International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

COVID Cases Prediction using Time Series Models

Aksh Mehta[1], AyushMattoo[2], Dylan Pereira[3]

[1, 2, 3]Information Technology, Fr. CRCE, Bandra, Mumbai, India

Abstract: ***The disastrous outbreak of Covid-19 has brought a global threat to the living society. This incident of COVID-19 in India was conveyed on 30th January 2020 instigated discovery in Wuhan, China. Every nation is putting incredible efforts into the fight against the spread of this deadly disease in terms of infrastructure, finance, data sources, protective gears, important treatments and several other resources. Artificial intelligence researchers are focusing their specialised knowledge to develop mathematical models to analyse the situation using nationwide shared data. To contribute towards the well-being of living society, this article proposes to utilise machine learning and deep learning models to understand its everyday Behaviour to be exponential along with the prediction of the outbreak across the nations by utilising the real-time information from the live covid website(covid19india.org). Machine Learning can be called one Such area that uses various algorithms to understand the correlation between the given data, visualise and predict the future forecast. The whole world is currently facing a devastating situation due to the covid-19.To control the spread and rising number of active cases in India, we did some research to demonstrate the future forecasting of the total number of active cases in India in the upcoming few days. We did our research on various time series models such as the Arimamodel, Fb-prophet model, LSTM, out of which LSTM proved to give the best result. We collected the real-time data from the live covid website. After which we did data pre-processing and data wrangling. The data set is then turned into the training set and testing set. Finally, the model is trained and tested for accuracy. After completion of testing and training, the model is ready to predict future forecasts.***

Keywords: COVID-19, ARIMA Model, FB Prophet, Machine Learning, Time Series Analysis, forecasting, R-squared score, Root Mean Squared Error, Mean Squared Error.

1. Introduction

As we have seen, nowadays everything is technology based. And in the upcoming years the technology would be more ahead. Most trending and booming technologies in the next few years would be Artificial Intelligence, Machine Learning, Deep Learning, Data Science, Neural Network. ML has also proved to solve many real-time problems in the past few years such as in fields of image processing, medical diagnosis, robotics and many more.One of the major projects on which Elon Musk worked is “Tesla electric self-driving cars” which uses Artificial intelligence. Predicting the weather in the next few days, Predicting the house Price, Predicting the Covid cases in upcoming days is all done with the help of Machine Learning. So, Machine Learning can be defined as the branch of Artificial Intelligence in which the system can learn from data, Identify Patterns and make decisions with minimal human Intervention. The speed at which this virus infects humans is rapidly increasing day by day. The need to minimize deaths and stabilize the country’s economy has now become the priority. After doing lots of research and understanding the demand of Machine Learning in the next few years, our team decided to work on the Project which is based on “Machine Learning” and “Neural Networks”. So our topic is “Covid-19 Outbreak Prediction using Machine Learning and Deep Learning algorithms”. CoronaVirus can be defined as a large family of viruses that are known to cause illness ranging from common cold to more severe cold. Coronavirus disease caused by SARS-CoV-2, was first reported from Wuhan City, China, in December 2019.The most common early symptom is fever, dry cough, tiredness and difficulty in breathing. People with weak immunity mostly aged people are suffering from this disease. Unlike other influenza, covid-19 has a high R0 value (the basic reproduction number, representing viral infectivity) of 3.25–3.4 characterized by human-to-human transmission through the air, which means that it is not easily controlled[3].Our prediction results can also help public health providers and policy makers make the necessary arrangements to respond to the potential changes in covid trend. The experiments are based on a set of covid cases as of between 5th may,2021 to 28th august,2021. Our research was done on various Time series models such as ARIMA, FB-Prophet and LSTM, out of which LSTM proved to give better accuracy than the other two. The loss in LSTM seemed to decrease significantly with increasing epochs. We also tested our LSTM model using different optimizers such asadam,r msprop. The results and details about those algos and optimizers will be discussed as we move ahead.

2. Literature Review

In the last few decades, digital technologies played critical roles in major health sectors for disease prevention, the present worldwide health emergency also trying to seek technological support to tackle covid. In some research papers, authors searched for trending digital technologies such as internet of things (IoT), big-data analytics, artificial intelligence (AI), deep learning and technology to develop various strategies for monitoring, detection and prevention of pandemic; and also to understand the impact of the situation to the healthcare sector [4].In a research work proposed by Benvenuto, authors proposed Arima model to predict the spread of Covid. In the previous research paper, the author forecasted the various parameters for the next few days based on the study about the incidence of the covid-19. Their research work also demonstrated the correlogram and Arima forecast graph for the epidemic incidence and prevalence. The authors found that ARIMA (1,0,4) was found to be the best ARIMA Model[5]. Given the severity of the outbreak, predicting when the pandemic will end is very important for the production and life of affected countries. Many Researchers have made efforts in the fight against the pandemic, and a number of predictive models based on mathematical model, infectious disease model and machine learning model are designed to forecast the trend covid [6], [7].Mathematical and machine learning techniques has been quite successful and widely used in time series prediction, such as electricity price prediction[8], stock price prediction [9] and air pollution forecast[10], covid cases prediction. Recently, time series prediction models such as Auto-Regressive Integrated Moving Average (ARIMA), Nonlinear Auto Regression Neural Network (NARNN) and Long-Short Term Memory (LSTM) approaches are used to model the prediction of epidemic diseases. According to the results of the survey[11], LSTM was found the most accurate model.

Also, there are some shortcomings in the existing LSTM prediction literatures. For eg, to obtain the long-term prediction curve, Yang et al. [12] trained the LSTM model which was based on the SARS data in 2003, as the time series data of SARS are complete. As there are many differences between SARS and COVID-19 in terms of latency rate, latency period, mortality rate and how developed the world is at the time of the outbreak that simulations can be unreliable. Some scholars have also designed the prediction models with LSTM, using cumulative number of confirmed cases as training set to train the model, however, it is only able to predict the rising trend of the epidemic within the next 30 days, and cannot predict when the epidemic will decline or end .Since the cumulative confirmed data shows an overall upward trend and is not a smooth sequence LSTM is not very applicable with it.

3. Materials and Methods

1) Dataset

Our Dataset consists of data of covid-19 from May, 2021 to August, 2021 which was retrieved from official Covid Website (covid19ind.org) [13]. The dataset consists of daily cases in various states at different time. It also consists of the status of cases in each state such as (Recovered, Death, Confirmed).It is easy to observe the exponential growth of the spread which needs to be controlled. We have used 75% data for training purpose and 25% for testing purpose. Below is the overview of our dataset after cleaning the dataset:

Table 1: COVID-19 Time-Series of India on Daily Basis.

	Num Cases	Date Announced
0	443.0	05/05/2021
1	151.0	05/05/2021
2	483.0	05/05/2021
3	52.0	05/05/2021
4	326.0	05/05/2021
….	….	….
175433	13.0	28/08/2021
175434	1.0	28/08/2021
175435	6.0	28/08/2021

2) Time-Series analysis

a) Arima Model:

Auto- regressive integrated moving average is popularly used time-series prediction model. Arima model assumes that there is a linear relationship between the variable and time. This model is dynamic and can be used again and again over a period of time .Arima is a category associated with models which unravel a specified census sustained by the previous values, its lags, and also lingered estimate errors. Any 'un-seasonal' statistic which showcases patterns and not belong to the non-linear noise needs to get shaped based on the present model. Arima model is characterized by 3 terms: p, q, d where p is the order of the AR term, q is the order of the MA term and d is the number of differencing which is required to make time series stationary.

Arima was first categorized into two "AR" and "MA". Mathematical formula for the AR and MA models: Auto Regressive (AR) model is one where Y_t depends only on its lags. So, $Y_t$is a function of the lags of Y_t.

$$Y_t = \alpha + \beta 1 Y_{t-1} + \beta 2 Y_{t-2} + \ldots.. + \beta p Y_{t-p} + \epsilon p$$

where, $Y_t{-}1$ is the lag 1 of the series, β_1 is the coefficient of lag 1 that the model estimates, and α is the intercept term, also estimated by the model.

Likewise, a pure Moving Average (MA) model is one where $Y_t$depends only on the lagged forecast errors.

$$Yt = \alpha + \epsilon t + \phi 1 \epsilon t\text{-}1 + \phi 2 \epsilon t\text{-}2 + .. + \phi q \epsilon t\text{-}q$$

where error terms are the errors of the auto-regressive models of the respective lags. The errors ϵ_t and ϵ_{t-1} are the errors from the following equations:

$$Ytt\text{-}1 = \beta 1 Y 1t\text{-}1 + t\text{-}2 + \beta 2 Y 2t\text{-}2t\text{-}3 .. + \beta 0 Y 0 + \epsilon t \ Y = \beta \ Y \ \beta \ Y .. + \beta 0 Y 0 + \epsilon t\text{-}1$$

An ARIMA model can also be said as the one where the time series was different at least once to make it stationary and you combine the AR and the MA terms. So, the equation becomes:

$$Yt = \alpha + \beta 1 Yt\text{-}1 + \beta 2 Yt\text{-}2 + .. + \beta p Yt\text{-}p\text{-}_{\epsilon t} + \phi 1 \epsilon t\text{-}1 + \phi 2 \epsilon t\text{-}2 + .. + \phi q \epsilon t\text{-}q$$

So, Predicted Y_t= Constant + Linear combination Lags of Y (up to p lags) + Linear Combination of Lagged forecast errors (up to q lags) [14].

Objective here, therefore, is to identify the values of p, d and q.

In our Research, we evaluate the performance of the learning models in terms of R-squared (R2) score, Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Square error (RMSE).

b) Facebook Prophet Forecasting Model:

A perishable model comprising three principal elements, namely trend, seasonality, and holidays is practised in the prophet forecasting model [15]. They're consolidated within the subsequent equation:

$$y(t) = g(t) + s(t) + h(t) + \epsilon t$$

- g(t): section-wise linear or logistic growth curve used to model non-periodic fluctuations in the statistics.
- s(t): periodical variations which can be weekly or yearly seasonality.
- h(t): impacts of holidays provided by the user with variable schedules.
- error term approximations for any significant changes not implemented by the model[16] . FB prophet attempts to readjust numerous linear and nonlinear functions of time by utilizing it as an independent variable. Exponential smoothing, as well as a prophet, practise the same strategy of modelling seasonality as a supplement component. In exponential smoothing the prevailing observations are given more weight in forecasting compared to the earlier observations, since exponentially decreasing weights are ascribed due to the emergence of the observations[17] .

c) LSTM Model:

In past few years, deep learning methods such as recurrent neural networks (RNN) have proved to be effective for prediction [18]. However, the limitations of RNN are that there are problems of gradient disappearance and gradient explosion, and since RNNs can only remember part of the sequence, it is much less accurate than short sequences when performing on long sequences, which result in reduced accuracy once the sequence is too long. To overcome these limitations we use LSTM which is a structural type of RNN model [18].To do prediction related tasks, LSTM proved to give most feasible solutions, and their future forecasts are dependent on various highlighted features present in dataset. With LSTM, the data moves through various components which are known as cell states [19].In many issues, LSTM has been quite successful and widely used, such as highway trajectory prediction, stock price prediction and air pollution forecast. LSTM cells are also said to be similar to RNN with hidden units replaceable with memory blocks.

The states of the gates can be represented mathematically as given as below:

ft= σσ(Wxfxt+ Whfht-1 + (1)
Wcfct-1 + bf) it =(Wxixt+ (2)
Whiht-1 + Wcict-1 + bi) (3)
ct= ittanh(Wxcxt+ Whcht-1 +
bc) + ftct-1
ot=(Wxoxt+ Whoht-1 + Wcoct-1 (4)
+ bo) (5)
h_t= oσttanh(c_t)

In the equation given below i,o,f,c represents the input gate, output gate, forget gate, cell and σ is the logistic sigmoid activation function which are of the same size as the hidden vector. W represents the weight matrices where Wci represents the cell input gate matrix. Input gate helps to determines how much information need to be passed based on its significance in the current time step, and it protects the cell from irrelevant inputs. Forget gate helps to decide which information should be deleted that is not relevant from the previous timestamp. The output gate helps to control the flow of information in the rest of the network. LSTM turns off a memory block if it is generating irrelevant outputs [19]. In this paper, we have used LSTM for testing on various optimizers of LSTM such adam, rmsprop, adagrad, sgd, adadelta to check which gives best results.

3) Various Evaluation Parameters

a) R-Squared Score:

R-squared score is a metric of confidence that is easy to compute[20]. It tells us how good the regression line predicts the real values. An R^2 score of 1 symbolises that the regression predictions fit the data faultlessly about its mean. Higher the R^2score better the model performance. R^2 monotonously improves with the increase in the number of variables but never diminishes.

The formula to find the R-squared (R^2) is given by

$$R^2 = \frac{\text{Variance Explained by Model}}{\text{Total Variance}}$$

b) Mean Absolute Error (MAE):

Mean absolute Error is a quantity accustomed to evaluate how close the projections are to the eventual outcomes, the sum of differences between the model forecasts and the true values[21]. It is an undeviating score hence all the discrete differences are balanced evenly within the mean. The MAE can span from 0 to infinity. The lower the worth, the better the model execution[22].

$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n}|y_j - y_j|$$

c) Mean Squared Error (MSE):

Mean Squared Error is the mean of the square of the magnitude of the differences between the actual value and the predicted value[21]. The mean separation between the precise point and the projection is computed, then squared to urge the error. The squaring is extremely important to get rid of negative sign, which provides more importance to more substantial variation. MSE can be calculated as[23]:

$$\text{MSE} = \frac{1}{n}\sum_{j=1}^{n}(y_i - y_i)^2$$

d) Root Mean Square Error (RMSE):

Root Mean Square Error is the variance or the root of the residuals, where residuals tell how distant the regression curve is from the precise data points [21]. It's the measure of how the residuals disperse around the line of best fit. It will easily be deciphered because its units match the output units. Again, this is usually negatively-oriented and an inferior RMSE value improves the model performance [22]. RMSE can be calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(yi - yi)2}$$

4. Methodology

Novel coronavirus formerly known as covid-19 has grabbed special attention from all over the world. To control the spread and rising number of active cases, this study attempts to demonstrate the future forecasting in India in the upcoming 30 days[2]. According to our Research study, it has come to our notice that the LSTM model can be the right choice to predict the future forecast. The data is retrieved from the live Covidwebsite(covid19ind.org)[13].The data set first undergoes data pre-processing after which the clean data is retrieved which then gets stored in the local drive. Now the pre-processed data is analysed for a perfect overview of the data set. The dataset is splitted as 75% as training data and 25% as testing data. The 15% test data, taken from the same dataset, is unrevealed to the model during the training period. By hiding some part of the data it helps in finding out whether the model is overfitting or under-fitting, which are few of the biggest complications while training any model. Facebook Prophet (Facebook prophet is an online open-source for Time-Series Analysis and future forecasting) gives the dates for the next 30 days along with the timestamp to forecast. Finally, the test data set gets appended with the date and time stamp given by Facebook Prophet. As of now, the model is trained on the total number of active cases patterns. LSTM Model has been evaluated based on important metrics such as R-Squared Score, MAE, MSE, and RMSE and reported in the results. The train data set is uploaded with the daily total number of cases so that the model trains better every time it predicts the forecast.

Table 2: Data For Predicting the Future Using Fb Prophet

ds	yhat	yhat_lower	yhat_ upper	y	cutoff
2021-09-05	42.52	-759.45	748.92	8	2021-08-30
2021-09-05	42.52	-643.34	755.84	2039	2021-08-30
2021-09-05	42.52	-679.43	740.14	1	2021-08-30
2021-09-05	42.52	-763.02	729.86	2	2021-08-30
2021-09-05	42.52	-695.26	687.92	1	2021-08-30

(yhat: forecast yhat_lower, yhat_upper: uncertainty interval)

The block diagram in figure.1 depicts the process we will follow for doing future forecasting.

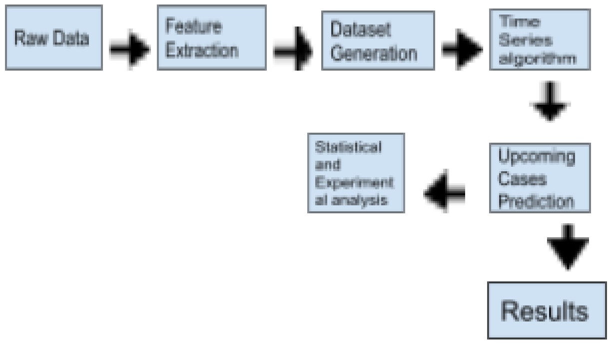

Figure 1: Block Diagram

5. Results

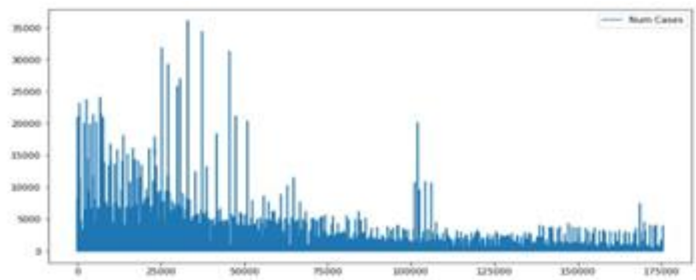

Figure 2: Total Cases in India

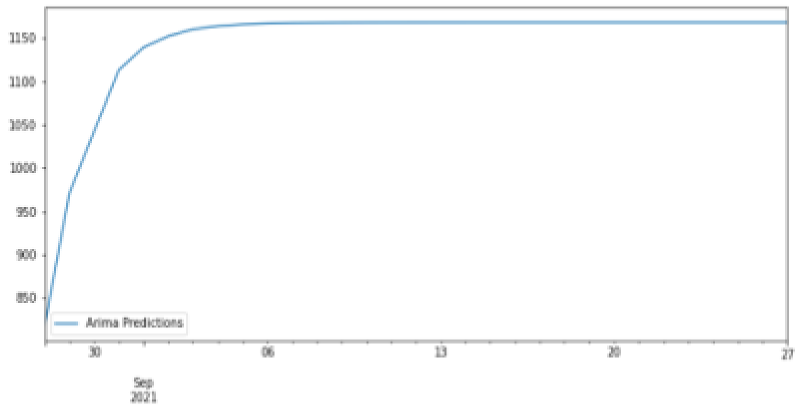

Figure 3: Forecast of the Future cases using Arima Model

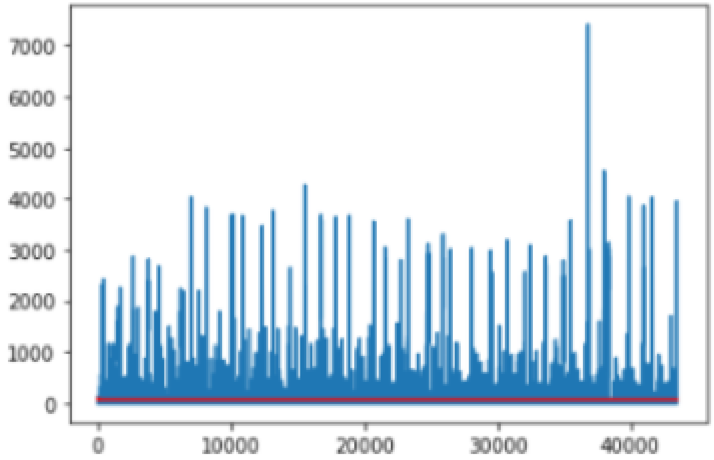

Figure 4: Prediction on testing data using Arima Model

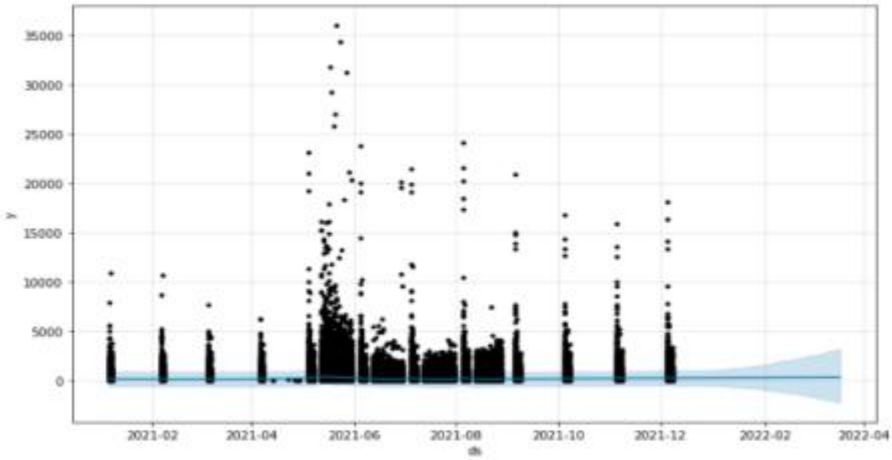

Figure 5: Forecast of future cases using FB-Prophet model

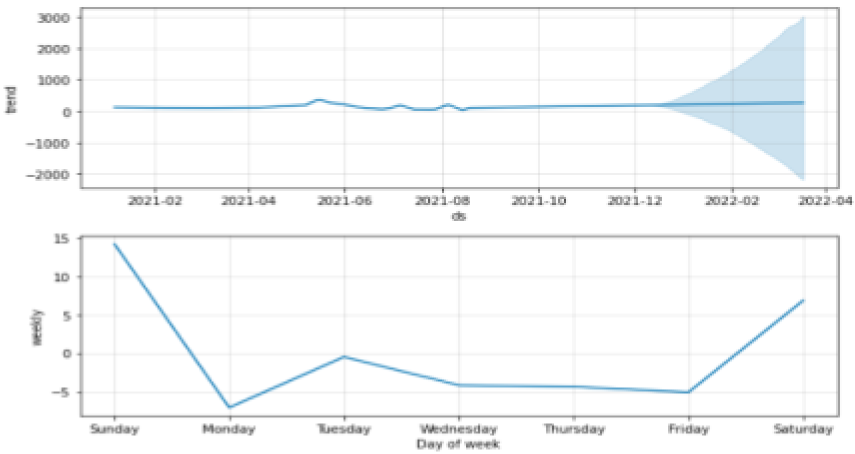

Figure 6: The trend of total cases (weekly) using FB-prophet

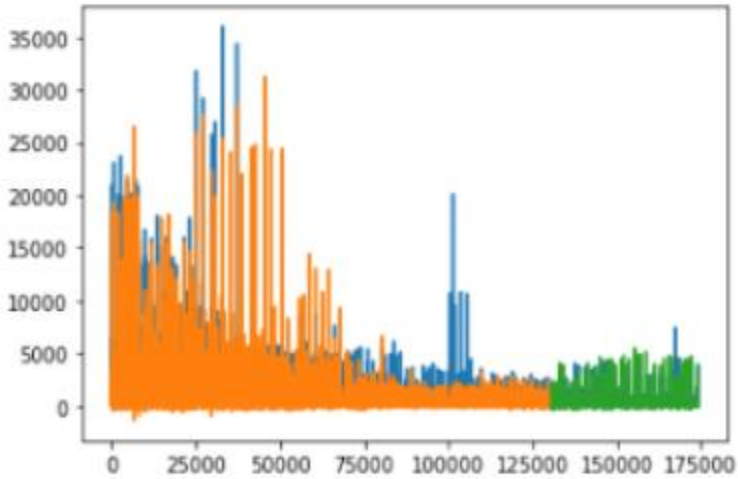

Figure 7: Prediction on dataset using LSTM model with rmsprop as the optimizer

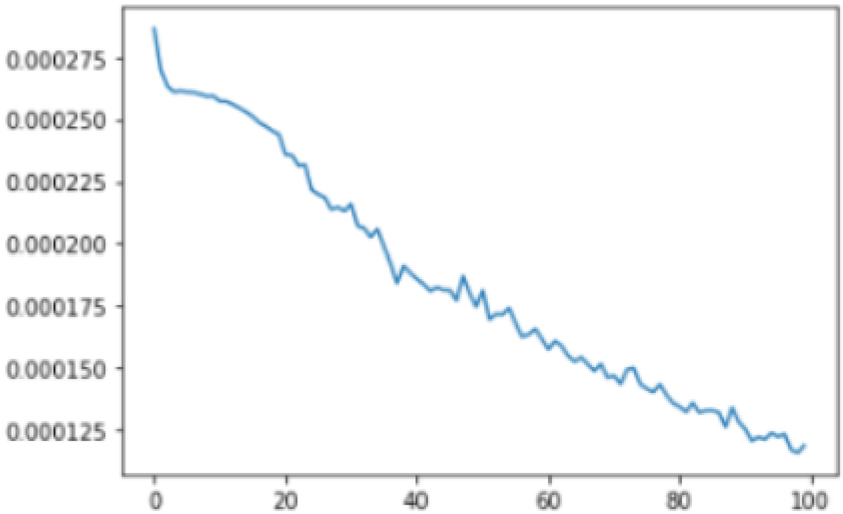

Figure 8: Loss Per epoch using LSTM with rmsprop as the optimizer

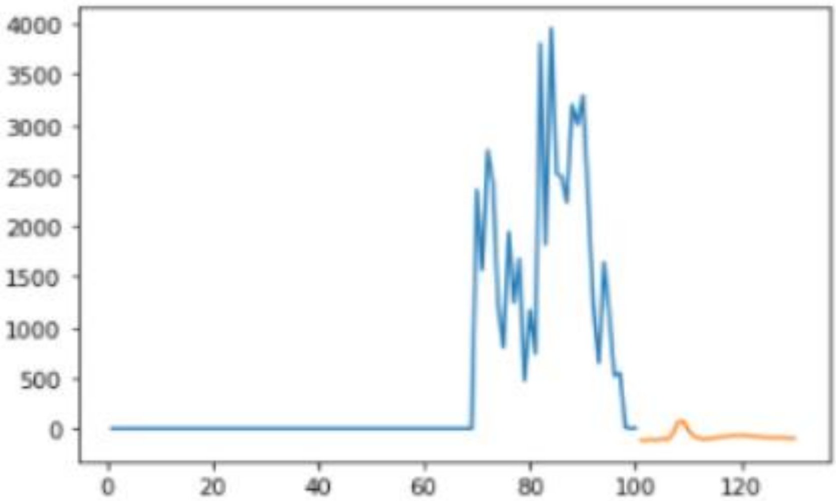

Figure 9: Forecast of future cases using LSTM model with rmsprop as the optimizer

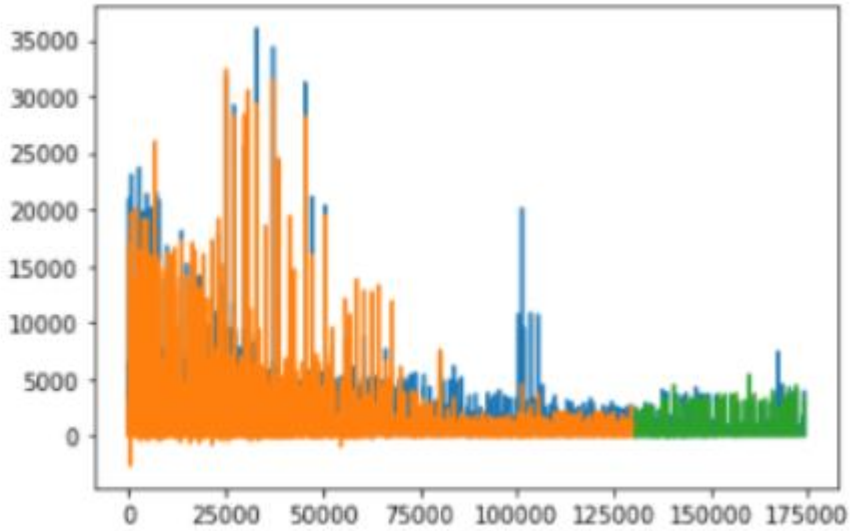

Figure 10: Prediction on dataset using LSTM model with adam as the optimizer

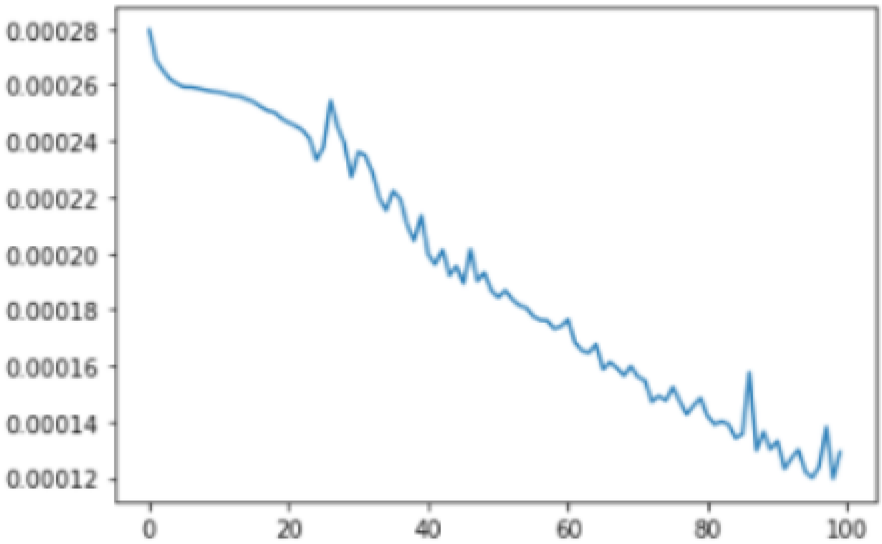

Figure 11: Loss Per epoch using LSTM with adam as the optimizer

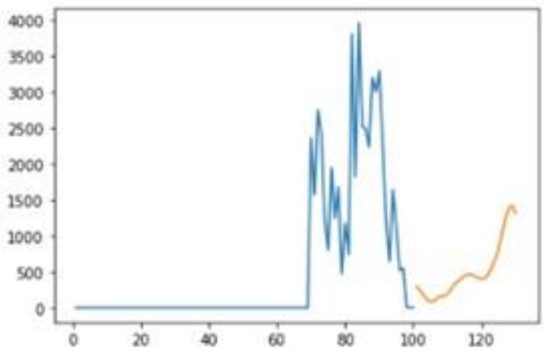

Figure 12: Forecast of the Future cases using LSTM Model with adam as the optimizer

1) **Predictions Using Arima Model**

Table 3: Accuracy Results

Model	R2	MAE	MSE	RMSE
ARIMA	-3.14E-05	102.30	75978.88	275.64

2) **Predictions using Fb-Prophet model**

Table 4: Accuracy Results

horizon	mse	rmse	mae	mape	mdape	coverage
7 days	459334.94	677.74	206.04	12.69	1.5	0.93
8 days	91123.12	301.86	111.22	17.45	4.77	0.97
9 days	48037.59	221.21	88.72	21.02	8.62	0.97
36 days	810158.1	900.08	294.33	8.97	0.96	0.96
37 days	358694.88	598.91	179.93	7.35	0.96	0.98

3) **Predictions using LSTM model Using Different Optimizers**

Table 5: Accuracy Results

Optimizer	MAE	MSE	RMSE
Adam	100.61	77202.84	277.85
Rmsprop	100.96	81803.54	286.01

6. Discussion

The total number of cases concerning COVID-19 in the upcoming few days is the main objective of this study. The data set consists of the total number of cases between May to August as shown in Figure-2.As per objective of this study, we retain only the required information in the data set. Collecting the required amount of information of past four months now, the trend for future cases using Arima model can be observed which is area under curve with the assistance of time series analysis as shown in Figure 3.In this Research, experiments are performed to analyze the effectiveness of our proposed model for predicting the curve of COVID-19 in next few days. The data we collected here is from May 5, 2021 to August 28, 2021. In Figure4. We can see that the red line indicates the predicted cases which are a straight horizontal line whereas blue lines indicate the actual cases on our dataset using the Arima Model. So, from Figure4. we can see that the Arima model does not give good results and hence is not a good fit. In Figure5. Blue horizontal line indicates the cases in the next 30 days whereas back dots indicate actual cases using the Fb-Prophet model which also indicates that fb-prophet is not a good fit as there is great difference in the actual and predicted curve. Fig6. Shows the weekly trend of total cases using fb-Prophet which was initially a straight horizontal line then went upwards and then again down and then again a straight horizontal line. In Figure7. we have taken 75% of the data as training data i.e. orange whereas 25% as the testing data i.e. green using lstm with rmsprop as the optimizer.So we can see that the predictions result of rmsprop optimizer on training and testing part on our dataset is almost equivalent to the actual cases. Figure8. shows the loss per epoch for rmsprop optimizer of lstm which keeps on decreasing with increasing epoch. In Figure9.orange curve indicates the prediction of covid cases in next 30days whereas blue lines indicates cases of previous 100 days with rmsprop optimizer. In Figure10. we have taken 75% of the data as training data i.e. orange whereas 25% as the testing data i.e. green using lstm with adam as the optimizer.So we can see that the predictions result of adam optimizer on training and testing part on our dataset is almost equivalent to the actual cases. Figure11. shows the loss per epoch for adam optimizer of lstm. In Figure12.orange curve indicates the prediction of covid cases in next 30days whereas blue curve indicates cases of previous 100 days with adam optimizer. Table3. shows the accuracy results on various parameters using Arima model. Table4. shows the accuracy results on using fb-prophet model. Table5. shows the accuracy results on various parameters using Lstm models.

7. Conclusion

Our objective for this study is to compare how well can LSTM, ARIMA and PROPHET handle time-series data with no seasonality, has random patterns, and with minimum observations by using COVID-19 cases data. The forecast was done in a time period for both model from May 5, 2021 to August 28, 2022.The limited data on COVID-19 is quite challenging for modeling and prediction. According to the results, it was determined that the LSTM approach has much higher success compared to ARIMA and fbProphet. Later on, forward estimates were made with LSTM with high performance. According to the 2-week prospective estimation study, the total case increase rate is expected to decrease slightly. The study is carried out entirely by considering statistical data and methodologies, the effects of measures taken during the epidemic, compliance with hygiene rules or lockdown are ignored.

Nevertheless, the rate of conformity of the developed prediction model with real data is very satisfactory and offers a strong projection for the near future. However, it is too early to draw a definitive conclusion due to the differences in available data, human behavior. It is observed that the model achievements will increase when the number of days in which the outbreak data is diversified and the data collected is increased. Both ARIMA and PROPHET were found to be fairly inaccurate in forecasting as time progresses. Both models have Mean Forecast Error (MFE) which indicate that both models have positive biases . Positive bias shows that the model is under-forecast, which more often than not, the forecast is less than the actual data. Negative bias shows vice versa, which more often than not, the forecast is more than the actual data. Both model also have negative R2 values. This means that the Sum of Squared Errors of the regression (the distances between actual data points and the regression line) are very far and even greater than the total of Sum of Squared Error. In other words, the regression line going to the way that does not match the harmony of the actual data (going further than the actual data). The greater the negative number of the R2, the greater the distance between the actual data and the predicted data, and the less accurate the predicted data will be. The RMSE score of Adam and Rmsprop optimizer are 277.85 and 286.01 respectively as shown in Figure. So from these results we can see that LSTM model with adam as the optimizer is best suited for prediction as its RMSE score is less as compared to rmsprop optimizer.

References

[1] Punn, Narinder&Sonbhadra, Sanjay & Agarwal, Sonali. (2020). COVID-19 Epidemic Analysis using Machine Learning and Deep Learning Algorithms. 10.1101/2020.04.08.20057679.

[2] Poleneni, Venkatbharat& Rao, Jahnavi&Hidayathulla, Syed. (2021). COVID-19 Prediction using ARIMA Model. 860-865. 10.1109/Confluence51648.2021.9377038.

[3] Wang P, Zheng X, Ai G, Liu D, Zhu B. Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos Solitons Fractals*. 2020;140:110214. doi:10.1016/j.chaos.2020.110214

[4] Ting, Daniel Shu Wei et al. "Digital Technology And COVID-19". *Nature Medicine*, vol 26, no. 4, 2020, pp. 459-461. *Springer Science And Business Media LLC*, https://doi.org/10.1038/s41591-020-0824-5.

[5] Benvenuto, Domenico et al. "Application of the ARIMA model on the COVID-2019 epidemic dataset." *Data in brief* vol. 29 105340. 26 Feb. 2020, doi:10.1016/j.dib.2020.105340

[6] Wang, Peipei et al. "Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics." *Chaos, solitons, and fractals* vol. 139 (2020): 110058. doi:10.1016/j.chaos.2020.110058

[7] Zheng, Nanning et al. "Predicting COVID-19 In China Using Hybrid AI Model". *IEEE Transactions On*

Cybernetics, vol 50, no. 7, 2020, pp. 2891-2904. *Institute Of Electrical And Electronics Engineers (IEEE)*, https://doi.org/10.1109/tcyb.2020.2990162.

[8] J. Contreras, R. Espinola, F. J. Nogales and A. J. Conejo, "ARIMA models to predict next-day electricity prices," in IEEE Transactions on Power Systems, vol. 18, no. 3, pp. 1014-1020, Aug. 2003, doi: 10.1109/TPWRS.2002.804943.

[9] Persio, Luca Di and OleksandrHonchar. “Artificial Neural Networks architectures for stock price prediction: comparisons and applications.” (2016).

[10] Q. Zhang, J.C.K. Lam, V.O.K. Li, Y. Han, 2020. Deep-AIR: A Hybrid CNN-LSTM Framework forFine-Grained Air Pollution Forecast.

[11] Kırbaş, İsmail et al. "Comparative Analysis And Forecasting Of COVID-19 Cases In Various European Countries With ARIMA, NARNN And LSTM Approaches". *Chaos, Solitons & Fractals*, vol 138, 2020, p. 110015. *Elsevier BV*, https://doi.org/10.1016/j.chaos.2020.110015.

[12] Yang Z., Zeng Z., Wang K., Wong S.-S., Liang W., Zanin M., Liu P., Cao X., Gao Z., Mai Z., Liang J., Liu X., Li S., Li Y., Ye F., Guan W., Yang Y., Li F., Luo S., Xie Y., Liu B., Wang Z., Zhang S., Wang Y., Zhong N., He J. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J Thorac Dis. 2020;12:165–174. doi: 10.21037/jtd.2020.02.64. [PMC free article] [PubMed] [CrossRef] [Google Scholar] https://covid19india.org/,5thMay-28th August

[13] “Arima model - complete guide to time series forecasting in python — ml+,” https://www.machinelearningplus.com/time-series/arima-modeltime-series-forecasting-python/,August 2021.

[14] I. Yenidoğan, Işil et al. “Bitcoin Forecasting Using ARIMA and PROPHET.” *2018 3rd International Conference on Computer Science and Engineering (UBMK)* (2018): 621-624.

[15] Taylor SJ, Letham B. 2017 Forecasting at PeerJ 5:e3190v2 scale Preprints https://doi.org/10.7287/peerj.preprints.3190v2

[16] “Time series forecasts using facebook’s prophet,” https: //www.analyticsvidhya.com/blog/2018/05/generate-accurate-forecasfacebook-prophet-python-r/#:~:text=Prophet%20is%20an%20open %20source,of%20custom %20seasonality%20and%20holidays!

[17] Wang, Jinyong, and Ce Zhang. "Software Reliability Prediction Using A Deep Learning Model Based On The RNN Encoder–Decoder". *Reliability Engineering & System Safety*, vol 170, 2018, pp. 73-82. *Elsevier BV*, https://doi.org/10.1016/j.ress.2017.10.019.

[18] Arora P, Kumar H, Panigrahi BK. Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. Chaos Solitons Fractals. 2020 Oct;139:110017. doi: 10.1016/j.chaos.2020.110017. Epub 2020 Jun 17. PMID: 32572310; PMCID: PMC7298499.

[19] Lupón J, Gaggin HK, de Antonio M, Domingo M, Galán A, Zamora E, Vila J, Peñafiel J, Urrutia A, Ferrer E, Vallejo N, Januzzi JL, Bayes-Genis A. Biomarker-assist score for reverse remodeling prediction in heart failure: The ST2-R2 score. Int J Cardiol. 2015 Apr 1;184:337-343. doi: 10.1016/j.ijcard.2015.02.019. Epub 2015 Feb 17. PMID: 25734941.

[20] Willmott, CJ, and K Matsuura. "Advantages Of The Mean Absolute Error (MAE) Over The Root Mean Square Error (RMSE)

[21] In Assessing Average Model Performance". *Climate Research*, vol 30, 2005, pp. 79-82. *Inter-Research Science Center*, https://doi.org/10.3354/cr030079.

[22] Wang, Weijie, and Yanmin Lu. "Analysis Of The Mean Absolute Error (MAE) And The Root Mean Square Error (RMSE) In Assessing Rounding Model". *IOP Conference Series: Materials Science And Engineering*, vol 324, 2018, p. 012049. *IOP Publishing*, https://doi.org/10.1088/1757-899x/324/1/012049.

[23] Witzgall, Hanna &Tarr, A. & Goldstein, J.S.. (2000). ROCKET: A reduced order correlation kernel estimation technique. 1. 406 410 vol.1. 10.1109/ACSSC.2000.910987.