date, however, all efforts have been in vain. In addition, the symptoms of COVID-19 patients have been changing over time. Early symptoms include coughing, fever, and depression, followed by pneumonia and severe respiratory illness.

Data is a very important requirement for a framework for the intelligence community to address this issue. It is difficult to construct and diagnose sufficient models due to lack of data for patients of COVID-19. Because people are reluctant to go to the doctor to avoid any possible human confinement, limited information about patients tested for COVID-19 is accessible. In addition, some infected patients do not show any symptoms of COVID-19, making it difficult to determine the exact number of positive cases in most people. The longer it takes to diagnose a person with the virus, the more dangerous it is for others. As a result, the availability of large amounts of data is important in analyzing COVID-19 disease and predicting behavior. Several published publications that look at the effects of COVID-19 from a mathematical, mathematical, and statistical perspective. Susceptible Infectious Recovered (SIR) is a mathematical model widely used to study the potency of COVID19 virus. It uses time-dependent systems to measure statistics to measure the growth of epidemics. Different types of SIR have been used to investigate and test Ebola and AIDS. In this regard, a standard SEIR epidemiologic model was developed to measure the prevalence of SARS CoV2 in Italy.

Particle Swarm Optimization (PSO) is used as a solution for the appropriate boundary model, providing a possible approach. Berger et al. investigates the impact of testing and detention regulations in the United States using the same SEIR model. The authors have shown that increasing the number of tests performed and selective segregation can reduce the economic impact of corona infection and reduce the burden on hospitals. With the use of a certified epidemic, it is important to diagnose and diagnose those affected as soon as possible, as the virus can spread if it is not detected and undiagnosed. Cases as a basis, Roda et al. It has been shown that the SIR model exceeds the SEIR model. The Akaike information criterion is used to select a model. In another study by Weissman et al. proposed. Using the SIR model of the COVID-19 Hospital Impact Model for Epidemics (CHIME), the researchers were able to predict when the hospital volume would run out and the ventilators would be extended in number, focusing on three hospitals in the Philadelphia area. CHIME predicted when available resources would be available to deal with COVID-19 patient surgery.

The actual number of COVID-19 data represents a timed view set. Machine-based methods, Meta-predictors, and structural-based methods are examples of time-series predictors that are natural in the field of mathematics. By predicting time series, artificial neural networks (ANNs) are often used. One of the important advantages of ANN methods is that they can be fed raw data and determine the required representation automatically. ANNs produce reliable results based on various parameters such as performance, accuracy, delay, speed, compression, and size. Brazil and Mexico have the highest number of diseases in the world, with millions estimated. In addition, it is estimated that the number of infected people in Brazil and Mexico exceeds the reported prevalence due to the spread of the disease in poorer communities. This study uses ANNs to predict the problem of a time series related to the COVID-19 study in Brazil and Mexico. In addition, by setting the maximum number of ANN parameters, we used the prey predator algorithm (PPA) to improve the performance of the ANN model. There is a lot of analysis of time series based on limited time visuals and time series variables. Below is how the whole article is organized. Section 2 consists of the literature review of the different approaches used for covid outbreak prediction. Section 3 defines the proposed approach for the outbreak prediction which is further followed by the conclusion of the work in Section 4.

2. LITERATURE REVIEW

ZY ZU, MD Jiang et.al December 2019. Outbreaks of corona-acute respiratory infections 2 occurred in Wuhan, China and specialized throughout China and beyond on February 12, 2020, the World Health Organization officially named the novel Corona virus. as COVID-19. [1] [4]

Tartaglione et.al opportunity to use a wide and simple chest x-ray image for early detection. COVID-19 patients are attracting great interest from both the clinician and the AI community. In this study they provide information and also suggest warnings about what to expect by using in-depth study on COVID image classification of CXR imagery. Provides practical and comprehensive study of a comprehensive set of mathematical results that can be obtained using currently available databases. [2] [6]

Kong, W .; Agarwal et.al presented a series of three certified cases of COVID-19 infection from China from Wuhan, China. Symptoms at presentation included paroxysmal cough, productive sputum, fever. Chest CT shows abnormalities of the lower extremity of the small nodular vertebrae in the upper left hemisphere, developed in 3 days to multifocal nodular and peripheral ground-glass opacities involving both upper extremities. A CT scan of 5 days from the introduction showed a small hole. [3] [5]

Chen, Y .; Cheng, J.et.al parameters involved in the dynamic novel system are based on time, changes in the effect of isolation and the degree of spread can be considered time dependent in order to improve the agreement between real data and limited resolution. Apart from the complex network and stochastic process it will be used to provide a better predictor of COVID-19 outbreaks. [7]

Khaleque et.al patients living in Wuhan, China, older patients and those with health problems often have more severe clinical symptoms and higher CFR. [8] [11]

Cohen et.al collect pre and lateral visual cues and metadata such as time from early symptoms, intensive care unit (ICU), and survival status and hospital status. They present many cases of possible data use such as predicting the need for an ICU, predicting patient survival and understanding patients duration during treatment. [10]

Zakary et.al is based in two countries, researching available data such as case. The disease appears to affect a small portion of Guinea’s population and a large number in Liberia. Considering the time when the data reaches the peak, one can conclude that the disease has been present in Sierra Leone and Guinea. [12]

Fu, L.et.al using clinical features, using Particle Swarm Optimization (PSO), a solution to improve predictive reliability in the medium term. They analyzed official data and predictable evolution in an Italian epidemic and compared the results with data and predictions. [14] [15].

3. PROPOSED APPROACH

3.1 Structure

ANN was instrumental in solving a myriad of current problems. ANN is made up of three layers: one input layer, one or more hidden layers, and one or more output layers. In ANN, processing elements, or neurons, produce outputs based on their predetermined functions.

The Multilayer Perceptron Neural Network (MLPNN) is the ANN feed we are currently using. We constructed MLPNN with 1 layer and 10 neurons using the sigmoid activation function on hidden neurons (see figure 1). It is noteworthy that connecting weights connect neurons in one layer with neurons in another layer. In this study, the input weights connect the input layer to the hidden layer and the output weights connect the hidden layer. Also, in the output neurons, we have adopted hyperbolic tangential transfer function as shown in figure (2). It is important to note the figure (2) gives the result between 1 and 1.

$y_i=1/(1+e^{(-\sum w_{ki}*x_k)})$ .... (1)

where x, k is the input value in the input neuron i, w, k, the node that binds the node k from the input layer with n.

$y_j=2/((1+e^{(_2\sum w_{(ij.)} y_i)})-1)$ .... (2)

where w ∗ i j is the output current between the hidden neuron and the output neuron j. yˆ j outgoing number of outgoing neurons j.

Lifting the ANN model involves two steps.

MLPNN properties must first be defined, depending n the data to be presented. The MLPNN formats of the two models we used are the same , depending on the data sets used in the study (Brazilian data set and Mexico data set) (ANN Brazil model and Mexican model). In both models, we used one input neuron, 10 hidden neurons, and 3 outgoing neurons. As a result, two ANN models were proposed, one for Brazil and one for Mexico.

Number of disease, recovery, and death are three outcomes in each model weights of the Brazilian MLPNN (BMLPNN) model do not match the appropriate parameters of the Mexican MLPNN model.

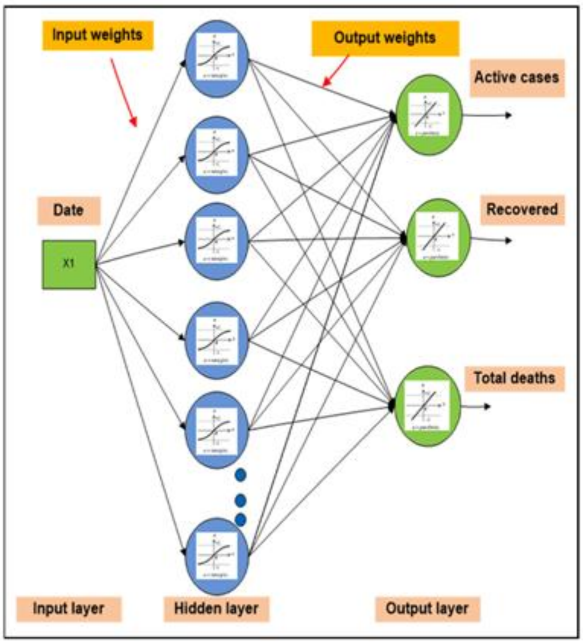

Fig 1. The Structure of the MLPNN that have been used for models (Brazil, India, Mexico and Italy ANN model).

The second phase is the ANNs' training process, which compares real and expected results. As a result, the goal of training is to achieve the smallest possible difference between the physical output of a data set (training data) and the corresponding values of the ANN model, ANN must first go through the training step.

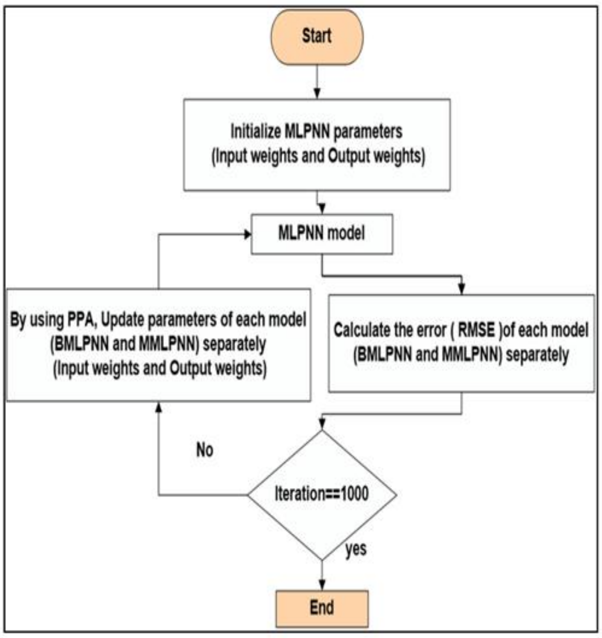

Fig 2. Training process of using prey predator algorithm (PPA) for both MLPNN models (PPA-BMLPNN and PPA-MMLPNN).

The PPA is used to study the difference between the two, and weights are adjusted to reduce the difference (RMSE). It is worth noting that the MLPNN boundary are determined as follows:

(1)Appropriate settings are determined using the development method. The PPA was used in this study to determine the best settings (Input weights and output weights).

(2)A set of data used in the training process (training data). This study used two different data. sets: A Brazilian data set and a Mexican data set. According to two different data sets (MMPLNN), the appropriate parameters (input and output) weights) for the Brazilian MLPNN model (BMLPNN) do not match the optimal parameters for the Mexican MLPNN model.

The PPMLPNN training process is illustrated in Figure 2. Performance measurement and performance improvement is achieved using the root mean square (RMSE) and correlation coefficient (R). Equations (3) and (4) can be used to calculate RMSE and R.

$$RMSE = \sqrt{\frac{(O-E)^2}{n}} \text{ ....(3)}$$

$$R = \frac{n\sum OE - \sum O \sum E}{\sqrt{[n\sum O^2 - (\sum O)^2][n\sum E^2 - (\sum E)^2]}} \text{ …………..................................(4)}$$

where n is the number of occurrences, ∑O is the sum of all observed events, ∑E is the sum of all expected values, $\sum O^2$ is the sum of all the square marks, and $(\sum O)^2$ is the square of the sum. In all, $\sum E^2$ is the square root of all expected values, and $(\sum E)^2$ is the sum of all expected values.

3.2 Prey Predator Algorithm (PPA)

The PPA is considered one of the most effective ways to determine the best solution to a problem. Continuous preparation, integrated development, and effective implementation of challenges are among the issues we are addressing. It can also deal naturally and effectively with more indirect and multifaceted problems. MLPNN is a problem for performance, and we used PPA to find the best MLPNN models by finding the right weight for the models. As mentioned earlier, the weight values are adjusted to reduce the RMSE.

In predatory animal relationships, animal interactions promote the development of PPA. An important concept of the PPA is that the attacker is looking for his prey while the deer is trying to escape and hide. Animals and predators are two types of solutions found in the PPA search engine.

In this paper, PPA is used to improve the survival rate of predators through the use of the RMSE function to obtain a minimum of performance. The best prey among the target species is the one that is able to escape and find a safe haven. This type of predator is known as the best victim in the search area, ad has the highest survival rate. In each case, the invader hunts the weaker prey, trying to escape from the invader by following other predators. Find a safe place to hide. The hunter’s movement to the hunt may be determined by the direction the attacker or the victim is moving and the distance he will cover in that area. Figure (5) shows the direction of the solution in the search field:

$$y_i = \sum e^{sv(x_j)^v - \|x_i - x_j\|}(x_i - x_j) \text{……………..……………………..(5)}$$

When the different values of the parameter v determine the jump size for the solution xi.

If going a certain way increases the survival value of the solution, that approach is called correct. The step length while researching the step length of the small and large step, where min-max, is difficult during the solution review. Equations (6)–(9) can be used to represent the invader and the movement of predators, as discussed in Deer movement is given in Equation (6) after tracking access is achieved.

$$x_i(t+1) = x_i(t) + \left(\frac{y_i}{\|y_i\|}\right) \mathrm{rand}(\lambda_{max} + \lambda_{min}) \ldots\ldots\ldots\ldots\ldots (6)$$

Alternatively, the movement of prey is determined by Equation (7).

$$x_i(t+1) = x_i(t) + \left(\frac{y_i}{\|y_i\|}\right) \mathrm{rand}(\lambda_{max} + \lambda_{min}) \ldots\ldots\ldots\ldots\ldots (7)$$

The best prey and predator's movements are respectively given by Equations (8) and (9).

$$x_i(t+1) = x_i(t) + \left(\frac{y_i}{\|y_i\|}\right) \mathrm{rand} * \lambda_{min} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

$$x_{predator}(t+1) = x_{predator}(t) + \left(\frac{y_i}{\|y_i\|}\right) \mathrm{rand} * \lambda_{max} * \lambda_{min} \left(\frac{x'_i - x_{predator}(t)}{\|x'_i - x_{predator}(t)\|}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

4. CONCLUSION

Epidemiological analytical models have shown sufficient accuracy to predict long-term epidemic due to a high levels of uncertainty and a lack of vital data. For the first time, ANN algorithms are used to describe the escalation of the COVID-19 epidemic, both in terms of infection and death, in excess of continuous fluctuations, to identify high rate of change, and to predict outbreaks. The algorithm can be adapted to different countries by different populations, climatic conditions. With strong communication coefficients (0.99) in respect of reported cases in all tested countries, the ANN-created algorithm produced COVID-19 outbreaks that are very close to the actual report dates, fulfilling fulfilment of your future predictions of both infection and death. COVID-19 cases of death in Qatar, Spain, and Italy were accurately predicted using the ANN architecture based on feed-forward back-propagation distribution algorithms.

With the exception of low data rates, ANN predictions were sufficient. The results showed ANN's high capacity to predict long-term outbreak of the COVID-19 epidemic. Due to the very complex nature of the COVID-19 epidemic and changes in national and national governance and defense, outbreak trends are evolving in each country. Current work serves as the starting point for future research on ANN potential.

5. REFERENCES

[1] ZU, Z.Y., Jiang, M.D., XU, P.P., Chen, W., Ni, Q.Q., Lu, G.M., and Zhang, L.J. ZU, Z.Y., Jiang, M.D., XU, P.P., Chen, W., Ni, Q.Q., Lu, G.M., and Zhang, L.J. Chinese view on Corona virus Disease 2019 (COVID-19). 200490, Radiology 2020.

[2] Tartaglione, E.; Barbano, C.A.; Berzovini, C.; Calandri, M.; Grangetto, M. Reveals COVID-19 from chest X-ray for in-depth reading: Obstacle racing with little data. Int. J. Nature. Res. Public Health 2020, 17, 6933.

[3] Kong, W.; Agarwal, P.P. Appearance image of chest infection with COVID-19. Radiol. Cardiothorac. Illustration 2020, 2, e200028.

[4] Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; et al An outbreak of pneumonia associated with a new corona virus may be the origin of bats. Nature 2020, 579, 270–273.

[5] Li, X.; Geng, M.; Peng, Y.; Meng, L.; Lu, S. Molecular immune pathogenesis and diagnosis of COVID-19. J. Pharm. Anal. 2020, 10, 102–108.

[6] Wang, L.; Lin, Z.Q.; Wong, A. Covid-net: Design of a deep neural network designed to detect COVID-19 cases in chest x-ray images. Science Rep. 2020, 10, 19549.

[7] Chen, Y.; Cheng, J.; Jiang, Y.; Liu, K. A flexible time-delayed system with an external source for the 2019-nCoV outbreak. Application. Indunu. 2020, 1–12.

[8] Fu, L .; Wang, B .; Yuan, T .; Chen, X .; AO, Y .; Fitzpatrick,T .; Li, P .; Zhou, Y .; Duan, Q .; et al. Clinical features of corona virus 2019 (COVID-19) in China: A systematic review and meta-analysis. J. Pour. 2020, 80, 656-665.

[9] Cascella, M.; Rajnik, M.; Cuomo, A.; Dulebohn, S.C.; Di Napoli, R. Features, Diagnosis and Treatment of Corona virus (COVID-19). Statpearls; StatPearls Publishing: Treasure Island, FL, USA, 2020.

[10] Cohen, J.P.; Morrison, P.; Dao, L.; Roth, K.; Duong, T.Q.; Ghassemi, M. COVID-19 image data collection: Predictable predictions are future. arXiv 2019, arXiv:200611988.

[11] Gao, Z.; Xu, Y.; Sun, C.; Wang, X.; Guo, Y.; Qiu, S.; Ma, K.. Systematic review of asymptomatic infection with COVID-19.J. Microbiol. Immunol. Pour 2020.

[12] Khaleque, A.; Sen, P. A powerful analysis of the Ebola outbreak in West Africa. Science. Rep. 2017, 7, 42594.

[13] Zakary, O.; Larrache, A.; Rachik, M.; Elmouki, Outcome of awareness programs and prevention activities in controlling the spread of HIV/AIDS: A multidisciplinary SIR model. Adv. Differ. Equ. 2016, 2016, 169.

[14] Godio, A.; Pace, F.; Vergnano, A. SEIR SARS Italian Epidemic Model- CoV-2 Using Computer Intelligence. Int. J. Nature. Res. Public Health 2020, 17, 3535.

[15] Berger, D.W.; Herkenhoff, K.F.; Mongey, S. An infectious disease model with diagnostic and conditional classification. Natl. Bur. Econ.