Journal of Information Technology Research
Volume 15 • Issue 1

Day-Level Forecasting of COVID-19 Transmission in India Using Variants of Supervised LSTM Models:
Modeling and Recommendations

Elangovan Ramanuja, Department of Computer Science and Engineering, School of Engineering and Technology, Christ University (Deemed-to-be), Bengaluru, India

C. Santhiya, Department of Information Technology, Thiagarajar College of Engineering, India

S. Padmavathi, Department of Information Technology, Thiagarajar College of Engineering, India

ABSTRACT

The novel Corona virus SARS-CoV-2 started with strange pneumonia of unknown cause in Wuhan City, Hubei Province of China. On March 11, 2020, the World Health Organization declared the COVID-19 outbreak as a pandemic. Due to this pandemic situation, countries all over the world suffered from economic and psychological stress. To analyze the growth of this pandemic, this paper proposes a supervised LSTM model and its variants to predict the infectious cases in India using a publicly available dataset from Johns Hopkins University. Experimentation has been carried out using various models and window hyper-parameters to predict the infectious rate ahead of a week, 2 weeks, 3 weeks, and a month. The prediction results infer that every individual in India has to be safe at home and to follow the regulations provided by ICMR and the Indian Government to control and prevent others from being infected by this complicated epidemic.

KEYWORDS

Coronoavirus, COVID-19, Deep Learning, Forecasting, Infectious Cases, LSTM Models, Pneumonia Attack, Prediction, RNN Models

INTRODUCTION

World Health Organization (WHO) named the disease caused by novel Coronavirus SARS-CoV-2 as COVID-19 on February 11, 2020 (Li et al., 2020; Dong et al., 2020; Law et al., 2020; Hamid et al., 2020). COVID-19 has initially started with strange new pneumonia of unknown cause in Wuhan City, Hubei province of China (Hui et al, 2020) on December 31, 2019 (WHO, 2020; Sarkodie & Owusu, 2020; Shi et al., 2020). On March 11, 2020, the WHO declared COVID-19 outbreak as pandemic as it widely spreads and infects many people all over the world (WHO, 2020). The entire trajectory of COVID-19 inflammation is growing rapidly in all the countries but awareness about the disease is very low among the population. Many countries are quarantining the affected peoples to reduce the cause of spread and even some people do not know their infection because of no such large number of testing laboratories available in all cities and provinces.

In India, the first case was reported on 30th January 2020 that the patient has travel history from some other country (ICMR, 2020). On 4th March 2020, India has a sudden hike in the increase of

DOI: 10.4018/JITR.299376

infectious rate as then the rate increases very slowly day by day. On 29th March 2020, it crosses more than a thousand in the infectious rate and still, it is increasing in large number. On 27th June 2020, the total number of infected/ confirmed cases is about 67,3165 as per the Indian Council of Medical Research (ICMR) record (ICMR, 2020). Moreover, many countries as like India has more complications in handling COVID-19 due to an increase in globalization, population density in cities, non-compliance of supply and demand of medical facilities, drugs and shortcoming of medical equipment and protections. However, in this complicated pandemic situation forecasting of numbers ahead of month or bi-month may be a significant and alternate strategy to prevent and control the infections. To predict the temporal order of infectious rate, this paper proposes a supervised Long-Short Term Memory (LSTM) model and its variants to predict the number of possible infectious cases ahead of 7, 14 days and a month. The publicly available dataset has been used to predict the infectious rate through the variants of proposed supervised LSTM models such as LSTM, Bi-LSTM and Stacked LSTM.

This kind of infection rate prediction has been analyzed recently in previous studies for COVID-19 pandemic situations. However, the proposed studies mostly follow only a data driven approach and those methods are linear in nature that often neglects the temporal components of the data. Statistical methods such as Auto Regressive Moving Average (ARIMA), Moving Average (MA), Auto Regressive (AR) methods, GARCH models tremendously depends on assumptions and such models are also difficult for forecasting real transmission of COVID-19 epidemic situation (Tomar and Gupta, 2020; Elmousalami and Hassanien, 2020). Certain other studies (Chimmula & Zhang, 2020 and Ayyoubzadeh et al, 2020) have proposed deep learning models for the analysis of infectious cases and death rates using traditional training and testing rather than prediction. Moreover, these studies have better inferences in prediction for shorter periods with minimum loss. In case of longer prediction for a month or above, these technique produces more error.

To analyze the epidemic level for a shorter to longer duration this paper proposes an LSTM model and its variants using one-step prediction with windowing technique. The proposed LSTM model uses R0 reproductive number - a statistical method to infer the prediction of an infectious cause for a month in India. The empirical analysis of infectious case prediction mainly depends on the past historical data collected. However, the COVID-19 epidemic depends on a large number of external and global factors. So this paper suggests a recommendation to all Indian individuals about the awareness and cause of COVID-19 ahead of the month using various analyses.

The paper composed of 5 sections, including the introductory one. Section 2 discusses the deep learning based prediction models of pandemic and endemic diseases, Section 3 details the materials and methods used. Section 4 presents the results and discussion of LSTM models and its variants with research finding. Finally, Section 4 concludes the paper.

2. RELATED WORK

The novel Coronavirus started its pandemic with strange new disease COVID-19 caused by pneumonia in Wuhan, Hubei province of China (Hui et al, 2020) on December 31,2019. From the day of the first reported case, the World Health Organization and the medical researchers are in an urge to provide a stronger vaccine that controls this pandemic spread of COVID-19. On the other side, the academic researchers concentrate more on identifying the best prediction model to help the front line workers in the process to control the spread of COVID-19 (Kirbas et al, 2020). Recently, during January - July 2020, there has been a wide variety of research works proposed to predict the COVID-19 positive, death and infected cases in various geographical locations all over the world (Shorabhi et al, 2020).

Rafig et al, 2020 have proposed a prognostic yet deterministic state space model to forecast the spread and causes ahead of 30 days for 10 most affected states in India. Singhal et al, 2020 have proposed two contrasting mathematical predictive models such as Fourier Decomposition Method (FDM) on Discrete Cosine Transform (DCT) and Gaussian mixture models for the prediction ahead of 30days. This proposed method predicts the cause for India, Italy and the United States of India. Gola et al, 2020 have proposed exponential and Least square curve fitting model with a process of fine tuning to achieve better performances. Dhamodharavadhani et al, 2020 have investigated suitable statistical neural network models and their hybrid version for COVID-19 mortality rate prediction in India. The proposed technique uses Radial Basis Function Neural Network model and Generalized Regression Neural Network model has been utilized. Choudhury et al, 2020 have proposed a validated semi-mechanistic stochastic model to generate short term forecasts. This analysis uses a simplified transmission model using Markov Chain Monte Carlo simulation with Metropolis Hastings updating.

In addition to the proposed models, most of the techniques have proposed Long Short Term Memory (LSTM) model that best predicts the time series data for a shorter duration (Kirbas et al, 2020). Chimulla et al, 2020 have proposed a real time forecasting model for COVID-19 transmission to help the frontline health workers and government policy makers using the LSTM model. This proposed LSTM model predicts the trends of different countries ahead of 14 days and the performances are compared with Canadian data. Hridoy et al, 2020 have proposed a LSTM model and Logistic curve method to predict the possible number of cases in Bangladesh ahead of a month. Tomar and Gupta, 2020 have proposed a LSTM model and curve fitting for the prediction of the number of COVID-19 cases in India 30 days ahead. This work also suggested preventive measures like social isolation and lockdown on the spread of COVID-19. Bedi et al, 2020 have proposed a modified Susceptible-Exposed-Infected-Recovered-Deceased (SEIRD) model for long term prediction and LSTM model for short term prediction of COVID-19 cases for next 30 days. Jana and Ghose, 2020 have proposed a short term adaptive prediction model upto 3 - 4 weeks using the SEIRD model. Anirudh 2020 has analyzed the challenges and outcomes of various mathematical models used for prediction of COVID-19 spread, peak and death cases. Kırbaş et al, 2020 have predicted the confirmed COVID-19 cases of various European countries such as Denmark, Belgium, Germany, France, United Kingdom, and Turkey using Auto-Regressive Moving Average Model (ARIMA), Nonlinear AutoRegressive Neural Network Model (NARNN), and Long Short-Term Memory model (LSTM). This paper predicts the confirmed cases 14 days earlier to the infected history and the models were evaluated using MSE, PSNR, RMSE, NRMSE, MAPE and SMAPE.

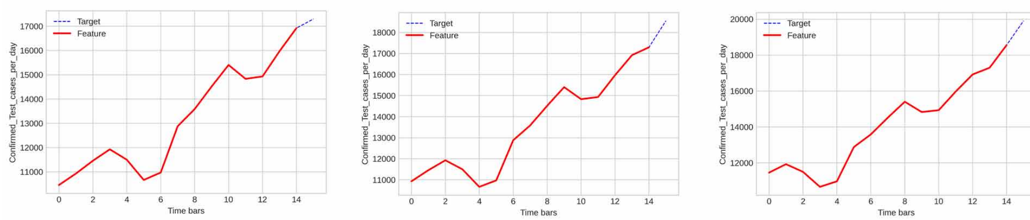
Though various research works have proposed LSTM models in the literature, there is no specified model that best predicts the data earlier to 7, 14 and 30 days. Moreover, no research work has analyzed the variants of LSTM models such as Bi-LSTM and Stacked LSTM models, that may show better performances in certain cases. Finally, the LSTM model has dependencies on window size of the training data that may predict the data in most accurately. To incorporate the issues identified in state-of-the-art techniques and to provide better accuracy in prediction, this paper proposes the LSTM model and its variants to predict the spread of COVID-19 cases.

3. MATERIALS AND METHODS

3.1 Data Collection

The proposed model uses publicly available COVID-19 dataset recorded from 22nd January 2020 for nearly 187 countries and territories all over the world (COVID-19). The dataset comprises 8 attributes representing time-stamp, name of the country, name of state or province, latitude, longitude and the time series data tracking the number of infectious cases, recovered cases and death cases. Every day, the data will be updated automatically from the upstream repository maintained by John Hopkins University Centre for System Sciences and Engineering (CSSE). In this proposed work, the data is collected on 5 July 2020 for the time period of 22 January 2020 (considered as 0th day) and 4 July 2020 (164th day). The clean data collected is transformed for the process of the proposed supervised LSTM model using a sliding window (lag observation). The sliding window step temporal data are scaled as features to the model that predicts one-step value (target) from the current sliding window. Naturally, the features are the current window temporal data and the target is the first time step of the next window. Figure 1 shows the sample consecutive three sliding windows in which the feature and target are highlighted in red and blue color respectively. Iteratively, the sliding window temporal data are used as input to derive one-step prediction as output from the LSTM model.

Figure 1. The feature and target of second consecutive sliding window of size 15.



3.2 Supervised LSTM Models

The proposed supervised LSTM (Gers et al, 1999) model forecasts the time series data of infectious cases in India over a period of days using the past historical data. The LSTM models are popularly used in time series forecasting as it has certain advantages over RNN and other sequences models. First, the data signals can travel in backward directions as well as the network has the feedback connections so that the LSTM model vanishes gradient problem occurred in other models. Further, it has the capability to maintain long term dependency of data as time series have an unknown lag of duration between the events. These advantages are majorly due to the usage of serious of gates and cells that are maintained in memory blocks connected through layers rather than neurons used in traditional feed forward networks. Figure 2 shows the LSTM cell that has three gates, Input gate, Forget gate and Output gate. Input gate scales the input to the cell to perform a write operation, Forget gate scales and reset the old cell value, and Output scales the output to the cell to perform read. Gates in the LSTM model operates like a switch that automatically controls the read/ write operation to incorporate long-term memory function in the model. The mathematical model and the working principle of LSTM cell is clearly given in Yao et al, 2015. Even more, if the model identifies any significant feature initially in the input sequence, it can easily transmit the information over a long distance to formulate the final decision. Finally, LSTM units formulate the decisions concerning the current memory through specific gates in the network. Moreover, to add up an advantage of LSTM mode, Bi-LSTM and Stacked LSTM have been proposed for prediction by Graves et al, 2005 and Yao et al, 2015 respectively. To utilize the advantage of those models, the proposed work also concentrated on implementing Bi-LSTM and Stacked LSTM models.

Bidirectional LSTM (Graves et al, 2005) also known as Bi-LSTM is an extension of the traditional LSTM model. This model has an advantage in training the available input information both in the past and in the near future for a specific time frame which improves the performance of forecasting. Further, Bi-LSTM duplicates the first recurrent layer to provide two layers that allow the network to have both forward and backward directions during data transmission. The Stacked LSTM model simulates a Multiple LSTM architecture layer. This works like one output per one input time step rather than one output time step for all input time steps. Stacked LSTM (Yao et al, 2015) model is a stable technique generally attributed to forecasting a wide range of challenging problems. Diagrammatic representation of all these three models are shown in Figure 3. Using these time related advantages of the LSTM model in forecasting, the LSTM model and its variants such as Bi-LSTM and Stacked LSTM have been proposed to forecast temporal near future and infectious progress of COVID-19 in India.

Figure 2. The LSTM cell

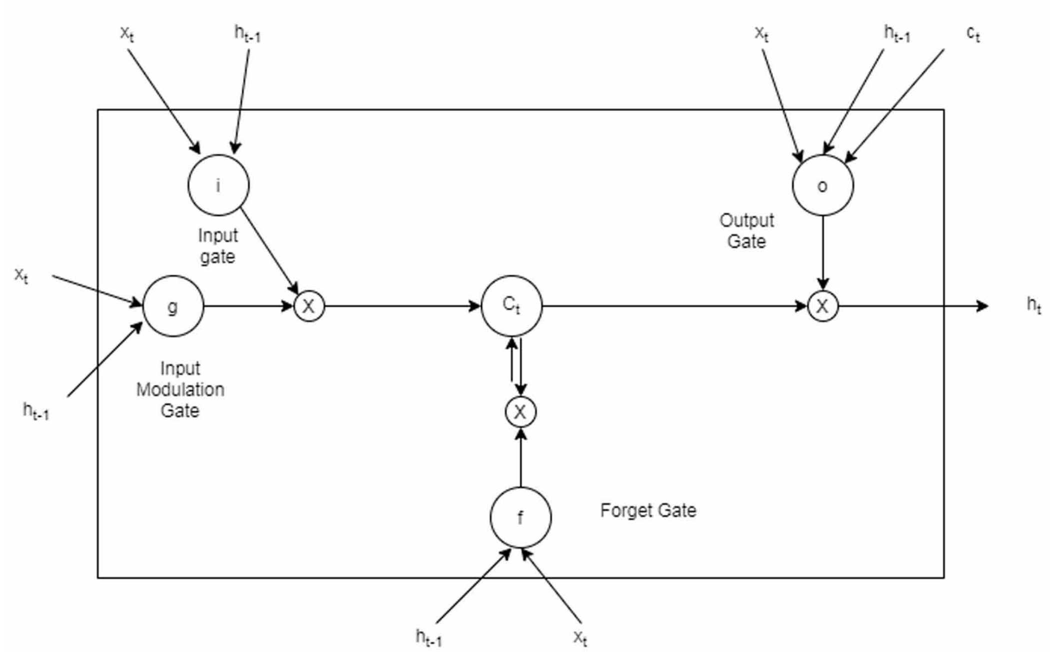
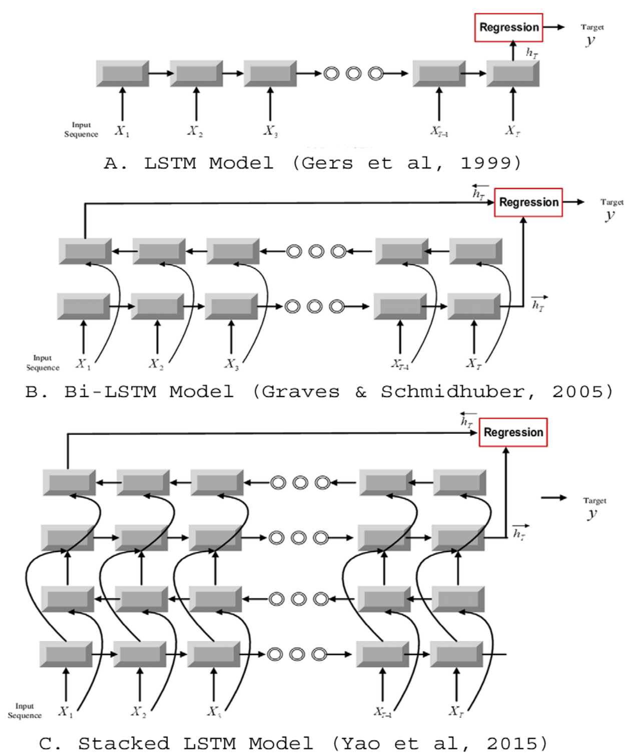


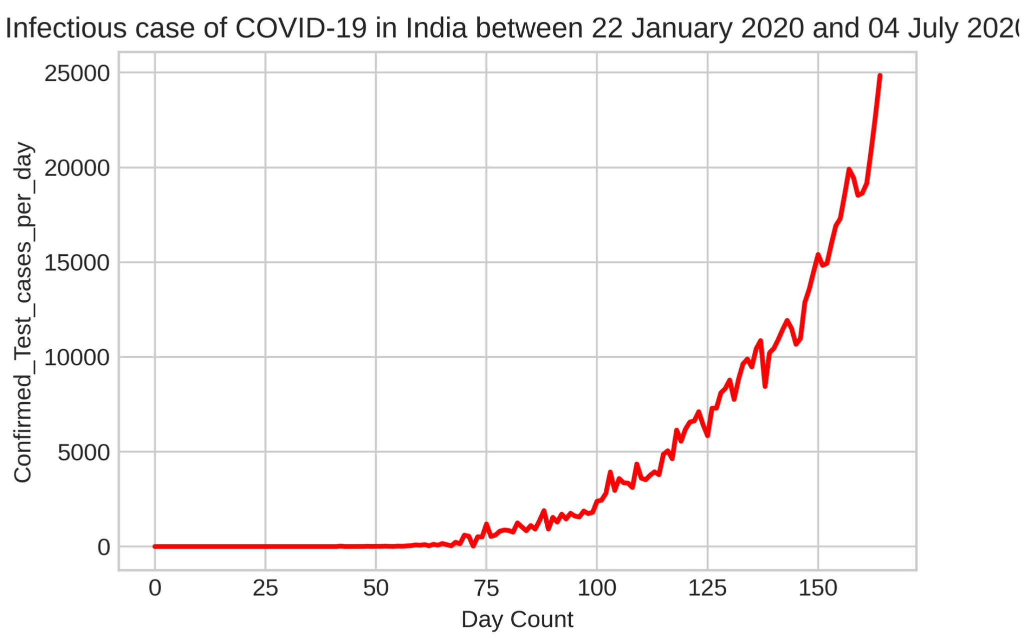
Figure 3. LSTM Model and its variants



4. EXPERIMENTAL RESULTS AND DISCUSSIONS

To utilize the time related advantages of LSTM models, the COVID-19 infectious cases of India from 22 January 2020 (considered as 0th day) to 4 July 2020 (considered as 164th day) have been used to evaluate the variants of LSTM models. Figure 4 shows the COVID-19 infectious cases of India from 22nd January 2020(0th day) and 4 July 2020(164th day).

Figure 4. COVID-19 infectious cases of India between 22 January 2020 (0th day) and 04 July 2020 (165th day)



4.1 Unit Root Testing

Generally, LSTM models perform better for the case of non-stationary signals (Bala and Singh, 2019) rather than stationary signals. To validate the stationary in the signal i.e trend or seasonality, Augmented Dickey-Fuller (ADF) (Cheung and Lai, 1995) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) (Li, 1999) test have been used. The test results are interpreted using the p-value. ADF test is a unit root test with consideration of null hypothesis H_0 and alternate hypothesis H_1 are

H_0 : The input data signal doesn't have unit root and so it is stationary

H_1 : The input data signal have unit root and so it is non-stationary

If the p-value is less than 0.05 (p-value < 0.05) it rejects the alternate hypothesis H_0 and the signal is confirmed to be stationary. In the case of a p-value greater than 0.05 (p-value > 0.05), it supports alternate hypothesis H_1 and the signal confirms to be non-stationary. Figure 5 shows the

ADF test results for the collected COVID-19 confirmed test cases and that confirms the signal to be non-stationary.

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) is another unit root testing method that is contrary to the ADF test. KPSS test with the consideration of Null hypothesis H_0 and alternate hypothesis H_1 is

H_0 : The input time series data have unit root and so it is non-stationary

H_1 : The input time series data doesn't have unit root and so it is stationary

In the case of a p-value less than 0.05 it rejects the alternate hypothesis H_1 and confirms the signal to be non-stationary and the case is vice-versa for p-value greater than 0.05. Figure 6 shows the KPSS test of COVID-19 cases in India and that confirms the signal to be non-stationary.

Figure 5. Augmented Dickey-Fuller Test for the collected COVID-19 confirmed test cases

```
Augmented Dickey-Fuller Test Results:
ADF Test Statistic          4.668814
P-Value                     1.000000
# Lags Used                12.000000
# Observations Used       145.000000
Critical Value (1%)        -3.476273
Critical Value (5%)        -2.881688
Critical Value (10%)       -2.577513
dtype: float64
Is the time series stationary? False
```

Figure 6. KPSS test for the collected COVID-19 confirmed test cases

```
KPSS Statistic: 0.9377391947504115
p-value: 0.01
num lags: 14
Critial Values:
   10% : 0.347
   5% : 0.463
   2.5% : 0.574
   1% : 0.739
Result: The series is not stationary
```

The proposed supervised LSTM model and its variants are implemented with Adam optimizer and Mean Squared error as loss function in Google Colabs using Tensor Flow as backend for the collected data signal as it is confirmed to be non-stationary. The Keras API has been used wherever necessary to define the LSTM model and its variants with the number of neurons in all the three models are considered to be 5 and activation function as 'relu' and the number of epochs to be 200. The performances of all the LSTM models are measured using familiar evaluation metrics such as RMSE and MAPE as given in equation (1) and (2). N is the length of the data signal, y_t is the observed time step and $\widehat{y_t}$ is the predicted time step. The hyper-parameters are fixed that fit our case best for forecasting according to different papers (Lipton et al, 2015) that are highly valuable.

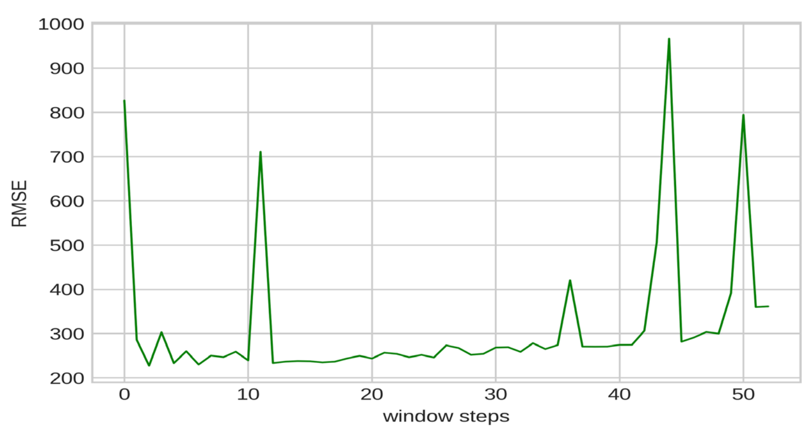
$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[y_t - \widehat{y_t}\right]^2} \tag{1}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y_t - \widehat{y_t}}{y_t}\right| \tag{2}$$

4.2 Prediction

To validate the performance of proposed LSTM model and its variants, infectious case for the period of 7 days(June 28, 2020 to July 4, 2020) has been predicted using the data signal collected between January 22,2020 (0th day) and June 27, 2020 (157th day). As the proposed model requires an optimal window size for iterative one-step prediction, an empirical test has been carried out using the LSTM model to identify the optimal window of size 30% of the original data signal. In this case, the LSTM model is trained with 80% of data (126 days out of 158 days) to identify the best optimal window size that has the smallest training RMSE value. Figure 7 illustrates the empirical testing for finding the optimal window length. From Figure 7, the window size 2, 6, 4, 12, 16 and 13 that has least RMSE values are considered for the experimentation.

Figure 7. Illustrating empirical testing for finding the optimal length for the window hyper-parameter



Using the window size of 2, 6, 4, 12, 16 and 13 the LSTM models and its variants are implemented to forecast the COVID-19 infectious cases for a period of 7 days (June 28, 2020 to July 4, 2020) through the data collected from January 22, 2020 (0th day) and June 27, 2020 (157th day). The predicted results of all the models LSTM, Bi-LSTM and Stacked LSTM are shown in Table 1 for selected optimal window size with evaluation metrics RMSE and MAPE. On comparing the results from Table 1, it is clear that the three models have a minor deviation in its predicted MAPE value. Further, all the MAPE values are less than 10 which shows the excellent characteristics of the proposed model in prediction. The LSTM model with a window size of 12 shows the highest efficiency of 1102.92 as RMSE and MAPE of 4.86 that is very closer to Stacked LSTM of window size 13 with RMSE 1148.28 and MAPE 4.82.

Table 1. Forecast of COVID-19 infectious rate for the period of 7 days (June 28, 2020 to July 4, 2020)

S.No	Window Size	Model	RMSE	MAPE
1	2	LSTM	1864.74	8.16
2		Bi-LSTM	2018.68	8.80
3		Stacked LSTM	**1715.32**	**7.29**
4	6	LSTM	1872.08	8.05
5		Bi-LSTM	2262.55	10.03
6		Stacked LSTM	**1479.26**	**6.35**
7	4	LSTM	1511.03	6.59
8		Bi-LSTM	**1224.79**	**5.52**
9		Stacked LSTM	1372.19	6.38
10	12	LSTM	**1102.92***	**4.86***
11		Bi-LSTM	1442.08	6.10
12		Stacked LSTM	1396.35	6.07
13	16	LSTM	2057.61	8.75
14		Bi-LSTM	1892.04	8.17
15		Stacked LSTM	**1783.18**	**8.02**
16	13	LSTM	1304.82	5.42
17		Bi-LSTM	1554.44	6.34
18		Stacked LSTM	**1148.28**	**4.82**

Also, to view the changes between the observed and predicted outcomes, the results are also shown graphically in Figure 8, Figure 9, and Figure 10 for LSTM, Bi-LSTM and Stacked LSTM respectively. On comparing the graphical prediction in Figure 8 for the LSTM model, the predicted results for window size 12 and 13 are very closer to observed data and that has been proved with minimum RMSE and MAPE value. In viewing the predicted results of Bi-LSTM in Figure 9, the window size of 4 has a closer correlation with the observed value and that has minimal RMSE value of 1224.79 and MAPE value of 5.52. Further on comparing the predicted results of Stacked LSTM in Figure 10, the window size 4, 16 and 13 has closer coherence with observed value and that shows higher efficiency with minimal RMSE and MAPE values. By comparing all the prediction results by the three models, the LSTM model with window size 12 has achieved maximum efficiency and thus the LSTM model with window size 12 is used for further prediction and analysis.

Figure 8. LSTM Model with sliding window values (A) 2 (B) 6 (C) 4 (D) 12 (E) 16 and (F) 13

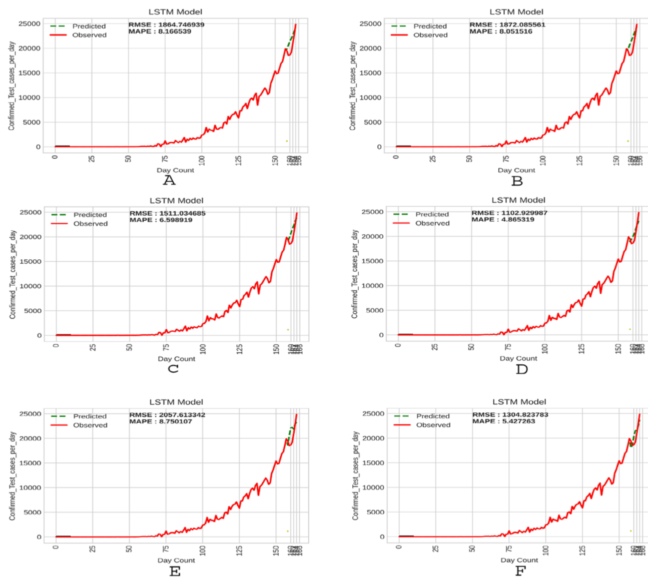


Figure 9. Bi-LSTM Model with sliding window values (A) 2 (B) 6 (C) 4 (D) 12 (E) 16 and (F) 13

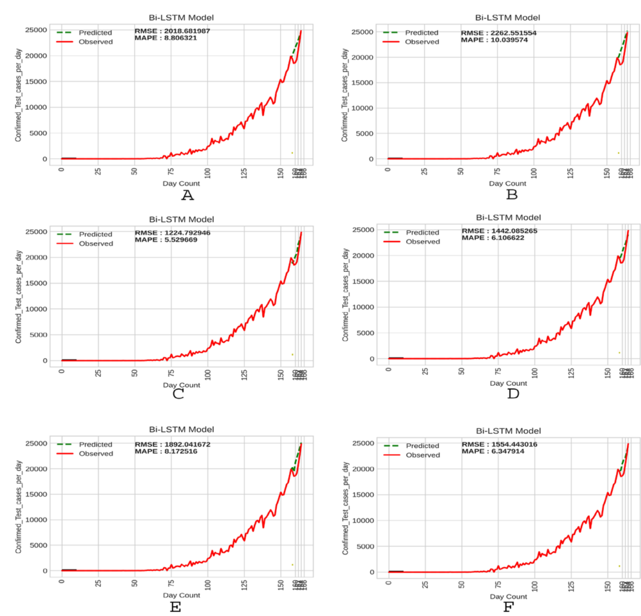
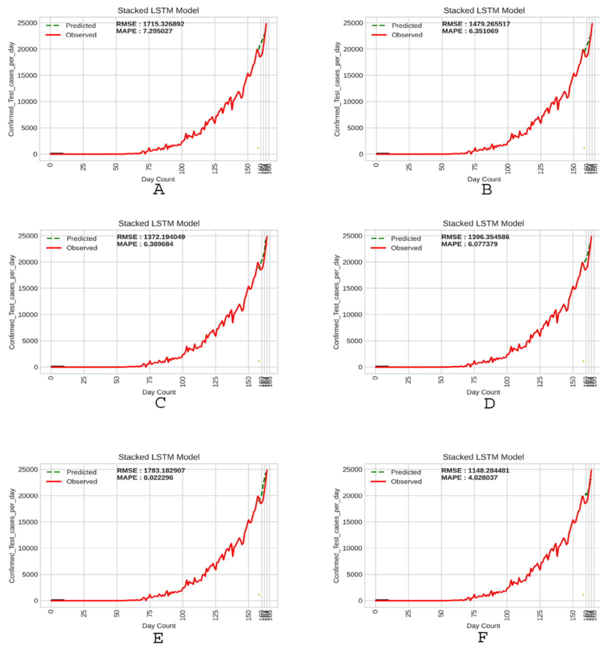
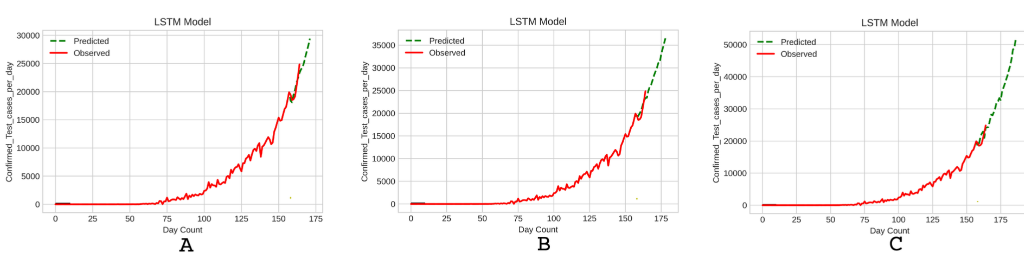


Figure 10. Stacked LSTM Model with sliding window values (A) 2 (B) 6 (C) 4 (D) 12 (E) 16 and (F) 13



To attain the other objective of the paper, the LSTM model with window size 12 has been used to predict the number of infectious cases per day for another period of 2 weeks (June 27, 2020, to July 11, 2020), 3 weeks (June 27, 2020, to July 18, 2020) and one month (June 27, 2020, to July 22, 2020). Figure 11 (a) shows the predicted result of the LSTM model for a period of 2 weeks. During the period of July 4, 2020, to July 11, 2020 the infectious count crosses 25,000 to 30,000 per day. According to prediction for another week as shown in Figure 11(b), this epidemic situation may exponentially grow and that may crosses more than 35,000 per day from July 11 to July 18, 2020. Further, Figure 11(c) infers the COVID-19 pandemic may be more severe during the last week of July 2020, and that reaches the count of more than 50,000. The prediction results using the supervised LSTM model alarms each and every Indian individual to follow the norms specified by the Indian government and ICMR to prevent and control this epidemic. However, the prediction shows severity, the cause, and spread can be controlled by external factors such as lockdown of cities, following social distance, wearing masks, etc. Moreover, this is not only a responsibility of the Government, it is the individual's responsibility to be safe at home and to avoid spreading the this COVID-19 pandemic.

Figure 11. Predicted results of LSTM Model with sliding window size of 12 (A) 2 weeks (B) 3 weeks (C) 1 month



5. CONCLUSION

In this paper, the supervised LSTM model and its variants Bi-LSTM and Stacked LSTM have been proposed to forecast the COVID-19 outbreak in India. The proposed model predicts the infectious count per day of India ahead of a week, 2 weeks, 3 weeks and a month using the historical data collected from John Hopkins University. The forecast trend and pattern infers, the COVID-19 pandemic may have high severity during the last week of July 2020 and the infectious count reaches the maximum of 50,000 per day. The proposed prediction using the LSTM model alarms the individual to follow the regulations provided by the Indian Government and Indian Council of Medical Research to prevent and control the epidemic rather depend only on Medical treatments. In future, the other hybrid mathematical models can also be applied for efficient results. Moreover, the prediction models can be integrated with Multi Criteria Decision making models to provide best preventive measures based on the rate of infections predicted. This model may be best suited for prevention and the quantity of infection with respect to the geographical locations all over the world.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

Anirudh, A. (2020). Mathematical modeling and the transmission dynamics in predicting the Covid-19-What next in combating the pandemic. *Infectious Disease Modelling*, *5*, 366–374. doi:10.1016/j.idm.2020.06.002 PMID:32666005

Ayyoubzadeh, S. M., Ayyoubzadeh, S. M., Zahedi, H., Ahmadi, M., & Kalhori, S. R. N. (2020). Predicting COVID-19 incidence through analysis of google trends data in iran: Data mining and deep learning pilot study. *JMIR Public Health and Surveillance*, *6*(2), e18828. doi:10.2196/18828 PMID:32234709

Bala, R., & Singh, R. P. (2019, July). Financial and Non-Stationary Time Series Forecasting using LSTM Recurrent Neural Network for Short and Long Horizon. In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

Bedi, P., Gole, P., Gupta, N., & Jindal, V. (2020). *Projections for COVID-19 spread in India and its worst affected five states using the Modified SEIRD and LSTM models.* arXiv preprint arXiv:2009.06457.

Cheung, Y. W., & Lai, K. S. (1995). Lag order and critical values of the augmented Dickey–Fuller test. *Journal of Business & Economic Statistics*, *13*(3), 277–280.

Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons, and Fractals*, *135*, 109864. doi:10.1016/j.chaos.2020.109864 PMID:32390691

Choudhury, L. P., & Kumar, B. R. (2020). Real-Time COVID-19 Forecasting for Four States of India Using a Regression Transmission Model. *Open Journal of Epidemiology*, *10*(4), 335–345. doi:10.4236/ojepi.2020.104027

COVID-19 dataset. (n.d.). https://github.com/datasets/covid-19

Dhamodharavadhani, S., Rathipriya, R., & Chatterjee, J. M. (2020). COVID-19 Mortality Rate Prediction for India Using Statistical Neural Network Models. *Frontiers in Public Health*, *8*, 441. doi:10.3389/fpubh.2020.00441

Dong, L., Hu, S., & Gao, J. (2020). Discovering drugs to treat coronavirus disease 2019 (COVID-19). *Drug Discoveries & Therapeutics*, *14*(1), 58–60. doi:10.5582/ddt.2020.01012

Elmousalami, H. H., & Hassanien, A. E. (2020). *Day level forecasting for Coronavirus Disease (COVID-19) spread: Analysis, modeling and recommendations.* arXiv preprint arXiv:2003.07778.

Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). *Learning to forget: Continual prediction with LSTM*. Academic Press.

Gola, A., Arya, R. K., Animesh, A., Dugh, R., & Khan, Z. (2020). *Fine-tuned Forecasting Techniques for COVID-19 Prediction in India.* medRxiv.

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, *18*(5-6), 602–610. doi:10.1016/j.neunet.2005.06.042

Hamid, S., Mir, M. Y., & Rohela, G. K. (2020). Noval coronavirus disease (COVID-19): A pandemic (Epidemiology, Pathogenesis and potential therapeutics). *New Microbes and New Infections*, *35*, 100679. doi:10.1016/j.nmni.2020.100679

Hridoy, A. E. E., Naim, M., Emon, N. U., Tipo, I. H., Alam, S., Al Mamun, A., & Islam, M. S. (2020). *Forecasting COVID-19 Dynamics and Endpoint in Bangladesh: A Data-driven Approach.* medRxiv.

Hui, D. S., Azhar, E. E., Madani, T. A., Ntoumi, F., Kock, R., Dar, O., & Zumla, A. et al. (2020). The continuing epidemic threat of novel coronaviruses to global health-the latest novel coronavirus outbreak in Wuhang, China. *International Journal of Infectious Diseases*, *91*, 264–266. doi:10.1016/j.ijid.2020.01.009

ICMR. (2020). *Indian Council of Medical Research*. https://www.icmr.gov.in/

Jana, S., & Ghose, D. (2020). *Adaptive short term COVID-19 prediction for India.* medRxiv.

Kırbaş, İ., Sözen, A., Tuncer, A. D., & Kazancıoğlu, F. Ş. (2020). Comperative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. *Chaos, Solitons, and Fractals*, *138*, 110015. doi:10.1016/j.chaos.2020.110015

Law, S., Leung, A. W., & Xu, C. (2020). Severe acute respiratory syndrome (SARS) and coronavirus disease-2019 (COVID-19): From causes to preventions in Hong Kong. *International Journal of Infectious Diseases*, *94*, 156–163. doi:10.1016/j.ijid.2020.03.059

Li, H., Liu, S. M., Yu, X. H., Tang, S. L., & Tang, C. K. (2020). Coronavirus disease 2019 (COVID-19): Current status and future perspective. *International Journal of Antimicrobial Agents*, *55*(5), 105951. doi:10.1016/j.ijantimicag.2020.105951

Li, Q. (1999). Convergence clubs: Some further evidence. *Review of International Economics*, *7*(1), 59–67. doi:10.1111/1467-9396.00146

Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). *A critical review of recurrent neural networks for sequence learning*. arXiv preprint arXiv:1506.00019.

Rafiq, D., Suhail, S. A., & Bazaz, M. A. (2020). Evaluation and prediction of COVID-19 in India: A case study of worst hit states. *Chaos, Solitons, and Fractals*, *139*, 110014. doi:10.1016/j.chaos.2020.110014

Sarkodie, S. A., & Owusu, P. A. (2020). Investigating the cases of novel coronavirus disease (COVID-19) in China using dynamic statistical techniques. *Heliyon*, 03747.

Shi, P., Dong, Y., Yan, H., Zhao, C., Li, X., Liu, W., He, M., Tang, S., & Xi, S. (2020). Impact of temperature on the dynamics of the COVID-19 outbreak in China. *The Science of the Total Environment*, *728*, 138890. doi:10.1016/j.scitotenv.2020.138890

Singhal, A., Singh, P., Lall, B., & Joshi, S. D. (2020). Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. *Chaos, Solitons, and Fractals*, *138*, 110023. doi:10.1016/j.chaos.2020.110023

Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., ... Agha, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery*.

Tomar, A., & Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *The Science of the Total Environment*, *728*, 138762. doi:10.1016/j.scitotenv.2020.138762

World Health Organization. (2020). *Pneumonia of unknown cause–China. In Emergencies preparedness, response, Disease outbreak news*. WHO.

Yao, K., Cohn, T., Vylomova, K., Duh, K., & Dyer, C. (2015). *Depth-gated LSTM*. arXiv preprint arXiv:1508.03790.

E Ramanujam (Member, IEEE) is currently working as an Assistant Professor in the Department of Computer Science and Engineering, School of Engineering and Technology, Christ (Deemed to be) University, Bengaluru, Karnataka, India. He has more than 10 years of teaching and research experience. He completed his PhD, Degree in Information and Communication Engineering from Anna University, Chennai in the year 2021. He has published research papers in peer-reviewed international journals from IEEE, Inderscience and IGI Global publisher and national and international conferences organized by IEEE. He has authored several chapters in the refereed edited books of IGI and Springer. He also acts as a Reviewer in International Journals such as IEEE, Springer, Elsevier, Emerald and IGI Global Publisher

S. Padmavathi received her Ph.D. from Anna University, Chennai, Tamil Nadu, and India. Currently, she is working as a Professor in the Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India. Her fields of interest are data mining, biomedical and high-performance computing. She has published several papers in reputed international journals and conferences.