

# Prediction of COVID-19 Wide Spread in India using Time Series Forecasting Techniques

**S. Shitharth**

Vardhaman College of Engineering

**Gouse Baig Mohammad**

Vardhaman College of Engineering

**Kadiyala Ramana**

Annamacharya Institute of Technology and Sciences

**Vidhyacharan Bhaskar** (✉ [meetvidhyacharan@yahoo.com](mailto:meetvidhyacharan@yahoo.com))

San Francisco State University <https://orcid.org/0000-0003-3820-2081>

---

## Research Article

**Keywords:** Machine Learning, Prophet, COVID-19, Time Series, Corona virus, Prediction

**Posted Date:** April 13th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-354432/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# **Abstract**

The assets of some of the enormous wealth are strain out due to the massive infectivity of COVID-19. India is a portion of the global wide spread of COVID- 19 engender by dreadful drastic respiratory syndrome corona virus 2. As of 15th July 2020, the Ministry of Health and Family Welfare has committed a total of 968857 instances, 612768 healings and 24914 demises in the country. Due to the heighten magnitude of number of instances, a professional working in the health departments, some forecasting methods would be required to forecast the number of instances in subsequent days. Due to a towering of uncertainty and lack of crucial information, quality models have shown stubby accuracy for long-term forecast. Among several machine learning models investigated, Time Series Forecasting like Facebook's Prophet showed promising results. In this paper, we have predicted the number of committed, healed, demise instances of COVID-19 in India 60 days' forwards, forecasted the Number of Committed instances, healed instances and demise instances of COVID-19 in India 30 days onwards. Relied on the consequences announced here, and due to the mostly composite nature of the COVID-19 eruption and variation in its deportment, this study suggests machine learning as an efficacious contrivance prototype to the eruption.

## **1. Introduction**

The corona virus COVID-19 broad outspread is the explaining worldwide health tragedy of our accent and the considerable provocation we have looked out on since world war – 2 [1]. Since its exposure in Asia late last year, the virus has proliferation to every continent excluding Antarctica. Instances are stand up daily in Africa, America, and Europe. On 31st December – 2019, the first announced instance in the COVID-19 eruption was communicated in Wuhan, city of China [2]. The earliest instance exterior of China was announced in Thailand on 13th January – 2020. Since then, this happening eruption has now proliferation to more than 180 supplementary countries. World Health Organization states COVID-19 eruption as a Public Health Emergency of International Concern (PHEIC) by on 30th January – 2020 [1].

The initial case of COVID-19 in India was announced on 30th January 2020 [3]. As of 15th July – 2020, the Ministry of Health and Family Welfare (MoHFW) has committed a total of 968857 instances, 612768 heals (including 1 migration) and 8,884 demises in the country [4]. India currently has the enormous number of committed instances in Asia and has the 4th highest number of committed instances in the world with the number of total committed instances breaching the 100,000 marks on 19th May, 200,000 on 3rd June and 900000 on 13 July. India's case fatality hire is relatively underneath at 2.80%, against the world wide 4.7%, as of 6th July 2020. Six cities account for around half of all announced instances in the country – Delhi, Chennai, Pune, Mumbai, Kolkata and Ahmadabad [3]. As of 24th May 2020, Lakshadweep is the only area which has not announced a case [4]. On June 10th, India's recoveries exceeded than the active instances for the first time by reducing 49% of total contamination followed by recovery rate crossing 60% till early July. The United Nations and the World Health Organization have commended India's reaction to the wide roll out as 'comprehensive and robust'; term the Lockdown

restrictions as aggressive but vital for containing the proliferation and building necessary healthcare infrastructure [1].

A contagious infection eruption is the incidence of the infection that is not ordinarily awaited in a distinct faction, geographical area, or time limit. A jumping contagious infection presumes fast proliferation; endangering the health of enormous numbers of people, and thus requires instant measures to intercept the infection at the faction level. COVID-19 is engendered by a noble corona virus which was formerly labeled 2019-nCoV by the World Health Organization [1]. It is the 7th member of the corona virus genealogy, together with MERSnCoV and SARS-nCoV that can proliferation to human beings. Ordinary signs of contamination encompass fever, coughing and breathing troubles. In acute instances, it can produce pneumonia, failure of different organs and finally demise. The incubation interval of COVID-19 is viewed to be between 1 to 14 days. It is infectious before symptoms appear which is why so many large numbers of human beings get contaminated. Infected sufferers can be additionally asymptomatic, means they do not exhibit any indication despite having the virus in their body.

As few of the data accessible to the communal may not be correct, it becomes dense for human beings to realize authentic origin and truthful counsel when they require it. WHO technological danger reporting and communal media party have been functioning closely to track and counter to myths and hearsay via its headquarters in Geneva, its 6 regional offices and its partners because of the elevated request for appropriate and truthful data about 2019-nCoV. WHO is making human beings health particulars and recommendation on the COVID-19 as well as tradition busters that are approachable on its communal media channels which includes Weibo, Instagram, Twitter, Facebook, LinkedIn, and Pinterest etc.

Viral wide lay outs are a consequential threat. COVID-19 is not the earliest, and it won't be the rearmost. But, like not at all before, we are gathering and sharing what we grasp about the virus. Hundreds of investigation teams around the world are merging their endeavor to gather data and evolve solutions. We want to radiate brightness on their efforts and show how machine learning is helping us to:

- • Recognize who is most at prospect
- • Identify sufferer
- • Evolve medicines quickly
- • Detecting subsist medicines that can assist
- • Forecast the proliferation of the infection,
- • Know virus finer
- • Chart where virus arrive from, and
- • Forecast the succeeding broad extend.

Machine learning is a foremost tool in contest to the current wide lay out. If we take this moment to gather information, puddle our grasp, and merge our expertise, we can rescue so many human beings –

both at the moment and in later also. The leading neutral of this article is to project and predict COVID-19 instances, demises, and heals by the Time Series Forecasting.

The arrangement of this work is as follows; 1st section presents COVID-19 as well as describes the gravity of this investigation. Second section presents on previous or interconnected works in predictive modelling of the work. Third section describes our proposed model and methodologies. Finally, the 4th section describes the experimental results, our findings in recent trends and predictive modelling and in fifth section we include the conclusion and future works.

## Related Work

As per several articles, online sources etc. obtainable in literature survey there are learning that focus on the trend analysis and forecasting on COVID-19 wide spread. Some of such works are described below.

Gupta, R. et al. [5] discuss COVID-19 eruption predictions in India. SEIR prototype and regression prototype were used to make predictions based on data gathered from the John Hopkins University repository between the 30th of January and the 30th of March, 2020. The resulting concert of prototypes was approximated using RMSLE, yielding 1.52 for the SEIR model and 1.75 for the regression model. The RMSLE error rate between the SEIR prototype and the regression prototype was 2.01. In addition, the value of  $R_0$ , which represents the infection's proliferation, was determined to be 2.02. In the next two weeks, estimated instances may vary between 5000 and 6000. This research would help the government and doctors in developing their plans over the next two weeks. These prototypes can be tuned for long-term intermission prediction based on short-term interval forecasts.

Singh, R et al. [6] suggested age-structured shattering of communal distancing in the COVID-19 epidemic in India. They research the progression of the COVID-19 epidemic in India using an age-structured SIR model with communal communication matrices derived from surveys and Bayesian imputation. Based on instance details, age dispersal, and communal contact structure, the basic reproductive ratio  $R_0$  and its time-dependent generalisation are computed. The bang of communal distancing steps - office non-attendance, school closure, and lockout - is then investigated, as well as their effectiveness over time. A three-week lockdown is considered inadequate to halt resurgence, and rules of encourage lockdown with periodic composure are proposed instead. Forecasts are based on a decrease in age-structured morbidity and mortality as a result of these forecasts.

Sahasranaman et al. [7] speculate on the network structure of COVID-19 proliferation and the gap in India's monitoring master plan. One of the studies used this tool to assess whether or not unique node clusters were developing. However, the authors only considered travel data junctions to determine which prominent areas are affecting Indian travellers returning to India. In addition, the study suggested using the SIR prototype to determine the rate of Corona Virus proliferation among patients in India. Prior writers have performed inspections on the examining laboratories and facilities.

Tanne, J. H. et al. [8] speculate on the efforts of doctors and frontline health staff. In India, the role of health workers was less emphasised because the corona virus was still in phase two or three of local communication rather than group communication, as opposed to other countries such as Italy, Spain, and the United States. However, it was also announced that the Indian healthcare infrastructure is not very well developed in accordance with WHO guidelines, and that in the event of population proliferation, the Indian government will find it difficult to control the proliferation.

Kucharski, A. J. et al. [9] suggested prior patterns of connectivity and COVID-19 dominance. They estimated how carrying in Wuhan varied between the middle of December, 2019, and February, 2020, by combining a statistical method of extreme SARS-CoV-2 transmission with four datasets from inside and outside Wuhan. It used various mathematical prototypes to detect the spread of the infection, forecast the number of infected patients, discuss the preparedness of each country in dealing with COVID-19 spread, and detect flattening curve trends in various environments.

M. Chinazzi et al. [10] speculated on the impact of travel restrictions on the spread of the 2019 noble corona-virus (COVID-19) outbreak. They use a global meta-population infection contact prototype to estimate the effect of travel restrictions on the infestation's national and international spread. The prototype is calibrated using publicly announced instances from around the world. Modeling results also indicate that sustained 90 percent movement restrictions to and from mainland China have only a slight effect on the disease route unless combined with a 50 percent or greater loss of human contact.

Roosa, K. et al. [11] suggested a real-time forecast of the COVID-19 outbreak in China between February 5th and February 24th, 2020. They used phenomenological prototypes validated during previous eruptions to obtain and analyse short-term forecasts of the increasing number of committed announced instances in Hubei province, the epidemic's epicentre, and for the overall orbit in China. Their findings suggest that the containment plan of action implemented in China was successful, and that the infestation's spread has slowed in recent days.

Grasselli, G. et al. [12] address the use of censorious protection during the COVID-19 eruption in Lombardy, Italy. The COVID-19 Lombardy ICU network's primary goal was to coordinate the deprecatory security response to the eruption. Two top priorities were identified: expanding surge ICU dimensions and implementing containment measures.

F. Petropoulos et al. [13] demonstrate an empirical approach for predicting the continuation of the COVID-19 using a simple but efficient mechanism. Assuming that the data used is accurate and that the future will continue to follow the disease's past trend, the projections indicate a continuing increase in reported COVID-19 cases with significant associated uncertainty. The risks are far from symmetric, as underestimating its spread like a pandemic and failing to do enough to contain it is far more serious than overspending and being overly cautious when it is not needed. They depict the timeline of a live forecasting exercise with massive potential consequences for preparation and decision making, as well as providing realistic predictions for COVID-19 verified events. In this case, they used univariate time series prototypes, which imply that the data is accurate and that previous trends as well as precautionary

assessments will continue to be applied. Important, compatible forecast errors should be correlated with shifts in perceived trends and the need for additional steps and interventions in the case of negatively biased forecasts.

S. Makridakis et al. [14] announce the results of a forecasting competition that offers information to aid in such decision-making. Seven experts forecasted up to 1001 series for six to eighteen time horizons using each of the 24 methods. The competition results are presented in this work, which aims to provide empirical evidence regarding differences discovered among the various extrapolative (time series) methods used in the competition.

Makridakis S. et al. [15] cover all aspects of M4, including its structure and management, the demonstration of its results, the top-performing procedures inclusive and by groups, its major findings and their recommendations, and the computational requirements of the various procedures. Finally, it summarises its key conclusions and expresses the hope that its series will serve as a testing ground for the creation of new procedures and the advancement of prediction practise, while also outlining some potential directions for the field.

Petropoulos F et al. [19] proposed using perception to improve the selection of a forecasting template. They compared the execution of a judgmental prototype selection to a standard procedure based on information standards. They also studied the effectiveness of a judgmental prototype-creation process, in which specialists were asked to reach a conclusion on the nature of structural components of a time series rather than explicitly selecting a prototype from a collection of options. Their behavioural survey drew on data from nearly 700 sources, including forecasting practitioners. According to the results of their assessment, selecting prototypes results in efficiency that is equal, if not better, than procedure selection. Furthermore, judgmental prototype selection aids in avoiding the worst prototypes that are often collated for algorithmic selection. Finally, a clear mixture of statistical and judgmental elections, as well as judgmental aggregation, outperforms all statistical and approximated elections.

To accomplish this, they devised a detectable experiment and tested the effectiveness of two judgmental mythologies for selecting prototypes, namely simple prototype election and prototype-creation. The final one was based on an approximation of time series characteristics detection. They compared the performance of these procedures to that of a statistical benchmark based on knowledge parameters. The development of a judgmental prototype outperforms both the creation of a statistical prototype and the creation of a judgmental prototype. The equal-weight mix of statistical and judgmental election resulted in significant execution changes over statistical election. The best execution of any mythologies they considered resulted from judicious aggregation. Finally, an exciting result is that humans outperform statistics in preventing the worst prototype. According to the findings of this study, businesses should regard judgmental election forecasting as a supplement to statistical prototype election. Furthermore, they believe that limiting the judgmental aggregation of a few experts to the most relevant items is a trade-off between capital and performance enhancement that businesses should be willing to accept. Forecasting bear systems with simple graphical interfaces and judgmental recognition of time series

characteristics, on the other hand, are needed for the efficient administration of do-it-yourself (DIY) forecasting.

Taylor JW – 2003 look over a latest damped multiplicative mode point of view. An observed survey, utilizing the monthly time series from the M3-Competition, gave uplifting outcomes for the latest method at a range of forecast skylines, when contrasted to the confirmed exponential smoothing approaches. In this work, they have initiated a latest damped exponential smoothing approach. The approach goes along with the multiplicative fashion formulation of Pegels (1969) but contains an extra variable to dampen the projected fashion. They used the 1,428 monthly time series from the M3- Competition to contrast the approach to the quality Pegels approach and the accepted exponential smoothing approaches. The accomplishment of the quality Pegels approach was alike to that of the quality Holt approach [18]. This is a compulsive outcome as there has been no previous empirical survey contrasting the post-sample predicting correctness of the quality Pegels approach with that of other exponential smoothing approaches. It indicates that the acceptance of a multiplicative fashion is not as treacherous as might have been contemplated. They get that the damped Pegels approach competently defeated the quality Pegels approach at all forecast skylines. Moreover, the latest damped variety of the approach also slightly defeated the popular damped Holt approach. The standardized Holt formulation is similar to damped Holt's either than that the  $\varphi$  variable is entitled to grab values greater than one. This happened for 203 of the 1,428 sequences.

An improved value greater than one for the generalized Holt's  $\varphi$  variable indicates that damped Holt's will not be capable to satisfactorily predict the fashion in these sequences and that other predicting approaches may be preferred. They go into whether the multiplicative fashion articulation of the quality Pegels and damped Pegels approaches is accepted for these sequences [19]. They contrasted the correctness of these approaches to the settled exponential smoothing approaches for the subset of 203 sequences. The quality Pegels approach exceeded quality Holt's according to the Symmetric Mean APE summary error quantity but not according to the Median APE.

Out of all the 7 approaches contemplated, the foremost outcomes were attained for both error quantities utilizing the damped Pegels approach. This encourages that the damped Pegels approach could at least be applicable as a substitute to the well liked and successful damped Holt approach for sequences for which the latter seems inappropriate [20]. In perspective of this, there would seem to be powerful call in as well as the damped 16 Pegels approach as a candidate in automated approach election methods, such as that of Hyndman et al. (2002). In summary, they feel that the outcomes for the 1,428 sequences and for the subset of 203 suggest that the latest damped Pegels approach is a appreciable development on the quality Pegels approach, and that it is a probably practical unwonted to the settled exponential settled approaches.

## Proposed Approach

The Proposed Model is shown in Figure 1. In the suggested approach, we apply Time Series Forecasting Techniques like Prophet to build a model which comes up with 60 days ahead predictions of Committed, Death and Recovered instances of COVID-19, India. System is assembled in following phases such as collection of information sets, pre-processing phases and applying Time Series Forecasting Techniques in Machine Learning like Facebook's Prophet Model because it supplies us with the capability to make time series forecasts with better accuracy by utilizing simple innate parameters.

### 3.1 Collection of Data Sets

Collecting information or data permits us to express a record of formerly events so that we can utilize analysis of data to detect patterns that are recurring in nature. From those patterns, we create predictive prototypes using machine learning mythologies that deem for trends and forecast hereafter changes. Collection or gathering of data is the procedure of collecting and quantifying information from innumerable several origins. In order to utilize the data, we collect to expand practical forecast solutions [21], it must be gathered and deposited in a manner that creates sense for the prediction problem at hand. The Data for the forecasting of COVID-19 in India is collected from covid19india.org – Corona virus Eruption in India and Ministry of Health and Family Welfare (MoHFW), Government of India and the data sets are prepared for processing. The data used in this paper is time Series Data of COVID-19 instances in India having the attributes like date, state, committed instances, recovered, death etc.

### 3.2 Pre-processing Step

In machine learning, this phase is a very crucial phase. Pre-processing of Covid-19 instances data consists of transforming Raw Covid-19 Data Day wise into understandable format by grouping the Covid-19 instances according to State wise so that we can explore the Data and derive some Observations from that Data. But for Fitting into Prophet Model and for Predictions, Data should be in Time Series.

### 3.3 Splitting the Dataset into Train Sets and Test Sets

Generally, when we divide the data set into a train set and test set, most of the data is utilized for training, and a little part of the data is utilized for testing. Investigation Services randomly samples the data to assist make sure that the testing and training sets are alike. The data or information we utilize is typically broken into train and test data. The train data accommodates a well-known outcome and the prototype master on these data in order to be generalized to other data subsequently. We have the test data in order to test the prototype's forecast on this test set.

While doing Prediction, we need to select training data in order to train the forecasting model and test data, upon which the forecasting model will be utilized to check the accuracy. Instead of implementing codes to split the dataset into training and test set, to make it easy for everyone, we just simply chose the time period from 30th January 2020 (Date of First Committed Covid-19 Case in India) to till now i.e. 15th July 2020 and we fitted into the Prophet Model.

### 3.4 Machine Learning Models

In this paper we have applied Time Series Forecasting technique such as Facebook's Prophet Model for 60 days onwards predictions of Committed, Death and healed instances of COVID-19 India which are as follows.

### 3.4.1 Prophet Model from Facebook

It is a step-by-step procedure to create predicting models for data that relies on time series. It supplies us with the potential to create time series forecasts with accurate accuracy by utilizing easy innate parameters. Unlike the ancient methods, It seeks to match a.k.a. 'curve fitting' additive regression models.

At its key, the Prophet course of action is an additive regression prototype with 4 major constituents: A piecewise straightaway or logistic magnification curve trend. Prophet instinctively determines changes in trends by electing change positions from the data. Prophet is an open-source library produced by Facebook that is relied on decomposable (fashion, seasonality and day off) prototypes. It supplies us with the capability to create time series forecasts with acceptable precision using simple intuitive variables and has fund for including impact of custom seasonality and holidays [22].

The Prophet Prototype work relied on the following tread;

1. Construct a prototype relied on the training data. Then, we can construct the predicting prototype relied on the training data with the Prophet.
2. Utilize the prototype to predict for the test data period.
3. Contrast one and the other forecasted and genuine.
4. Square them.
5. Figure out the average (mean value).
6. Root it.

It is really versatile when the data fed to this Algorithm is the best thing about this method. You will have NAs because you don't need both dates and times in sync.

It works correctly and equally by default, without setting any parameters explicitly. And if you know the domain, you can customise all of the parameters to further optimise the model, but it's easy to understand the parameters.

We utilize a decomposable time series model with 3 important prototype constituents: inclination, seasonality, and day off. They are merged in the following calculation:  $x(t) = p(t) + q(t) + r(t) + \epsilon$  (1)

- $p(t)$  models trend, which reports long-term grow or enlarge in the data. Prophet assimilates 2 trend prototypes, a drenching growth prototype and a piecewise linear prototype, depending on the type of forecasting issue.
- $q(t)$  prototype seasonality with Fourier series, which reports how information is overripe by seasonal components such as the time of the year

- $r(t)$  models the effects of holidays or enormous events that impact business time series
- $\epsilon$  describes an irreducible error phrase

Utilizing hour as a regressor, prophet is demanding to befitting different linear and nonlinear functions of time as constituents. The seasonal model of an additive is the same method as the Holt Winter methodology for exponential smoothing. In effect, we are framing the predicting issue as a curve-fitting exercise rather than examining explicitly at the time dependent vulnerability of each and every inspection inside the time series.

## Experimental Result And Performance Analysis

### *4.1 Data Exploration*

Exploration of data is a procedure alike to prime analysis of data, where an analyst utilizes vision observation to realize what is in a dataset and the features of the data, rather than across conventional data administration systems [23]. Data exploration can assist cut down the enormous data set to a practicable size where we can aim our attempts on examining the most pertinent information. It is one and the other a science as well as an art. There is the science of working into and altering the information [24].

From the Data Exploration, we came to know that, as of, 08:00 IST, 15th July 2020, In India, there are 968857 Covid-19 committed instances. Out of which, 612768 are either Cured or Discharged or Migrated and 24914 are Dead. Till now, 35 States/ Union Territories in India are affected and Number of Active Instances is 331175. Figure 3, 4 and 5 shows the Present COVID-19 Condition of India.

Humans do not have immunity to this virus, allowing its easy and rapid spread among human populations through contact with an infected person. SARS-CoV-2 is more transmissible than SARS-CoV. The two possible reasons could be (i) the viral load (quantity of virus) tends to be relatively higher in COVID-19-positive patients, especially in the nose and throat immediately after they develop symptoms, and (ii) the binding affinity of SARSCoV-2 to host cell receptors is higher than that of SARS-CoV.

### *4.2 Prediction Results using Prophet Model*

Generating 60 days ahead forecast of COVID-19 instances using Prophet, with 95% prediction interval by creating a base model with no tweaking of seasonality-related parameters and additional regressor [25,26]. The load to Prophet is every time a data frame with 2 columns: ds and y. The ds (date stamp) column should be of a format anticipated by Pandas, preferably YYYY-MM-DD for a date or YYYY-MM-DD HH:MM:SS for a timestamp. y column must be numeric, and constitutes the measurement we wish to predict. Here y refers to Covid-19 instances [27,28]. The foretell technique will allocate each row in later a forecasted utility which it tags as "yhat". If you pass in ancient dates, it will supply an in-sample fit. The forecast object here is a new data frame that includes a column yhat with the forecast, as well as columns for constituents and unpredictability intervals.

Here, The Following attributes are specified by Prophet Model for Prediction.

ds = Specific Dates of Covid-19 Instances

yhat = Number of Covid-19 instances that is Predicted by Prophet Model

yhat\_lower = Lower Limit of No. of Covid-19 instances that is Predicted by Prophet Model

yhat\_upper = Upper Limit of No. of Covid-19 instances that is Predicted by Prophet Model

#### *4.2.1 Committed Instances Prediction up to 60 Days*

60 days ahead forecast of COVID-19 Committed instances using Prophetic shown in Table 1 and the graphical representation is shown in Figure 6.

Table 1: Covid-19 Committed Instances Prediction from 16 July 2020- 13 September 2020

S.No.	ds	yhat	yhat_lower	yhat_upper
1	16-07-2020	907729.3	876242.6	938327.5
2	17-07-2020	926183.2	895940	953512.5
3	18-07-2020	944972.4	916841.5	975007.7
4	19-07-2020	963818.3	934143.9	993074.5
5	20-07-2020	982336	953866.3	1011730
6	21-07-2020	1001032	970155.1	1030980
7	22-07-2020	1019618	989167.9	1050080
8	23-07-2020	1034851	1003048	1066373
9	24-07-2020	1053305	1023225	1087382
10	25-07-2020	1072094	1040423	1105414
11	26-07-2020	1090940	1057969	1125705
12	27-07-2020	1109457	1076021	1143930
13	28-07-2020	1128153	1092142	1166741
14	29-07-2020	1146739	1108998	1184182
15	30-07-2020	1161972	1124326	1199832
16	31-07-2020	1180426	1142034	1217512
17	01-08-2020	1199215	1162248	1241080
18	02-08-2020	1218061	1172592	1264696
19	03-08-2020	1236579	1187241	1282396
20	04-08-2020	1255275	1203901	1304951
21	05-08-2020	1273861	1223824	1323386
22	06-08-2020	1289094	1233379	1341186
23	07-08-2020	1307548	1250327	1362841
24	08-08-2020	1326337	1271105	1382590
25	09-08-2020	1345183	1282065	1406210
26	10-08-2020	1363700	1295920	1424895
27	11-08-2020	1382396	1310297	1455543
28	12-08-2020	1400982	1327739	1471982

29	13-08-2020	1416215	1332725	1491820
30	14-08-2020	1434669	1353936	1504869
31	15-08-2020	1453458	1367301	1530624
32	16-08-2020	1472304	1381635	1553249
33	17-08-2020	1490822	1396180	1579058
34	18-08-2020	1509518	1409904	1599027
35	19-08-2020	1528104	1430333	1614559
36	20-08-2020	1543337	1433750	1643873
37	21-08-2020	1561791	1455384	1659072
38	22-08-2020	1580580	1468313	1686859
39	23-08-2020	1599426	1484940	1707585
40	24-08-2020	1617943	1497841	1732854
41	25-08-2020	1636639	1515539	1747307
42	26-08-2020	1655225	1528751	1781150
43	27-08-2020	1670458	1541850	1794393
44	28-08-2020	1688912	1544914	1818156
45	29-08-2020	1707701	1561719	1843265
46	30-08-2020	1726547	1577545	1858074
47	31-08-2020	1745065	1590651	1884790
48	01-09-2020	1763761	1604471	1913275
49	02-09-2020	1782347	1620467	1937216
50	03-09-2020	1797580	1635074	1953297
51	04-09-2020	1816033	1646536	1971998
52	05-09-2020	1834823	1658111	1999146
53	06-09-2020	1853669	1682157	2015943
54	07-09-2020	1872186	1693829	2046867
55	08-09-2020	1890882	1706916	2069866
56	09-09-2020	1909468	1726333	2094189
57	10-09-2020	1924701	1731204	2115282

58	11-09-2020	1943155	1743954	2137696
59	12-09-2020	1961944	1758631	2156359
60	13-09-2020	1980790	1771342	2179629

Here in Figure 6, The Black Dots describes the actual Covid-19 announced Instances in India. The Blue Straight Line represents Predicted Covid-19 committed instances in India by Prophet Model. The Blue Shaded Region in the Graph shows the Upper and lower limit of Predicted Covid-19 committed instances in India by Prophet Model.

#### 4.2.2 Death Instances Prediction up to 60 Days

60 days ahead forecast of COVID-19 Death instances using Prophet is shown in Table 2 and the graphical representation is shown in Figure 7.

**Table 2:** Covid-19 Death Instances Prediction from 16 July 2020- 13 September 2020

S.No.	ds	yhat	yhat_lower	yhat_upper
1	16-07-2020	24965.117339	24670.239291	25277.036350
2	17-07-2020	25420.959330	25127.552695	25727.258943
3	18-07-2020	25877.995894	25582.507197	26198.853987
4	19-07-2020	26329.779676	26007.316998	26644.408259
5	20-07-2020	26783.893825	26478.334679	27094.579539
6	21-07-2020	27316.037165	26974.676104	27655.628432
7	22-07-2020	27773.339298	27387.207216	28137.381796
8	23-07-2020	28208.055356	27800.471026	28606.112560
9	24-07-2020	28663.897347	28201.666836	29039.808100
10	25-07-2020	29120.933910	28614.631377	29546.403542
11	26-07-2020	29572.717693	28995.876727	30086.717038
12	27-07-2020	30026.831841	29396.789652	30617.770044
13	28-07-2020	30558.975182	29862.959164	31186.691236
14	29-07-2020	31016.277314	30223.869040	31724.533390
15	30-07-2020	31450.993372	30578.447743	32201.676799
16	31-07-2020	31906.835363	30963.936554	32642.379239
17	01-08-2020	32363.871927	31417.058991	33237.129758
18	02-08-2020	32815.655709	31764.524918	33725.507701
19	03-08-2020	33269.769858	32117.977116	34298.435284
20	04-08-2020	33801.913198	32555.172364	34762.952494
21	05-08-2020	34259.215330	32918.675072	35368.927746
22	06-08-2020	34693.931389	33338.691663	35952.381996
23	07-08-2020	35149.773379	33598.431642	36495.911246
24	08-08-2020	35606.809943	34059.635582	36905.294876
25	09-08-2020	36058.593725	34435.842355	37551.657883
26	10-08-2020	36512.707874	34805.507183	38054.085879
27	11-08-2020	37044.851215	35185.042975	38696.966303
28	12-08-2020	37502.153347	35664.516963	39212.311606

29	13-08-2020	37936.869405	35930.039682	39724.193226
30	14-08-2020	38392.711396	36254.125782	40305.912233
31	15-08-2020	38849.747959	36685.087282	40818.331828
32	16-08-2020	39301.531742	37001.400935	41444.831683
33	17-08-2020	39755.645891	37507.343009	41978.291077
34	18-08-2020	40287.789231	37934.516953	42572.743216
35	19-08-2020	40745.091363	38246.783678	43121.762842
36	20-08-2020	41179.807421	38534.337420	43649.361773
37	21-08-2020	41635.649412	38985.017765	44284.446296
38	22-08-2020	42092.685976	39278.455651	44847.166422
39	23-08-2020	42544.469758	39588.523204	45337.275914
40	24-08-2020	42998.583907	39891.400445	45916.313186
41	25-08-2020	43530.727247	40326.173396	46528.914481
42	26-08-2020	43988.029380	40724.203184	47078.195530
43	27-08-2020	44422.745438	40854.621602	47702.026710
44	28-08-2020	44878.587429	41160.063993	48230.333729
45	29-08-2020	45335.623992	41655.106087	48767.978812
46	30-08-2020	45787.407775	42096.063807	49295.815245
47	31-08-2020	46241.521923	42294.196013	49928.176152
48	01-09-2020	46773.665264	42721.086112	50589.989289
49	02-09-2020	47230.967396	43154.288457	51080.266615
50	03-09-2020	47665.683454	43458.286209	51654.248036
51	04-09-2020	48121.525445	43723.300299	52268.456890
52	05-09-2020	48578.562009	44108.571324	52786.061654
53	06-09-2020	49030.345791	44300.727216	53325.420742
54	07-09-2020	49484.459940	44684.796556	53837.661882
55	08-09-2020	50016.603280	45116.962939	54506.304654
56	09-09-2020	50473.905412	45509.257886	55121.774334
57	10-09-2020	50908.621471	45821.906440	55679.370149

58	11-09-2020	51364.463461	46141.480994	56293.982542
59	12-09-2020	51821.500025	46336.858804	56872.965286
60	13-09-2020	52273.283807	46767.209356	57447.857735

Here in Figure 7, the black dots show the actual Covid-19 demise instances in India. Blue straight line represents predicted Covid-19 demise instances in India by prophet model. The blue shaded region in the graph shows the upper and lower limit of predicted Covid-19 death instances in India by prophet model.

#### *4.2.3 Recovered Instances Prediction up to 60 Days*

60 days ahead forecast of COVID-19 recovered instances using prophet is shown in Table 3 and the graphical representation is shown in Figure 8.

**Table 3:** Covid-19 Recovered Instances Prediction from 16 July 2020- 13 September 2020

S.No.	ds	yhat	yhat_lower	yhat_upper
1	16-07-2020	573099.1	551998.1	593499
2	17-07-2020	586025.8	565012.5	606030.8
3	18-07-2020	599083.7	579185.1	621489
4	19-07-2020	611873.3	589556.8	631473.6
5	20-07-2020	624789.8	606476.7	646311.4
6	21-07-2020	637825.7	617336.4	658961.2
7	22-07-2020	650369.7	627864.5	671722.7
8	23-07-2020	661510	637474	684627.9
9	24-07-2020	674436.7	648525.5	695954.2
10	25-07-2020	687494.6	665388	710571.1
11	26-07-2020	700284.2	676644.8	725268.9
12	27-07-2020	713200.7	689073.5	737303.7
13	28-07-2020	726236.6	697686.3	753194.4
14	29-07-2020	738780.6	711467.5	764847.7
15	30-07-2020	749920.9	720805.9	777360.1
16	31-07-2020	762847.7	734063.6	794764.8
17	01-08-2020	775905.5	744707.2	808289.5
18	02-08-2020	788695.1	753147.9	823174.7
19	03-08-2020	801611.6	768682.8	836641.6
20	04-08-2020	814647.5	778467.5	852639.5
21	05-08-2020	827191.5	783635.9	867857.2
22	06-08-2020	838331.8	797074	879762
23	07-08-2020	851258.6	805903.4	895837
24	08-08-2020	864316.4	817051.3	907576.2
25	09-08-2020	877106	828109	923983.2
26	10-08-2020	890022.5	836518.3	939525.6
27	11-08-2020	903058.4	851518.7	955028.7
28	12-08-2020	915602.5	859883.8	968670.3

29	13-08-2020	926742.7	869555	983020.7
30	14-08-2020	939669.5	879744.8	999431.3
31	15-08-2020	952727.4	889306.4	1013551
32	16-08-2020	965516.9	895794	1031148
33	17-08-2020	978433.4	909057.2	1042243
34	18-08-2020	991469.3	917933.1	1061961
35	19-08-2020	1004013	930218.7	1077979
36	20-08-2020	1015154	940576.1	1093847
37	21-08-2020	1028080	952732.8	1104533
38	22-08-2020	1041138	962649.3	1120889
39	23-08-2020	1053928	969178.3	1135864
40	24-08-2020	1066844	980030.1	1154096
41	25-08-2020	1079880	991433.7	1172589
42	26-08-2020	1092424	999722.4	1183282
43	27-08-2020	1103565	1009061	1199648
44	28-08-2020	1116491	1011669	1211837
45	29-08-2020	1129549	1031164	1230202
46	30-08-2020	1142339	1040362	1250211
47	31-08-2020	1155255	1049823	1264128
48	01-09-2020	1168291	1055274	1278509
49	02-09-2020	1180835	1067086	1294023
50	03-09-2020	1191975	1079225	1311179
51	04-09-2020	1204902	1090889	1327839
52	05-09-2020	1217960	1097417	1341240
53	06-09-2020	1230750	1108997	1358896
54	07-09-2020	1243666	1120774	1377759
55	08-09-2020	1256702	1121754	1391683
56	09-09-2020	1269246	1131400	1410157
57	10-09-2020	1280386	1144642	1423466

58	11-09-2020	1293313	1146908	1441285
59	12-09-2020	1306371	1160454	1467942
60	13-09-2020	1319161	1172981	1472007

Here in Figure 8, The black dots represent the actual Covid-19 recovered instances in India. The blue Straight line represents predicted Covid-19 recovered instances in India by prophet model. The blue shaded region in the graph shows the upper and lower limit of predicted Covid-19 recovered instances in India by prophet Model.

## Conclusion And Future Work

The corona virus infection carries on to proliferation across world by following a direction that is very arduous to forecast. The health, humanitarian and socio- economic strategies acquired by countries will control the tempo and power of the healing. Governments over the world have foisted limitations, exceptionally acute in some countries such as India, to moderate the proliferation of the virus. Public health specialists have steered the calls for durable measures. However, various factors which influence any disease. The main thing is the capacity and the decision taken at the right time, among which population; geographical conditions; not considered the social distance; lack of diagnostic facilities and lack of doctors or clinical personnel.

Due to an elevated quantity of unpredictability and absence of vital data, quality models have shown squat accuracy for long-term forecast. Among a board range of machine learning models investigated, Time Series Forecasting Techniques like Facebook's Prophet showed encouraging consequences. In this proposed work, we have forecasted the Number of Committed instances, healed instances and demise instances of COVID-19 in India 30 days onwards. Relied on the consequences announced here, and due to the mostly composite nature of the COVID-19 eruption and variation in its deportment, this study suggests machine learning as an efficacious contrivance prototype to the eruption. The consequence of preventing measures like social isolation and lockdown has also been observed which manifests that by these inhibitory measures, proliferation of the virus can be turned down notably.

## Declarations

1) Funding (information that explains whether and by whom the research was supported)

None

2) Conflicts of interest/Competing interests (include appropriate disclosures)

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

3) Availability of data and material (data transparency)

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to containing information that could compromise the privacy of research participants.

#### 4) Code availability (software application or custom code)

The code that supports the findings of this study are available on request from the corresponding author. The code is not publicly available due to containing information that could compromise the privacy of research participants.

#### 5) Authors' contributions (optional: please review the submission guidelines from the journal whether statements are mandatory)

**S.Shitharth**: Conceptualization, Methodology, Software, Visualization **Gouse Baig Mohammad**: Data curation, Writing - original draft, Data Analysis, Investigation. **Kadiyala Ramana**: Software, Validation, Editing. **Vidya Charan Bhaskar**: Supervision, Writing - review & editing.

## References

1. World Health Organization (2020). Corona virus infection (COVID-19) Wide spread, WHO. Accessed from <https://www.who.int/emergencies/infections/novel-coronavirus-2019>
2. John Hopkins University (2020). Novel Corona Virus (COVID-19) Instances, provided by JHU CSSE. Accessed from <https://github.com/CSSEGISandData/COVID-19>
3. The Ministry of Health and Family Welfare (MoHFW), Government of India (2020) Corona virus infection (COVID-19) Wide spread. Accessed from <https://www.mohfw.gov.in/> on 13th June 2020.
4. org: Corona virus Eruption in India. Government of India (2020). Accessed from <https://www.covid19india.org/> on 13th June 2020.
5. Manikandan, T., S. Shitharth, C. Senthilkumar, C. Sebastinalbina, and N. Kamaraj, Removal of selective black hole attack in MANET by AODV protocol, *The International Journal of Innovative Research In Science, Engineering And Technology*, vol. 3, no. 3, pp. 2372-2377, 2014.
6. Pandey, Gaurav, Poonam Chaudhary, Rajan Gupta, and Saibal Pal, SEIR and Regression Model based COVID-19 outbreak predictions in India, *arXiv preprint arXiv:2004.00958*, 2020.
7. Devi, B. Thirumaleshwari, S. Shitharth, and M. A. Jabbar, An Appraisal over Intrusion Detection systems in cloud computing security attacks, *In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 722-727. IEEE, 2020.
8. Selvarajan, Shitharth, Masood Shaik, Sirajudeen Ameerjohn, and Sangeetha Kannan, Mining of intrusion attack in SCADA network using clustering and genetically seeded flora-based optimal

classification algorithm, *IET Information Security*, vol. 14, no. 1, pp. 1-11, 2019.

9. Singh, Rajesh, and Ronojoy Adhikari, Age-structured impact of social distancing on the COVID-19 epidemic in India, *arXiv preprint*, arXiv:2003.12055, 2020.
10. Sahasranaman, Anand, and Nishanth Kumar, Network structure of COVID-19 spread and the lacuna in India's testing strategy, Available at SSRN 3558548, 2020.
11. Shitharth, S., and D. Prince Winston, A New Probabilistic Relavancy Classification (PRC) based Intrusion Detection System (IDS) for SCADA network, *Journal of Electrical Engineering*, vol. 16, no. 3, pp. 278-288, 2016.
12. Shitharth, S., and D. Prince Winston, A novel IDS technique to detect DDoS and sniffers in smart grid, *In 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, pp. 1-6, IEEE, 2016.
13. Tanne, Janice Hopkins, Erika Hayasaki, Mark Zastrow, Priyanka Pulla, Paul Smith, and Acer Garcia Rada, Covid-19: how doctors and healthcare systems are tackling coronavirus worldwide, *BMJ* 368, 2020.
14. Kucharski, Adam J., Timothy W. Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M. Eggo et al, Early dynamics of transmission and control of COVID-19: a mathematical modelling study, *The lancet infectious diseases*, vol. 20, no. 5, pp. 553-558, 2020.
15. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak, *Science*, in press, 368(6489), pp. 395–400, 2020.
16. Shitharth, S, An enhanced optimization based algorithm for intrusion detection in SCADA network, *Computers & Security*, vol. 70, pp. 16-26, 2017.
17. Roosa, K., Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. M. Hyman, P. Yan, and G. Chowell, Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, *Infectious Disease Modelling*, Vol. 5, pp. 256-263, 2020.
18. Grasselli, Giacomo, Antonio Pesenti, and Maurizio Cecconi, Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast during an Emergency Response, *JAMA-Journal of the American Medical Association*, 2020.
19. Petropoulos, Fotios, and Spyros Makridakis, Forecasting the novel coronavirus COVID-19, *PLOS ONE*, vol. 15, no. 3, pp. 1-8, 2020.
20. Makridakis, Spyros, Allan Andersen, Robert Carbone, Robert Fildes, Michele Hibon, Rudolf Lewandowski, Joseph Newton, Emanuel Parzen, and Robert Winkler, The accuracy of extrapolation

(time series) methods: Results of a forecasting competition, *Journal of forecasting*, vol. 1, no. 2, pp. 111-153, 1982.

21. Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos, The M4 Competition: 100,000 time series and 61 forecasting methods, *International Journal of Forecasting*, vol. 36, no. 1, pp. 54-74, 2020.
22. Petropoulos, Fotios, Nikolaos Kourentzes, Konstantinos Nikolopoulos, and Enno Siemsen, Judgmental selection of forecasting models, *Journal of Operations Management*, vol. 60, pp. 34-46, 2018.
23. Taylor, James W, Exponential smoothing with a damped multiplicative trend, *International journal of Forecasting*, vol. 19, no. 4, pp. 715-725, 2003.
24. Sangeetha, S.Venkatesan, S.Shitharth, Security Appraisal conducted on real time SCADA dataset using cyber analytic tools, *Solid State Technology*, Vol.63, no. 1, pp. 1479-1491, 2020.
25. Shitharth, Masood Shaik and Sangeetha Sirajudeen, Integrated Probability relevancy classification (IPRC) for IDS in SCADA, in *Design Framework for wireless network Lecture notes in network and systems*, Springer, vol. 82, no. 1, pp. 41-64, 2019.
26. Senthil Kumar, C.Sebastian Albina, S.Shitharth, T.Manikandan, Modified TSR protocol to Support Trust in MANET using fuzzy, *International Journal Of Innovative Research In Science, Engineering And Technology*, Vol.3, no. S3, pp. 2551-2555, 2014.
27. Bhattacharya S, Maddikunta PK, Pham QV, Gadekallu TR, Chowdhary CL, Alazab M, Piran MJ. Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. Sustainable cities and society. 2021 Feb 1;65:102589.
28. MK M, Srivastava G, Somayaji SR, Gadekallu TR, Maddikunta PK, Bhattacharya S. An Incentive Based Approach for COVID-19 using Blockchain Technology. arXiv preprint arXiv:2011.01468. 2020 Nov 3.

## Figures

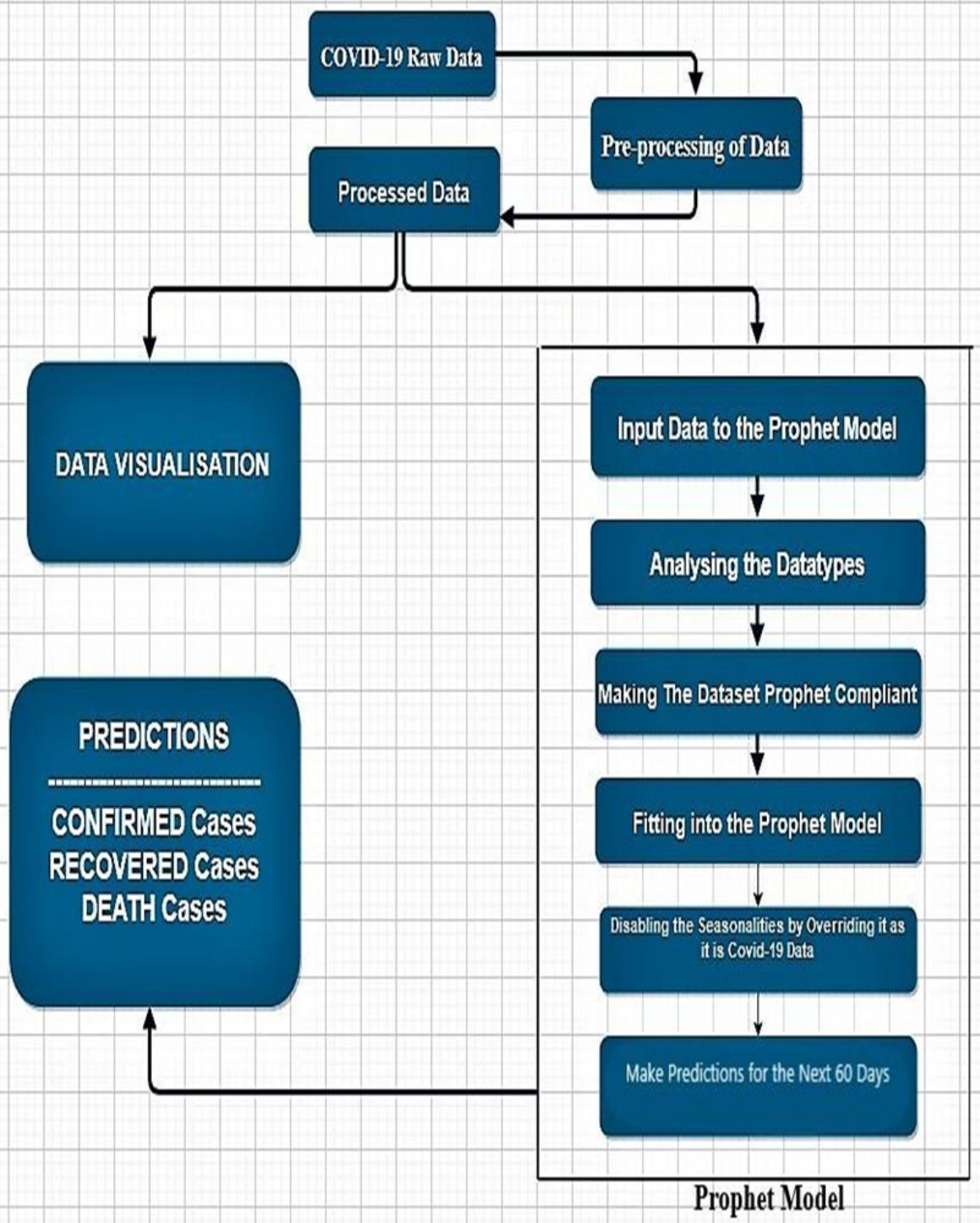
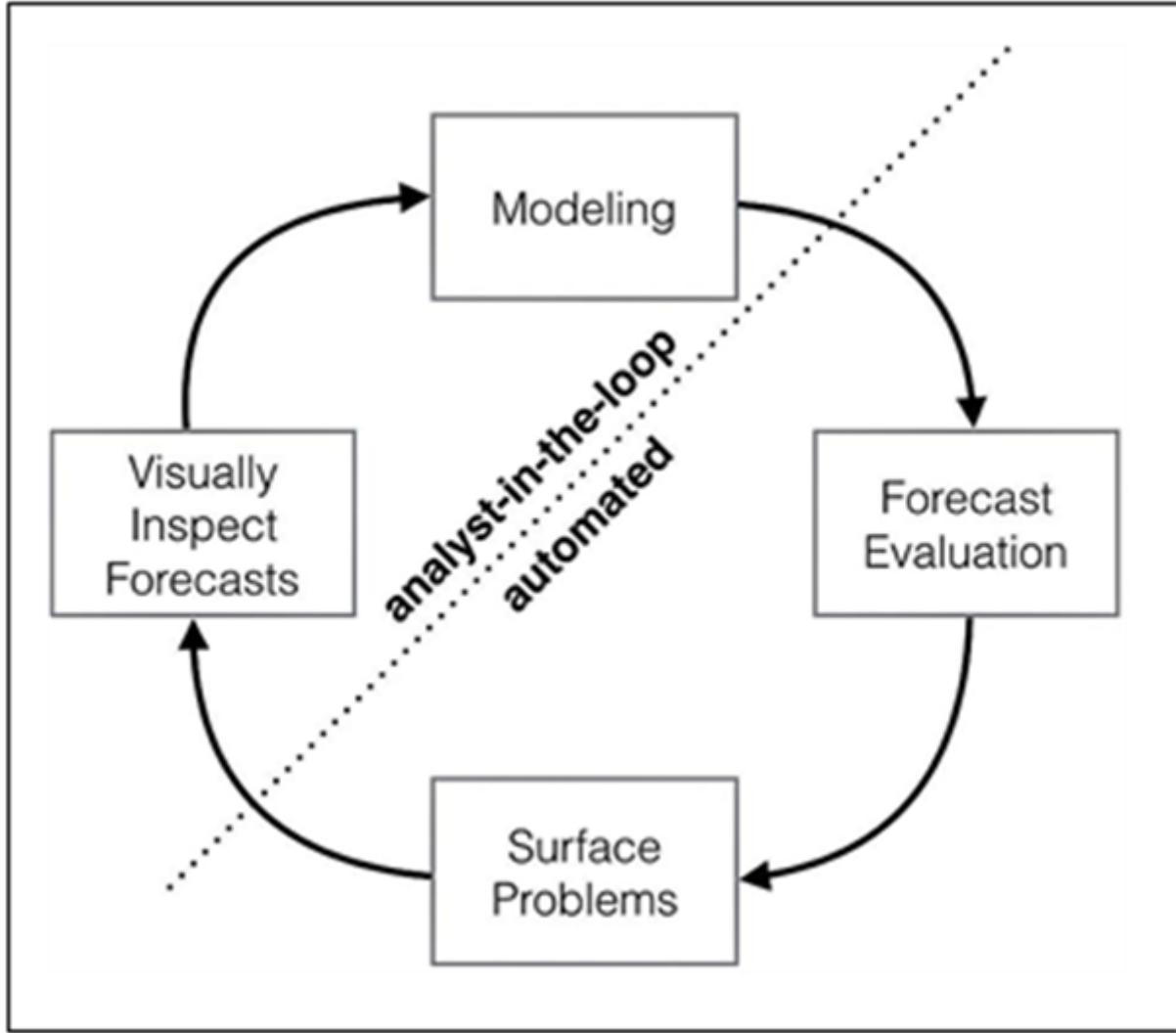


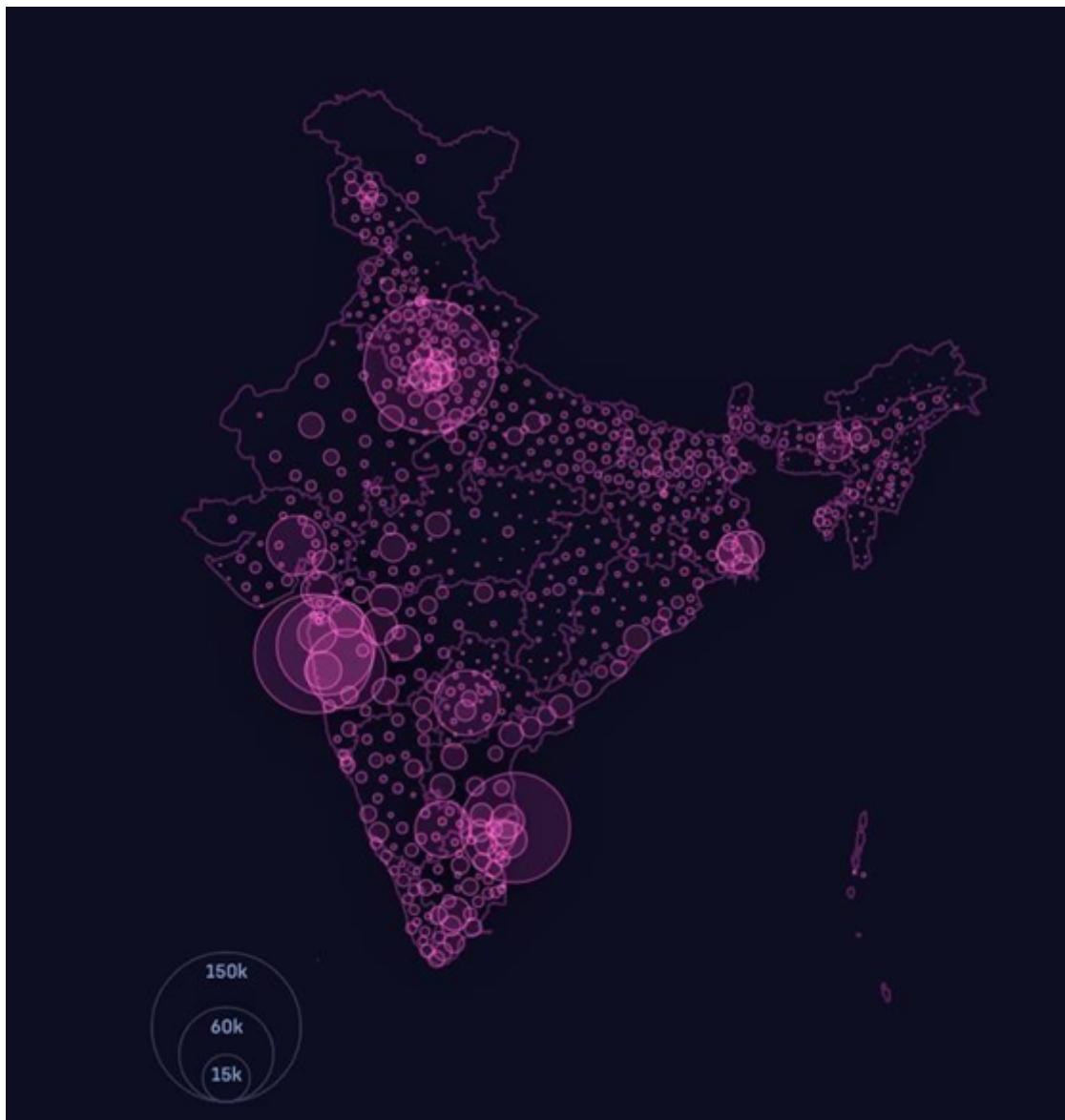
Figure 1

Proposed Model



**Figure 2**

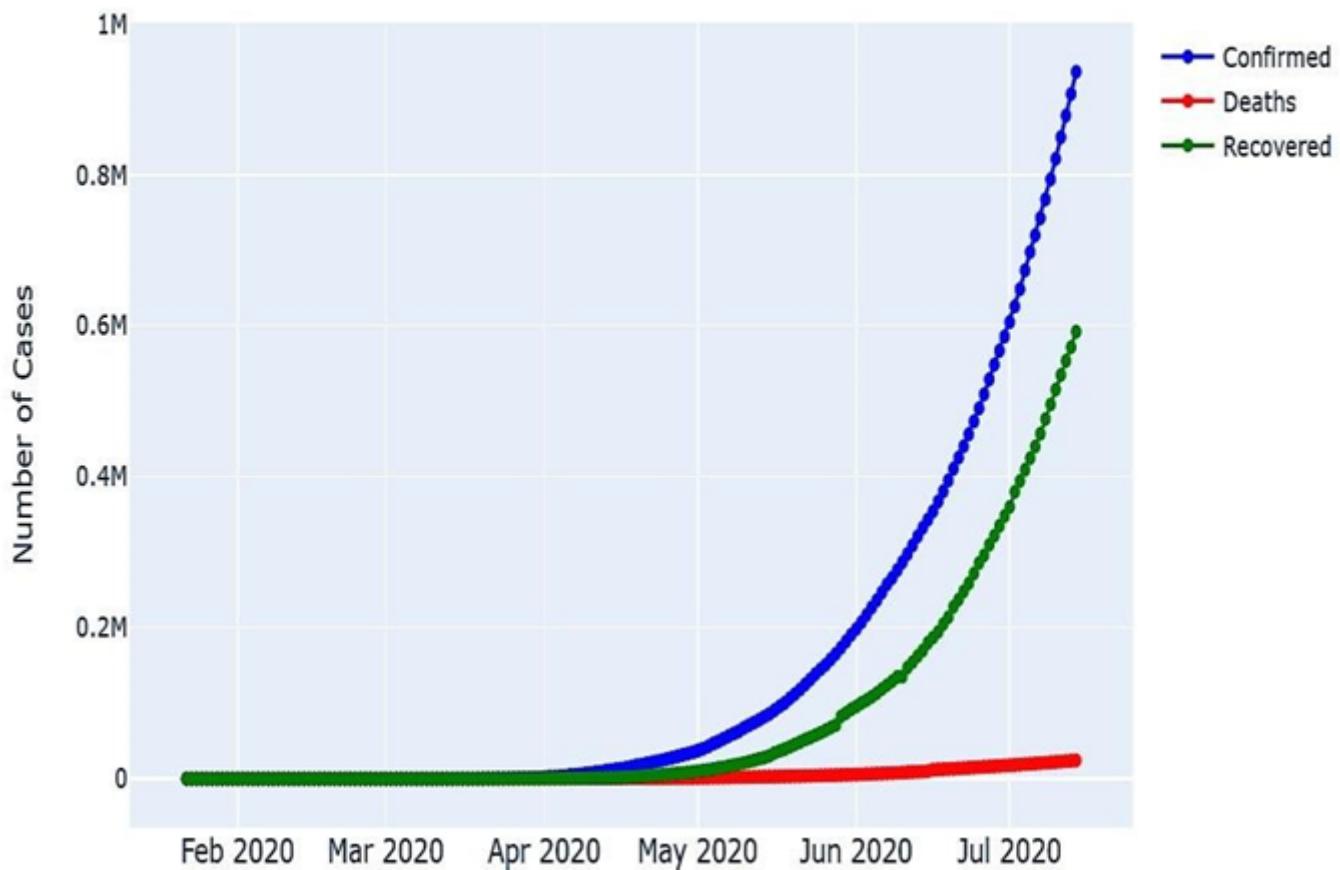
Prophet Forecast Flow



**Figure 3**

Covid-19 Confirmed instances in India

## Indian COVID-19 Cases



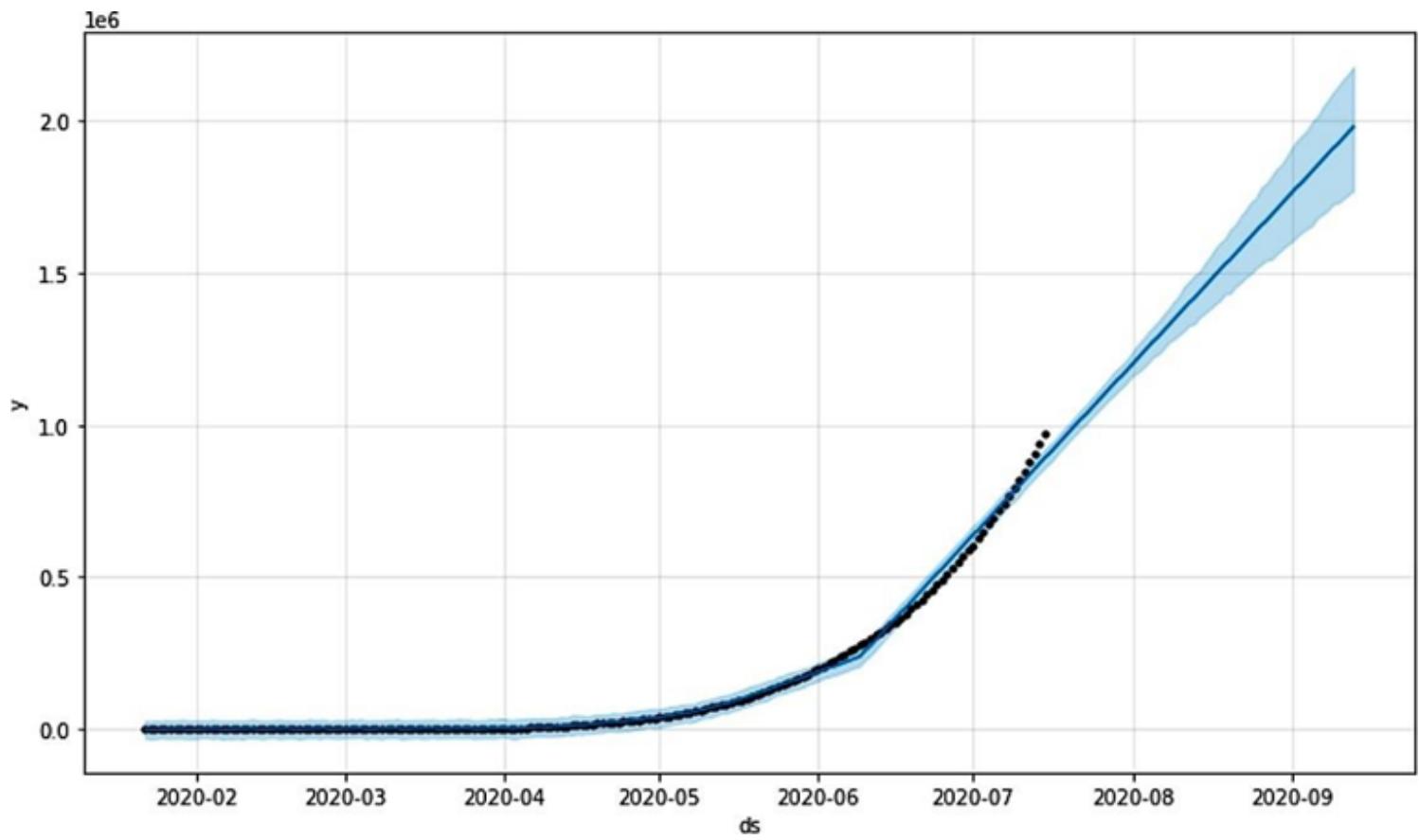
**Figure 4**

Day wise Covid-19 Instances in India

Name of State / UT	Total Confirmed cases	Active	Cured	Deaths
Andaman and Nicobar Islands	166	57	109	0
Andhra Pradesh	33019	15144	17467	408
Arunachal Pradesh	462	306	153	3
Assam	17807	6351	11416	40
Bihar	19284	6261	12849	174
Chandigarh	600	144	446	10
Chhattisgarh	4379	1084	3275	20
Dadra and Nagar Haveli and Daman and Diu	520	209	310	1
Delhi	115346	18664	93236	3446
Goa	2753	1128	1607	18
Gujarat	43637	11065	30503	2069
Haryana	22628	5226	17090	312
Himachal Pradesh	1309	347	951	11
Jammu and Kashmir	11173	4755	6223	195
Jharkhand	4091	1628	2427	36
Karnataka	44077	25845	17390	842
Kerala	8930	4458	4438	34
Ladakh	1093	146	946	1
Madhya Pradesh	19005	4757	13575	673
Maharashtra	267665	107963	149007	10695
Manipur	1672	702	970	0
Meghalaya	318	250	66	2
Mizoram	238	79	159	0
Nagaland	896	550	346	0
Odisha	14280	4342	9864	74
Puducherry	1531	684	829	18
Punjab	8511	2635	5663	213
Rajasthan	25571	5885	19161	525
Sikkim	209	122	87	0
Tamil Nadu	147324	47915	97310	2099
Telangana	37745	12530	24840	375
Tripura	2170	630	1538	2
Uttarakhand	3686	769	2867	50
Uttar Pradesh	39724	13758	24983	983
West Bengal	32838	11927	19931	980

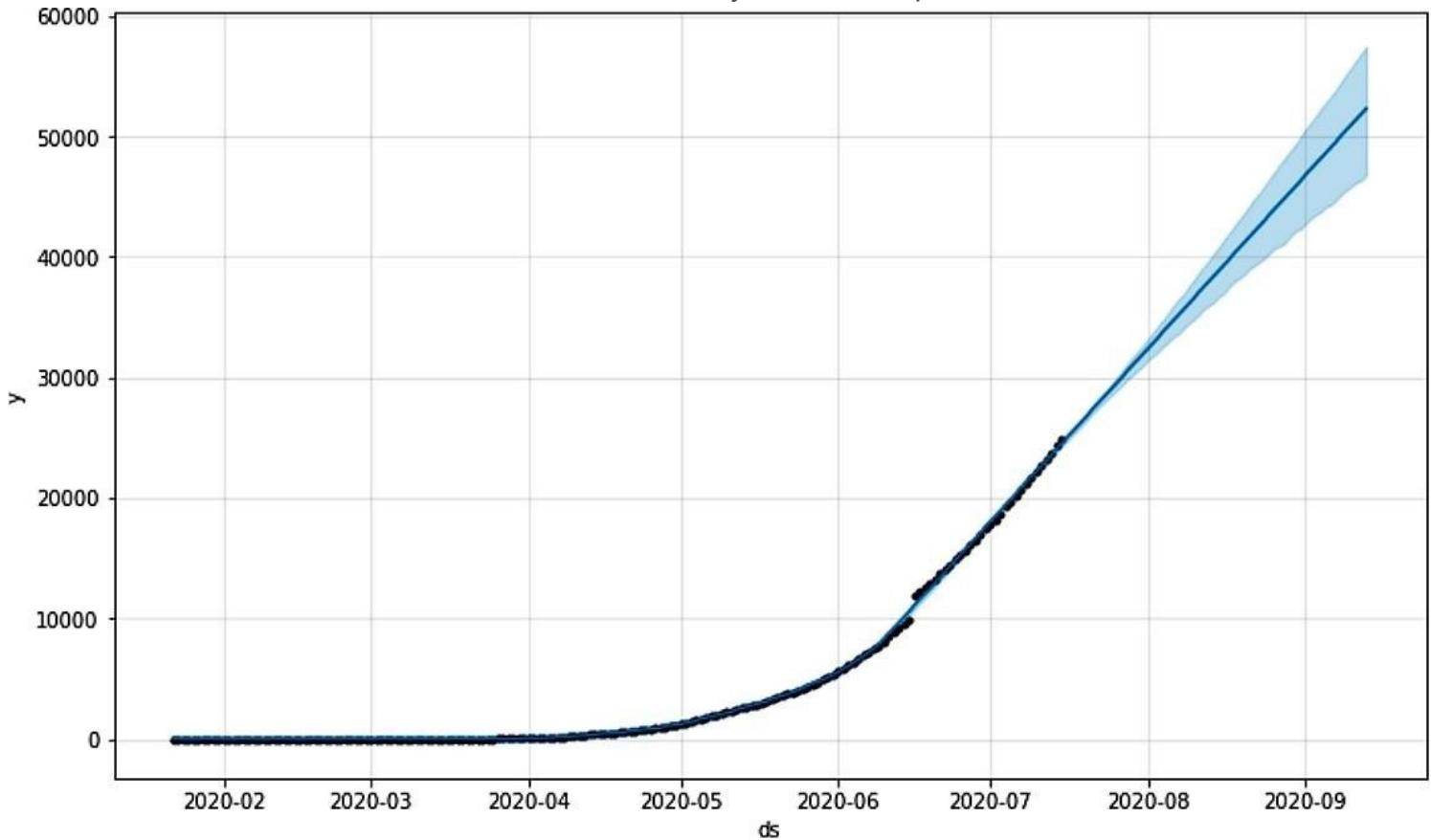
**Figure 5**

State wise Covid-19 Instances in India



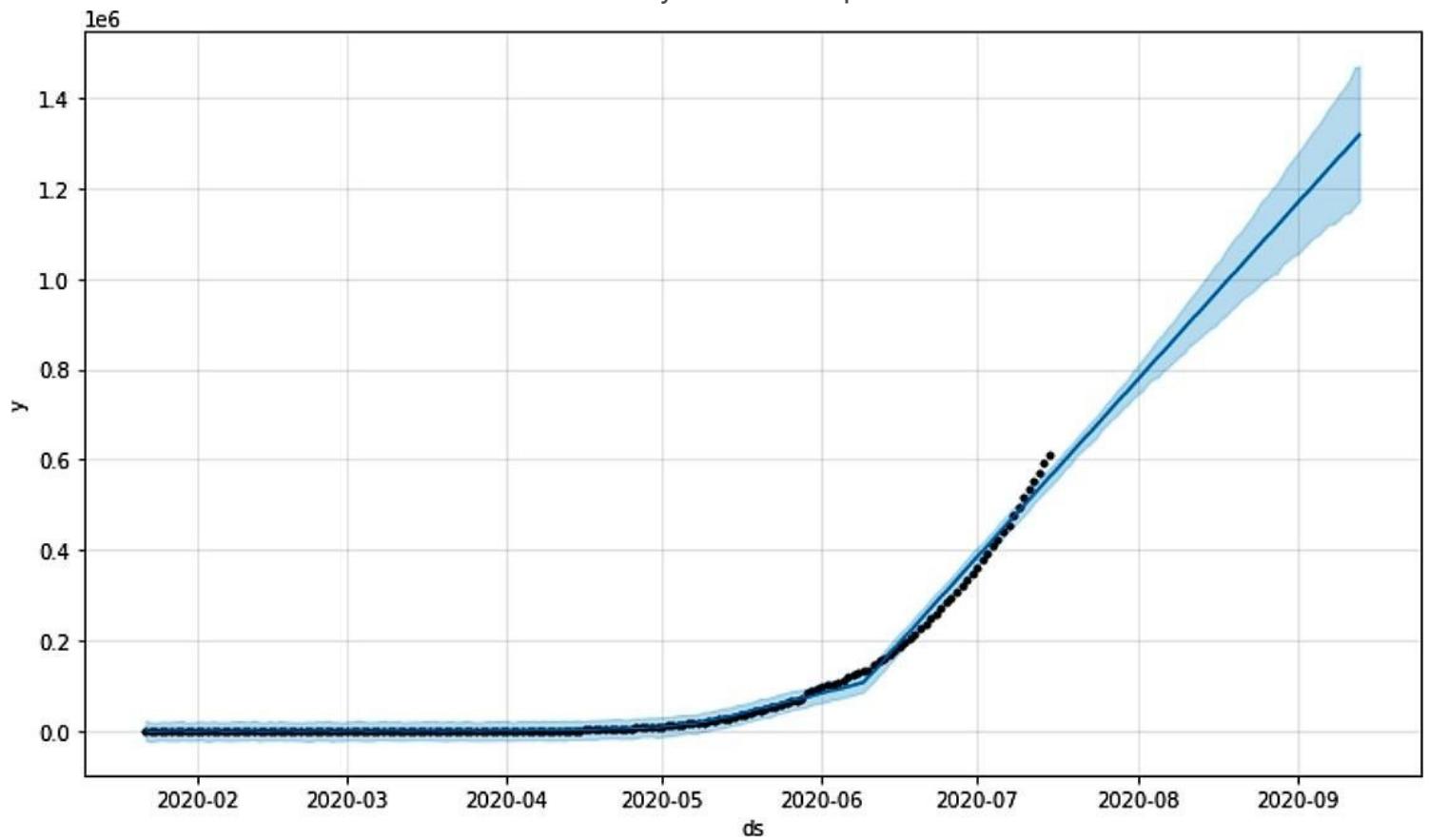
**Figure 6**

Covid-19 Committed Instances Prediction from 16 July 2020- 13 September 2020



**Figure 7**

Covid-19 Death Instances Prediction from 16 July 2020- 13 September 2020



**Figure 8**

Covid-19 Recovered Instances Prediction from 16 July 2020- 13 September 2020