Volume 11, Issue 3
June 2022

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

A Survey on COVID-19 Severity Prediction using Various Models

[1]Akanshika Gandhi, [2]Vedpal and [3]Harish Kumar

[1]Department of Computer Engineering, J.C. Bose University of Science and Technology, India

[2]Department of Computer Applications, J.C. Bose University of Science and Technology, India

[3]Department of Computer Engineering, J.C. Bose University of Science and Technology, India

E-mail: [1]akanshika.gandhi@gmail.com, [2]ved_ymca@yahoo.co.in, [3]htanwar@gmail.com

ABSTRACT

Making judgments about how to tackle the situation of COVID-19 requires a lot of data and information. Data on COVID-19, both demographics and aggregates, are processed and evaluated to give details about the circumstances and state of the COVID-19 epidemic in the world at the moment. The COVID-19 data is additionally used in predictive analysis to determine the expected count of COVID-19 cases in the near future. However, the rapid spread of COVID-19 requires the right approach to predict its spread in the process of fighting the virus. This demonstrates how and to what extent Deep Learning (DL) is important in developing and improving global health care systems. The limited amount of data is the main problem for Deep Learning (DL) to be able to generate predictions in COVID-19 time series data. In this article, various models are explored for the nitty-gritty of obtaining predictions long-term, namely by making the prediction results one step ahead of the new data.

Keywords: *Covid-19; deep learning; neural networks; prediction*

1. INTRODUCTION

The WHO and the Chinese government announced the emergence of respiratory illness in the city of Wuhan at the end of December 2019, with the seafood market as a potential source. By this time, there were already critical cases and others almost recovered. On January 9th, 2020, it was confirmed that the disease was a type of not-so-lethal coronavirus and had already arrived outside the city of Wuhan. Just four days later, it was known that the virus had affected people in other countries who had made recent trips to the city of Wuhan. Subsequently, on January 17, 2020, many governments in various major cities around the world began to carry out controls at airports. Three days later, the confirmation is announced that the virus can be transmitted from human to human and not just from animal to human-like initially thought, it was said that the virus could mutate and facilitate its transmission [1]. It is not until January 21st, 2020 that the virus manages to leave the Asian continent, reaching more specifically the United States by an individual who had made a trip to the city from Wuhan. By January 27, 2020, the disease had already managed to spread to different countries on various continents, such as Canada, South Korea, France, Vietnam, and Australia, infecting a total of 2846 people so far and leaving 81 fatalities as a result. Subsequently, severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) or COVID-19 is an infectious disease caused by the coronavirus family which can cause common cold-like illnesses [2]. The symptoms of this disease can manifest between the second and fourteenth days after exposure to the virus. To detect it, a C-reactive protein test (CRP) is done. If it gives high results, it indicates the presence of severe infection [3]. The infection rates of this disease are high. Therefore, one of the measures considered is the separation of patients who arrive at the emergency room due to COVID-19 infection. Thus, exclusive ICUs are available for this type of patient. The high demand for services in health care has caused most hospitals to be overwhelmed due to the lack of resources to provide care (beds, staff, supplies, among others) [4]. Among the numerous sanitary measures taken to mitigate the virus are mobility restrictions for the population, such as sanitary customs, sanitary cords, sanitary passports, temporary permits for displacement, and confinement by communes. The latter is carried out based on the following criteria:

1. The total number of active cases.
2. A rise in the count of active cases.
3. The integrated COVID-19 network's capacity for assistance.

2. LITERATURE SURVEY

In the fields of mathematics and epidemiology, there are models of equation differentials that seek to emulate the behavior of an epidemic or pandemic to have approximations of how it could affect the study population during a period of given time. The models chosen for this chapter are SIR [5] (susceptible-infected-recovered) and SEIR (susceptible-exposed-infectious-recovered) [6]. These models present some variants, which would be SIR and SEIRS, which consider the possibility that people are susceptible to catching the disease again after a while. These models in which there may be a re-infection will not be taken into account because the evidence in the world is not long enough to make such a claim, and some researchers suggest that this is because the tests are not 100% accurate, there are errors in their taking, or that simply the virus does give immunity, but it is not very strong, and that is why there are some patients who have tested positive after having tested negative, as investigated in [5, 6].

- **Forecast Models**

Researchers [6] suggested a time series forecast model that can simulate the most common commercial time series characteristics, such as trend by part, multiple seasonality, impact on the series on non-periodic relevant dates, and the presence of outliers. This is why a decomposed time series

Volume 11, Issue 3
June 2022

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

model is proposed (Harvey and Peters 1990) with three main components: trend, seasonality, and events. These combined have the following equation:

$$y(t) = g(t) + s(t) + h(t) + \in_{t,}$$ (Eq.no.1)

where $g(t)$ trend function that models the non-periodic changes in the value of the series over time, $s(t)$represents periodic changes, $h(t)$ represents the irregular effect produced by an event over one or more days and $y \in_{t,}$ represents any idiosyncratic change that the model cannot account for.

- **Trend models**

Two models are proposed to model the trend: the saturated growth model and the piecewise linear trend.

- *Nonlinear saturated growth*

To represent growth, a variation of the logistic growth model is used, since it is considered that the carrying capacity and the growth rate are not constant. The most basic logistic growth model is:

$$g(t) = \frac{C}{1+\exp\left(-k(t-m)\right)'}$$ (Eq.no. 2)

with C carrying capacity, k growth rate, and m displacement parameter.

Given the limitations of the above equation, which replaces the load capacity *C* with a capacity that depends on the time *C (t),* in addition to stating explicitly the points at which the growth rate is permitted to change. Assume that there exist S points of change s_j, $j = 1...$ *S.* Additionally, we define the adjustment vector $\delta \in R^s$, where δ_j is the rate of change that occurs at the moment s_j. The rate at any time t is then equal to the rate at base k plus all previous adjustments: $K + \sum_{j:t>s_j} \delta_{j.}$ The vector is then defined as $a(t) \in \{0,1\}^S$ such that:-

$$a(t) = \begin{cases} 1, Yes \;\; t \geq s_j \\ 0, on\ other\ case. \end{cases}$$ (Eq.no.3)

At time *t*, the rate is then $k + a(t)^T\delta$. When rate k is adjusted, the parameter of offset m must also be modified to connect the segment ends. The correct fit at switch point *j* can be easily calculated as:-

$$\gamma_i = \left(S_j - m - \sum_{l<j} \gamma l\right)\left(1 - \frac{k+\sum_{i<j}\delta_l}{k+\sum_{i\leq j}\delta_l}\right).$$ (Eq.no.4)

- *The linear trend with change points*

For problems that do not exhibit saturated growth, a growth rate that is constant and piecemeal provides an often useful model. That model is given by:-

$$g(t) = (k + a(t)^T\delta)t + (m + a(t)^T\gamma),$$ (Eq.no.5)

k represents the growth rate, δ has the rate settings, *m* represents the displacement parameter, and y $\gamma_j$is set as $-s_j\delta_j$to obtain a continuous function.

Then the logistic growth model by parts is given by:-

$$g(t) = \frac{C(t)}{1+\exp\left(-(k+a(t)^T\delta)\left(t-\left(m+a(t)^T\gamma\right)\right)\right)}$$ (Eq.no.6)

- *Automatic selection of change points*

The change point's s_j can be specified by the analyst using dates known as product launches and other changes in growth caused by some γ event. Or, the automatic selection of change points can be formulated through (Eq. no.5) and (Eq. no.6) using a small priori for δ. Because a large number of switch points is frequently specified, so $\delta_j \sim$ Laplace $(0, \tau)$ is used. The model's flexibility is controlled by the parameter τ.

- **Seasonality**

Time series often exhibit multi-period seasonality as an outcome of the human behaviors they portray. So you must specify models of seasonality that are periodic functions of *t.*

Let *p* be the expected regular period of the time series. You can approximate the seasonal smoothing effects with the standard Fourier series.

$$s(t) = \sum_{n=1}^{N}(a_n \cos\left(\frac{2\pi nt}{p}\right) + \; b_n \sin\left(\frac{2\pi nt}{p}\right)).$$ (Eq.no.7)

Adjusting for seasonality requires 2N parameters $\beta = [a_1, b_1, \ldots, a_N, b_N]^T$. We can construct a matrix of seasonality vectors *(X (t))* for every value of *t* in our historical and subsequent data. Then the seasonal component is:-

$$s(t) = X(t)\,\beta$$ (Eq.no.8)

For the generic model, we take $\beta \sim N(0, \sigma_\beta^2$ to enforce anti-aliasing before seasonal.

- **Events**

Events provide high volatility and often do not follow a periodic pattern, so their effect is not well modeled by a smooth cycle. For each event *i*, let D_i be the past and future data set for each event. Additionally, an indicator function is included that shows if time *t* is occurring during event *i*.

$$Z(t) = [1(t \in D_i) \ldots, 1\left((t \in D_L)\right].$$ (Eq.no.9)

Furthermore, the parameter k is assigned to each event, which denotes the change in corresponding toward the forecast. Then it remains in the form:-

$$h(t) = \; Z(t)k.$$ (Eq.no.10)

As with the seasonal component, the priori $k \sim N(0, \sigma_\beta^2)$

- **Recurrent neural networks**
- *Neural networks*

One of the first works that reported an idea regarding neural networks was that of McCulloch and Pitts (1943), as an attempt to mathematically model the behavior of a biological neuron. Two branches of research were born from this. For On the other hand, it continued to delve into the functioning of biological neurons, while from that, on the other hand, was born an approach applied to models in the area of artificial intelligence [7].

Researchers [8] consider neural networks as universal approximating functions, being able to find and reproduce complicated relationships without entering direct knowledge between input and output variables if there is enough training data.

Three layers constitute the most conventional neural network architecture: an input layer, a hidden layer, and a further output layer. In this type of network, the input layer is completely independent of the output layer. Additionally, one or more

Volume 11, Issue 3
June 2022

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

hidden layers are linked to the input and output layers via weight matrices, bias, and different activation functions.

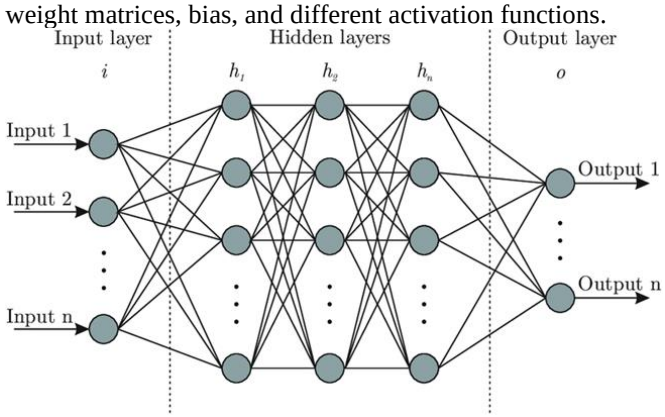

Fig. 1: Scheme of architecture on a simple neural network

The key difference between various types of neural networks involves the architecture of their networks, which is correlated to the multitude of hidden layers in addition to the multitude of neurons in each hidden layer.

Neural networks present advantages in predictions due to their non-linear adaptive power. But when talking about time series, traditional neural networks have shortcomings because they are not capable of capturing the temporal dependence that observations may have. That is why recurrent neural networks (RNN) were born, since they use previous events to be able to infer subsequent events.

In the 1980s, RNN was first developed [9]. An input layer, one or more hidden layers, in addition to an output layer, make up its structure. RNNs have repeating modules with a structure akin to a string, with the intention of using these modules to store crucial data from earlier processing steps. As a result, RNNs have been effective at learning sequences.

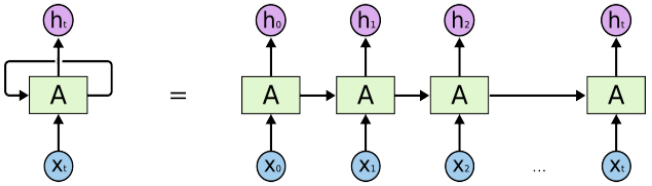

Fig. 2: Sequential processing of a recurrent neural network.

Figure 2 shows a simple RNN with an input unit, a unit output, and a recurring hidden unit spread over the entire network, where X_t is the input at a time t and h_t is perhaps the output at a time t.

RNNs use a back-propagation algorithm during training, which is frequently implemented in computing the gradient and adjusting the matrices that need to be adjusted after modifying the feedback process. This is why this algorithm is also commonly known as back propagation through time (BPTT). The BPTT attempts to minimize the error between actual and expected outputs in response to inputs by modifying the weights of a neural network. Consequently, if the training data is extensive, the network residuals that must be returned will drop rapidly, which will slow the network's update of its weights. That is, because the RNN has a gradient problem, it cannot accurately represent the effect of long-term memory in real-world applications; thus, an additional component is required to store this memory. Therefore, an RNN was proposed, namely the long short-term memory model, which features both short and long memory terms [10, 11].

- **Long short-term memory**

LSTM, or long short-term memory, is a revolution in RNN. It was presented by [12] to optimize RNN by adding further interactions per cell, learning long-term dependencies, and being able to retain information for a longer period of time.

According to [13], the LSTM model is arranged as a chain structure. The recurring module, however, is distinct. The LSTM includes four interactive layers with a different communication strategy than the regular RNN. The network structure that will be employed for this endeavor, as inspired by [13], is specifically as follows:

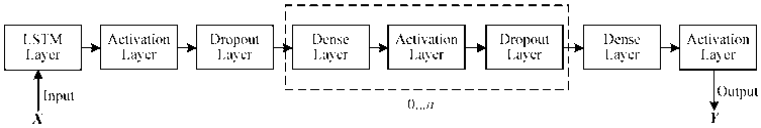

Fig. 3: Structure of a hybrid LSTM neural network.

In Figure 3, the hybrid LSTM neural network's structure has been depicted. It comprises an input combination layer, one output combination layer, and zero to more intermediate combination layers. The input combination layer consists of an LSTM layer, an Activation layer, plus a Dropout layer. Then, a Dense layer, an Activation layer, plus a Dropout layer make up the intermediate combination layer. The output combination layer also has an Activation layer and a Dense layer.

- *LSTM layer*

The LSTM layer is made up of several LSTM units, altogether termed as the LSTM model [13]. In contrast to other neural networks, the LSTM model consists of three multiplicative units, i.e., an input gate, an output gate, and a forget gate. In summary, the input gate is in charge of memorizing the current information, the output gate serves as the output, and the forget gate is utilized to selectively neglect certain past knowledge.

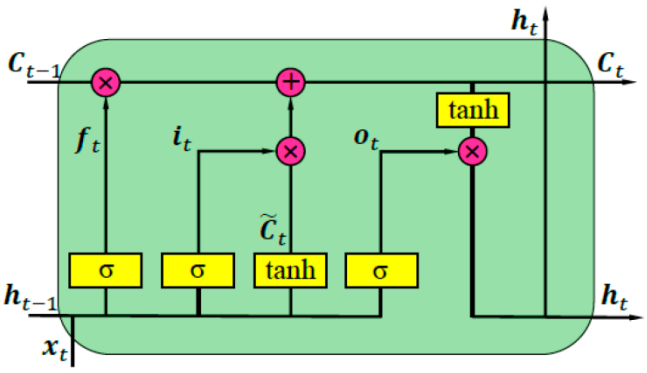

Fig. 4: LSTM unit with its respective components

In Figure 4, f_t is the forget gate, i_t represents the input gate and o_t is the output gate. The input and output gates collectively govern the cellular state C_t that holds past and forthcoming information. The LSTM unit's input and output information are denoted by x_t and h_t, respectively.

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \text{ (Eq.no.11)}$$

where σ is the sigmoidal function, W_{xf} and b_f are the weight and bias matrices, respectively, of the forget gate process.

Volume 11, Issue 3
June 2022

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad \text{(Eq.no.12)}$$

$$o_t = tanh(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad \text{(Eq.no.13)}$$

$$C_t = f_t * C_{t-1} + i_t * g_t \quad \text{(Eq.no.14)}$$

$$g_t = \sigma(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad \text{(Eq.no.15)}$$

$$h_t = o_t * tanh(C_t) \quad \text{(Eq.no.16)}$$

The forget gate expression in Eq. no. 11 uses the weighted sum of the cell's state at time t, output at time t-1, and input at time t as the input of the activation function to produce the forget gate's output. In most cases, the activation function is typically a logistic sigmoid function. The parameters and operating principles of Eq. No. 12's input gate expression are identical to those of Eq. No. 11. Eq. no. 15 determines the potential value of the memory unit, where the activation function may be either a logistic sigmoid function or a tanh function. Eq.no. 14 combines memories from the past and the present. The output gate expression in Eq.no. 13 has the parameters and principles identical to those of Eq. no. 11. In Eq. no. 16, the cell output is expressed, and the activation function is typically a tanh function.

- **Forecasting Accuracy**

Validation of forecasting methods, especially those above, cannot be separated from the indicators in measurement of forecasting accuracy. The overall accuracy of the model bias forecasting is determined by comparing its values foreseen with actual values. Forecasting errors are defined by [14]:

$$role\ play\ error = request - forecast$$

However, there are several indicators of measurement forecasting accuracy, but the most commonly used are the mean absolute deviation, mean absolute percentage error, and mean squared error.

- **Mean Absolute Deviation (MAD)**

MAD is the absolute total value of the forecast error divided by the amount of data. Or the simpler one is the absolute cumulative value, as [15]:-

$$MAD = \frac{\sum(absolute\ of\ forecast\ error)}{n}$$

- **Mean Absolute Percentage Error (MAPE)**

The mean percent error squared is a measure of accuracy using the absolute percentage error. MAPE shows the mean absolute error of the estimate as a percentage of the actual data [16].

$$MAPE = \frac{E\left|\frac{e}{X_t}\right|(100)}{n}$$

- **Mean Squared Error (MSE)**

According to [17], the mean squared error is usually referred to as the forecasting error. Forecasting errors cannot be avoided in forecasting systems. However, these forecast errors must be managed properly. The management of forecast error will be more effective if the forecasters can take appropriate action about the reasons for the occurrence of the forecast error. In the forecasting system, the use of various forecasting models gives different forecast values and degrees of forecast errors. Average squared errors reinforce the effect of large error numbers but minimize the estimated error rates to less than one unit [18].

$$MSE = \frac{\sum ei}{n}$$

- **Root Mean Square Error (RMSE)**

The calculation error (error) is a measure of how well Neural Networks are copied. This error calculation is measured by the provision of artificial neural networks on learning target data. The accuracy of a regression model can be seen from the Root Mean Square Error (RMSE). The RMSE shows how big the deviation of the estimated value is from the actual value. The fit of the model is said to be better if the RMSE approaches 0. With the equation according to [19, 20] as follows:

$$RMSE = \frac{\sqrt{\sum_{t=1}^{n}(x_t - f_t)^2}}{n}$$

With:

$\boldsymbol{x_t}$= actual value at time $\boldsymbol{t}$

$\boldsymbol{f_t}$ = estimated value (predicted result) at time $\boldsymbol{t}$

$\boldsymbol{n}$ = amount of predicted data

3. RELATED WORK

An innovative DL-based RCAL-BiLSTM model for COVID-19 diagnosis has been developed in this paper [22]. The steps taken by the described RCAL-BiLSTM model are BF-based preprocessing, feature extraction using RCAL-BiLSTM, and classification using SM. Three modules i.e. ResNet 152 based feature extraction, CAL, including the Bi-LSTM modules—are used in the RCAL-BiLSTM based feature extraction procedure. While the CAL plans to capture the discriminative class-based characteristics, the ResNet model primarily abstracts the fine-grained semantic feature map. Thirdly, the Bi-LSTM module generates structured object labels and explains the class dependencies in both directions. The feature vectors are then sorted into equivalent feature maps using the SM layer. The Chest-X-Ray dataset was employed to assess the provided RCAL-BiLSTM model, and the findings show that the model has reached its highest levels of sensitivity (93.28%), specificity (94.61%), precision (94.90%), accuracy (94.88%), F-score (93.10%), and kappa (91.40%).

In this research [23], Deep Learning techniques are used to predict the number of COVID-19 positive cases in Indian states. The increasing number of positive cases in India has prompted an exploratory data analysis. In order to implement more effective lockdown measures at the state level as opposed to a national lockdown, which could have detrimental socioeconomic effects, states are split into mild, moderate, and severe zones in accordance with the number of cases and daily rate of growth. As prediction models, long short-term memory (LSTM) cells based on recurrent neural networks (RNN) are employed. LSTM variants such as deep LSTM, convolutional LSTM and bi-directional LSTM models are chosen for testing on 32 states/union territories and based on absolute error, the most accurate model is selected. Bi-directional LSTM produces the best results in terms of

prediction errors, while convolutional LSTM produces the poorest. For all states, weekly and daily forecasts are computed. It is discovered that bi-LSTM provides maximum accuracy (error less than 3%) for short-term forecasting (1–3 days). A website designed for the common people has predictions that are open to the public. These forecasts will be beneficial for researchers, planners, and state and federal government officials in managing services and setting up medical infrastructure appropriately.

The rapid spread of the coronavirus disease leads to severe pneumonia and is anticipated to have a significant adverse impact on the healthcare system. By deploying a Genetic Deep Learning Convolutional Neural Network (GDCNN), this research [24] aims to provide a solution for detecting pneumonia caused by COVID-19 and healthy lungs using CXR. It is trained from scratch in order to extract features and distinguish them between COVID-19 and normal images. A dataset including more than 5000 CXR image samples is utilized for categorizing pneumonia, normal and other pneumonia diseases. Training a GDCNN from scratch shows that the suggested method outperforms other transfer learning methods. With a classification accuracy of 98.84%, a precision of 93%, a sensitivity of 100%, and a specificity of 97.0%, COVID-19 prediction has been successfully achieved. The method may be applied to a relatively large scale database to achieve better hierarchical categorization accuracy.

This work [25] presents a thorough evaluation of newly invented prediction models and forecasts the count of confirmed, recovered, and mortalities posed by COVID-19 in India. Autocorrelation and autoregression were utilized to increase the accuracy, along with multiple linear regression and correlation coefficients, which were applied for prediction. The predicted number of cases corresponds well with an R-squared score of 0.9992 to the actual values.

This study [26] aims to deploy machine learning models in order to forecast the disease's trend in Indonesia, estimating when things will revert to normal. It compares the productivity and reliability of Facebook's Prophet Forecasting Model and the ARIMA Forecasting Model on a dataset consisting of confirmed cases, fatalities, and recovery cases collected from the Kaggle website. The outcome demonstrates that both ARIMA and PROPHET have proved to be pretty inaccurate in predicting as time passes.

Recurrent neural networks, including long short term memory (LSTM), bidirectional LSTM, and encoder-decoder LSTM models, were used in this research work [27] to forecast COVID-19 infections in many steps over the short term. The initial (2020) and secondary (2021) waves of infections were captured, and a couple of months ahead, a prognosis was given for the Indian states having COVID-19 hotpots. The model indicates that there is a low probability that there will be another surge of infections in October and November 2021, but the authorities must remain cautious given arising variants of the virus. Nevertheless, the complexity in modelling remains due to the reliability of data and challenges in taking into account variables like population density, transportation, and social aspects like culture and lifestyle.

Sah et al., 2022; this research demonstrates a COVID-19 predictive analysis technique [28]. The Prophet, ARIMA, and stacked LSTM-GRU models were used to anticipate the count of confirmed as well as active cases in the Indian dataset. The results of the prediction were compared using state-of-the-art models like recurrent neural network (RNN), gated recurrent unit (GRU), long short-term memory (LSTM), linear regression, polynomial regression, autoregressive integrated moving average (ARIMA), and Prophet. After doing predictive research, it was discovered that the stacked LSTM-GRU model forecast had superior prediction outcomes and was more consistent than existing models. Even though the stacked model requires a vast dataset for training, it helps maximize the model's memory size and produces results with a higher degree of abstraction. The GRU, on the contrary, aids in the resolution of vanishing gradient problem. According to the study, the suggested stacked LSTM and GRU model surpasses all the other models when it comes to R squared and RMSE. Also, the coupled stacked LSTM and GRU model beats all the other models when it comes to R squared and RMSE. This forecasting assists in identifying the virus's potential future pathways of propagation.

4. CONCLUSION

The dynamic nature of the COVID-19 outbreak increases the possibility of a threat. The researchers are searching for novel ways to comprehend the dynamics of COVID-19 that will go beyond the constraints of current epidemiological models. This review investigated several forecasting models important for disease prediction and gave an idea of the varying trend of COVID-19 cases worldwide. It has been discovered that hidden layer architectural models containing LSTM can produce better predictions for time series-based prediction using deep learning. Therefore, the lack of data can be overcome by reducing the time value steps so that the data becomes more so that it allows the Deep Learning process to be better. Despite its limitations, forecasting can unquestionably be employed as a preventative strategy. To produce more accurate results, future research may be enhanced by incorporating new elements and algorithms into the hybrid model.

REFERENCES

[1] World Health Organization, Novel Coronavirus (2019-nCoV) SITUATION REPORT - 1 (2020), https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf.

[2] Yen-Chin Liu, Rei-Lin Kuo, Shengchung Shih, COVID-19: the First Documented Coronavirus Pandemic in History, Biomedical Journal. 43(4) (2020) 328-333.

[3] L. Wang, C-reactive protein levels in the early stage of COVID-19, Medecine et maladies infectieuses. 50(4) (2020) 332–334.

[4] E. Shadmi, Y. Chen, I. Dourado, Inbal Faran-Perach, J. Furler, P. Hangoma, P. Hanvoravongchai, C. Obando, V. Petrosyan, Krishna D. Rao, A. L. Ruano, L. Shi, L. Eugenio de Souza, Sivan Spitzer-Shohat, E. Sturgiss, R. Suphanchaimat, Manuela Villar Uribe and Sara Willems, Health equity and COVID-19: global perspectives, International journal for equity in health. 19(1) (2020) 104.

[5] Okabe, Y., & Shudo, A. (2020). A Mathematical Model of Epidemics—A Tutorial for Students. Mathematics, 8(7), 1174. https://doi.org/10.3390/math8071174

[6] Taylor, S. J., & Letham, B. (2017). Forecasting at Scale. The American Statistician, 72(1), 37–45. https://doi.org/10.1080/00031305.2017.1380080

[7] Arbib, M. A. (2000). Warren McCulloch's search for the logic of the nervous system. Perspectives in Biology and Medicine, 43(2), 193–216. https://doi.org/10.1353/pbm.2000.0001.

[8] Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. Neural Networks, 2(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

[9] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533–536. https://doi.org/10.1038/323533a0.

[10] Ushiku, Y. (2021). Long Short-Term Memory. Computer Vision, 768–773. https://doi.org/10.1007/978-3-030-63416-2_856.

[11] Graves, A. (2012). Long Short-Term Memory. Studies in Computational Intelligence, 37–45. https://doi.org/10.1007/978-3-642-24797-2_4.

[12] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[13] Xiao, Y., & Yin, Y. (2019). Hybrid LSTM Neural Network for Short-Term Traffic Flow Prediction. Information, 10(3), 105. https://doi.org/10.3390/info10030105.

[14] Culhane, P. J., Friesema, H. P., & Beecher, J. A. (2019). Forecast Accuracy: A Summary. Forecasts and Environmental Decisionmaking, 227–256. https://doi.org/10.1201/9780429035838-7.

[15] Fürnkranz, J., Chan, P. K., Craw, S., Sammut, C., Uther, W., Ratnaparkhi, A., Jin, X., Han, J., Yang, Y., Morik, K., Dorigo, M., Birattari, M., Stützle, T., Brazdil, P., Vilalta, R., Giraud-Carrier, C., Soares, C., Rissanen, J., Baxter, R. A., & Bruha, I. (2011). Mean Absolute Deviation. Encyclopedia of Machine Learning, 652–652. https://doi.org/10.1007/978-0-387-30164-8_524.

[16] Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. Neurocomputing, 192, 38–48. https://doi.org/10.1016/j.neucom.2015.12.114

[17] Murphy, A., & Moore, C. (2019). Mean squared error. Radiopaedia.org. https://doi.org/10.53347/rid-68955.

[18] Judgement, Bias, and Mean Square Error. (2021). Intermittent Demand Forecasting, 225–242. https://doi.org/10.1002/9781119135289.ch10.

[19] Karunasingha, D. S. K. (2021). Root Mean Square Error or Mean Absolute Error? Use their ratio as well. Information Sciences. https://doi.org/10.1016/j.ins.2021.11.036.

[20] Chai, Tianfeng & Draxler, R.. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?. Geosci. Model Dev.. 7. 10.5194/gmdd-7-1525-2014.

[21] Lusk, S. L. (1981). Test Anxiety, Level and Accuracy of Predicted Performance. Psychological Reports, 49(2), 527–532. https://doi.org/10.2466/pr0.1981.49.2.527.

[22] Pustokhin, D. A., Pustokhina, I. V., Dinh, P. N., Phan, S. V., Nguyen, G., Joshi, G. P., & K., S. (2020). An effective deep residual network based class attention layer with bidirectional LSTM for diagnosis and classification of COVID-19. Journal of Applied Statistics, 1–18. https://doi.org/10.1080/02664763.2020.1849057.

[23] Arora, P., Kumar, H., & Panigrahi, B. K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. Chaos, Solitons & Fractals, 139, 110017. https://doi.org/10.1016/j.chaos.2020.110017.

[24] Babukarthik, R. G., Adiga, V. A. K., Sambasivam, G., Chandramohan, D., & Amudhavel, J. (2020). Prediction of COVID-19 Using Genetic Deep Learning Convolutional Neural Network (GDCNN). IEEE Access, 8, 177647–177666. https://doi.org/10.1109/access.2020.3025164

[25] Kumari, R., Kumar, S., Poonia, R. C., Singh, V., Raja, L., Bhatnagar, V., & Agarwal, P. (2021). Analysis and predictions of spread, recovery, and death caused by COVID-19 in India. Big Data Mining and Analytics, 4(2), 65–75. https://doi.org/10.26599/bdma.2020.9020013.

[26] Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET. (2021). Procedia Computer Science, 179, 524–532. https://doi.org/10.1016/j.procs.2021.01.036

[27] Chandra, R., Jain, A., & Singh Chauhan, D. (2022). Deep learning via LSTM models for COVID-19 infection forecasting in India. PLOS ONE, 17(1), e0262708. https://doi.org/10.1371/journal.pone.0262708

[28] Sah, S., Surendiran, B., Dhanalakshmi, R., Mohanty, S. N., Alenezi, F., & Polat, K. (2022). Forecasting COVID-19 Pandemic Using Prophet, ARIMA, and Hybrid Stacked LSTM-GRU Models in India. Computational and Mathematical Methods in Medicine, 2022, 1–19. https://doi.org/10.1155/2022/1556025

AUTHOR PROFILES

Akanshika Gandhi received the Master's degree in computer science from the J.C. Bose University of Science and Technology, YMCA, Faridabad, Haryana, in 2022. She has completed her B.Tech from The NorthCap University, Gurugram. Her research interests include Machine Learning and Deep Learning.

Ved Pal is working as an Assistant Professor in the department of Computer Applications in J. C. Bose University of Science& Technology, YMCA, Faridabad, Haryana, India. He has completed Ph.D. in Computer Engineering from J. C. Bose University of science & Technology, YMCA in 2019, M. Tech (CE) From YMCA University of Science & Technology, Faridabad in year 2012, MCA from MD University Rohtak in year 2008. He has thirteen years of experience in teaching. His Research Areas include Software engineering, Software Testing, Object Technology, Data Science and Machine Learning.

Harish Kumar received his Bachelor and Master degrees in engineering from the YMCA Institute of Engineering in Faridabad, India which was affiliated with MDU Rohtak. He earned his Ph.D. from YMCA University of Science and Technology, Faridabad, Haryana, India. He is an Associate Professor in Department of Computer Engineering at J.C. Bose University of Science and Technology, YMCA, Faridabad, Haryana, India. He is a well-respected educator with a substantial body of research and has more than 16 years of teaching experience. He published more than 30 research papers in various national and international journals. Machine learning, software testing, project management, and computer programming are among his research interests.