# Chapter 10
# Predicting Epidemic Outbreaks Using IOT, Artificial Intelligence and Cloud

**S. Shitharth, Gouse Baig Mohammad, and K. Sangeetha**

## 10.1 Introduction

After celebrating a new year 2020, no one knows is sick. As usual, it feels like a very fine day. A few people around you are sick and suddenly, you get that everyone is sick and it sounds very threatening. It was happening very rapidly. This is the paradox of pandemic. In this article, we are going to analyze the outbreak of COVID-19 using Machine Learning.

At very end of the year December 2019 outbreak in China, the WHO organization had foundSARS-CoV-2 as a new type of coronavirus and at the drop of the hat outbreak spread around the world. Novel Coronavirus also known as COVID-19 is caused by SARS-CoV-2. This coronavirus is enough capable to infect dozens of people around it. The virus starts showing it's symptoms after 10 to 12 days, which is most worried thing. COVID-19 thread is not the first and last viral pandemic. However, like never before this virus killing people and spreading very massively.

On 30 January 2020, a very first report is generated by Kerala-based laboratory who confirmed case of COVID-19. The patient is student by profession who returns earlier from Wuhan. After some days, a 65-year-old man from Mumbai who had travel history to the UAE is reported as 10th victim found in India. The PMO and the MoHFW have close eyes on 2019-nCoV situation. When ministry saw things

S. Shitharth (✉) · G. B. Mohammad
Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India
e-mail: gousebaig@vardhaman.org

K. Sangeetha
Department of Computer Science and Engineering, KG Reddy College of Engineering and Technology, Hyderabad, India
e-mail: sangeetha@kgr.ac.in

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2021

P. Siarry et al. (eds.), *The Fusion of Internet of Things, Artificial Intelligence, and Cloud Computing in Health Care*, Internet of Things,
https://doi.org/10.1007/978-3-030-75220-0_10

are going worst, then the Prime minister of India came up with a decision and addressed on same day of 24th march 2020. He clarified real situation and requested to be self-quarantine.

Now it became extremely important and essential to control the novel corona virus not only in India but also throughout the world. Without getting late WHO announces COVID-19 outbreak as a pandemic. Now ministry needs to control the spread of virus and predict its risks of infection. The major priority is to identify the infected patient and collect as much as data by testing individual ones.

Here is main role of Healthcare services. They can able to collect the reports by doing testing, researching vaccines for this virus, curing patients. They are also able to provide death reports, confirm reports, bed requirements and all [1].

IT sector also having their own challenges in this pandemic, like collecting massive volume of data generated by clinics, medicals. These reports are essential to analyze outbreak, tracking virus, identifying risk, understanding virus better, diagnose current patients, predict the spreading of virus, predict further pandemics and most important securing our future [2].

All these magic could happen by using Machine Learning, Big Data, Deep learning and Artificial Intelligence and those techie words will be proven soon as a trump card in this war.

Therefore, what is CoronaVirus, how it infects and how pandemic works we have to understand it first. In addition, how can we take help from ML, AI, Deep Learning and Big Data to fight with COVID-19.

### *10.1.1 What Is COVID-19?* (the Problem)

Coronavirus is an infectious disease emerged in Wuhan, China. The new coronavirus spread through person to person. This virus spreads primary through droplets of saliva or by coughing and sneezing. It is diagnosed with a laboratory test. There is no vaccine available for this virus till date.

### *10.1.2 How Can We Detect?* (the Symptoms)

Coronavirus works on different people in a different way. Most of the people have good immunity to fight with this virus. So, they will recover naturally, without hospitalization. This type of people gets only mild to moderate level of Illinois.

Here are some common symptoms:

1. Fever
2. Dry Cough
3. Weakness

Here are some major symptoms:

1. Breathing problem
2. Chest pain
3. Loss of speech or movement

### *10.1.3 How Can We Break the Virus?* (the Solution)

This virus can infect bodies, if it cannot find bodies to infect it will end automatically. Social distancing, self-quarantine, sanitizing are the effective ways to stop infecting people and spreading the virus.

### *10.1.4 What Is Outbreak, Epidemic and Pandemic?*

COVID-19 has a unique property which makes this virus most dangerous. This virus grows **exponentially**. That means, it become double day-by-day and unfortunately there is no vaccine available to cure the patients.

However, it cannot go on last. The virus will eventually stop finding people to infect and ultimate will go slow down the count. This is called **logistic growth.**

An **outbreak** is when the disease happens in unpredicted multitude. It may stick in one zone or expand more extensively. An outbreak can last for few days or some years. Sometimes, authority reviews a single case of a contagious disease to be an outbreak. This could happen when if it is disclosed disease or virus, if it is latest to a community or if it is been missing from a community for a long-term [3].

An **epidemic** is when transmissible disease expands rapidly in regional community than experts/authority would expect. It usually infects a larger region than an outbreak [4].

A **pandemic** is when an epidemic occurs across countries or continents. It infects in large amount and takes more lives than an epidemic. The WHO announces COVID-19 as a pandemic when it became clear that the disease was severe and that it was growing rapidly over a large region [5].

## 10.2 Environment and Tools

### *10.2.1 Machine Learning*

**Overview**

ML is an emerging technology day by day in different sectors. Now-a-days 'health care' is the area where ML applications are in high demand. But the question is that, **what is machine learning?**

ML is a form of AI (artificial intelligence) that enables s/w applications to become more precise in predicting systems results without being explicitly programmed. ML is an approach of data analysis that robotizes analytical model building [6]. It is an arm of AI based on the goal that machines should be able to grasp and self-adjust through previous experience. To fight the COVID-19 Pandemic AI-Driven Informatics, Sensing, Imaging and Big Data Analytics are highly useful and its results are so authentic [7].

The process of grasping and learning starts with analyzing on data. The primary focus is giving access to the systems to learn automates without user interference and changes actions accordingly.

### Why Is Machine Learning Important?

ML can help to enhance 'health related data management and exchange of health statistics' with the aim of technologize updated workflows, ease access to clinical data and upgrading the precision and flow of health details [8].

It also help to pathologists make faster and more precise diagnose further more to identify patients who might sake of new types of treatments and therapies.

### Methods of Machine Learning

Two main trendy methods of ML are supervised learning and unsupervised learning. Supervised learning is about 70% of ML, although unsupervised learning is about 10–20% of ML. Reinforcement learning and semi-supervised learning methods are less used. Gaming, finance sector and manufacturing sector lie under reinforcement learning [9].

#### Supervised Machine Learning Algorithms

Supervised learning is all about 'Classification' and 'Regression'. This algorithm enables fraud detection, e-mail spam detection, diagnostics and image classification. It also helps in risk assessment and scores prediction. The technic is able to issue targets for any new input after sufficient training. The machine learning algorithm further compares its output with the right results, intended output and search errors to modify and customize the model accordingly [10, 11].

#### Unsupervised Machine Learning Algorithms

Unsupervised ML supports 'Dimensionality Reduction' and 'Clustering'. Dimensionality includes text mining, face recognition, big data visualization and image recognition. It also helps in biology, city-planning sectors. Unsupervised learning does study and analyze in order to systems could derive a function to set out a hidden structure from unlabelled data [12]. The system is not able to check correct output, but it can able to analyze the data and can draw inferences from provided datasets to describe hidden structures from unlabelled data. These algorithms do not need any pre-requirements like training with desired outcome data. Instead, they use an iterative approach called Deep Learning to review data and arrive at wind-up [13, 14].

Semi-Supervised Machine Learning Algorithms

This is a sort of combination of supervised and unsupervised learning and use both labelled and unlabelled data for analyzing. This type of ML can be used for methods like classification, regression and prediction. In semi-supervised learning, it would be like face and voice recognition techniques. In a typical situation, the algorithm will use a small amount of labelled data with a large amount of unlabelled data [15].

Reinforcement Machine Learning Algorithms

Reinforcement ML again uses same methods such as classification, regression and prediction. Reinforcement learning is very different from supervised learning. This ML algorithm is all about sequential decisions, in other hand in supervised learning decision made under starting inputs [16] (Table 10.1).

**The Machine Learning Process (Fig. 10.1)**

### *10.2.2 Deep Learning*

**Overview**

A survey on deep learning in medicine: Why, how and when? [19] shows the impact of the algorithm in the medicine field. It is an arm of machine learning. Alike ML, deep learning also has supervised, unsupervised and reinforcement learning in it.

**Table 10.1** Difference between supervised learning and unsupervised learning [17, 18]

| Factors | Supervised learning | Unsupervised learning |
|---|---|---|
| Input | Well-known and labelled data | Unspecified data |
| Complexity | Very complicated | Less complicated |
| Number of classes | Known | Undisclosed |
| Accuracy | Precise and authentic | Average in accuracy and reliable |

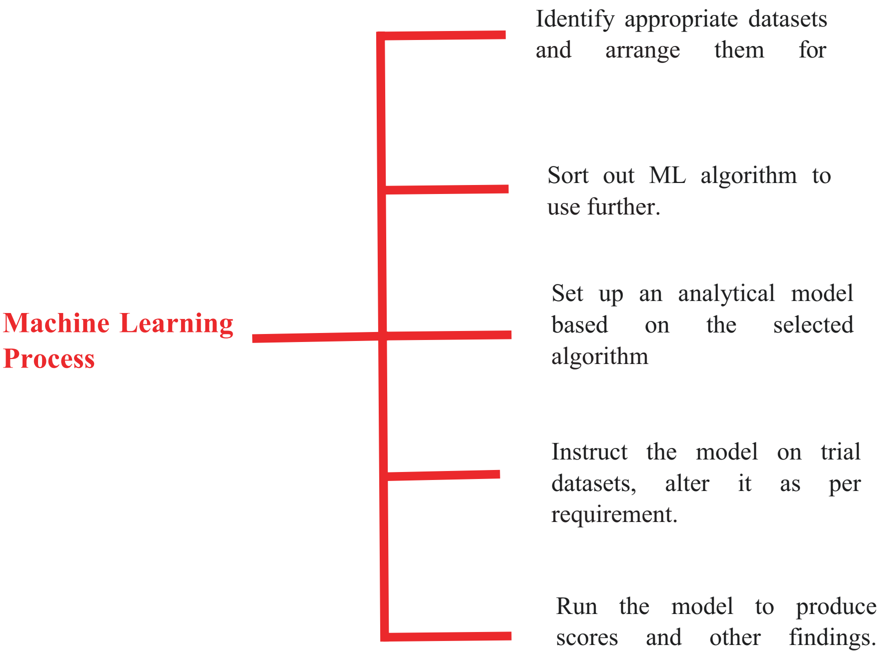


**Fig. 10.1** Machine learning process [3]

The word 'deep' in deep learning indicates use of multiple layers in network. Most of the latest models are built on artificial neural network, CNN, although they also have propositional formulas sorted in layer-wise (Fig. 10.2).

Deep Learning is all about integrating such unseen layers between the initial and the final layer. Even Deep learning-based cardiovascular image diagnosis also been in research to show the versatility of the algorithm [20].

### Methods of Deep Learning

There are some different methods implemented in deep learning. Every suggested method has a certain use case like the sort of data we have, so it is either supervised or unsupervised learning, what kind of task you would want to solve [21]. Therefore, it is all about on these factors, you choose one of the methods that can best solve your problem (Fig. 10.3) (Table 10.2).

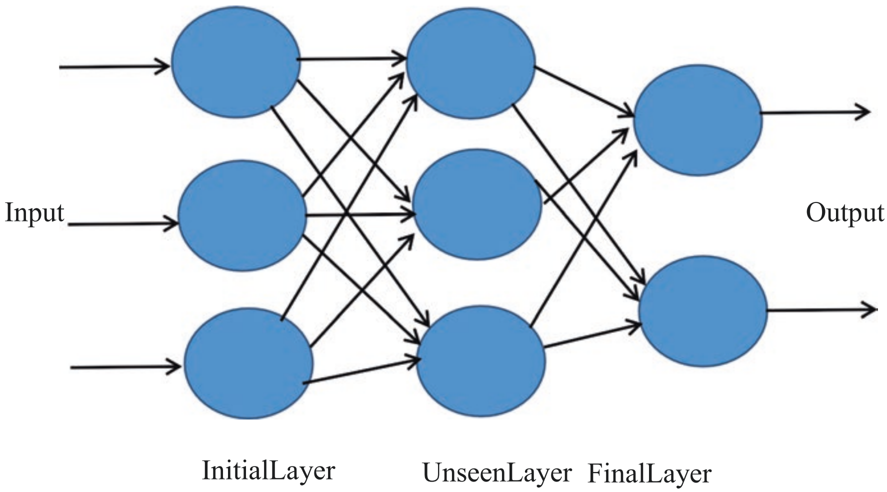


**Fig. 10.2** Shallow neural network [5]

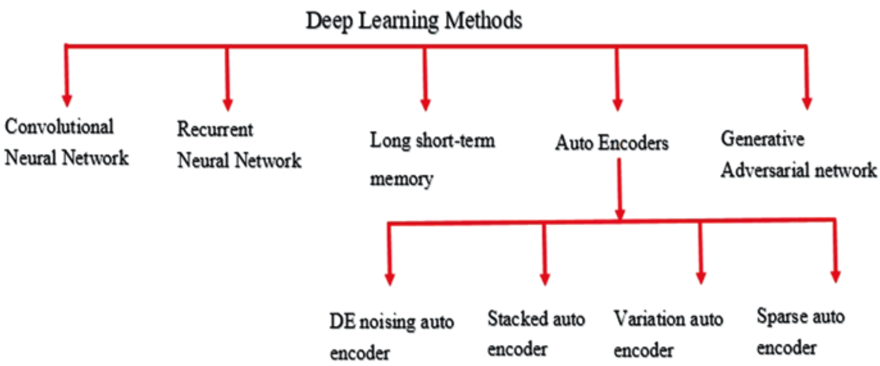


**Fig. 10.3** Deep learning method [5]

**Table 10.2** Deep learning versus machine learning [5]

| Factors | Deep learning | Machine learning |
|---|---|---|
| Data requirement | Requires large data. | Can train on less data. |
| Accuracy | Provides high accuracy | Gives less accuracy. |
| Training time | Takes longer to train. | Take lesser time to train. |
| Hardware dependency | Requires GPU. | Trains on CPU. |
| Hyper parameter tuning | Can be tuned in various different ways. | Limited tuning capabilities. |

## 10.3 Analyzing the COVID-19 Epidemic

### *10.3.1 Overview*

**Objective**

Idea is to come up with a really strong model that can able to predict how coronavirus could spread across different countries and in regions.

**Goal**

The task is to predict spread of the virus in next 7 days.

To analyze situation, we need to collect all type of medical data. On that data further, we can apply various methods and to get better understanding, we visualize this data graphically (pie chart, bar graphs, etc.) [22].

Here, we are using **python,** a scripting language. This programming language is very effective when it comes to analyzing on big data.

(a) First, let us understand why these libraries are essential and how we use it, in our analysis (Table 10.3).

### *10.3.2 Analyzing Present Condition in India*

It is important to analyze present condition in India. As we already discussed, India is on that stage 2 of pandemic progression, which is why shutdown is important.

We are now finding some similarities and differences between counts of confirm cases in India with other country's confirmation cases [25]. But, while we comparing India with other countries, we should select same trending countries. Therefore, we could analyze future domestic losses and we will be preparing for any other unknown risk.

We are also exploring worldwide data and keep updating to our healthcare sector and the dataset. We already have a dataset in form of excel file. Using that same data, we are creating a frame using ***Pandas.*** This library helps us to read tabular form of data.

**Table 10.3** Libraries [9]

| Libraries | Description |
|---|---|
| Pandas | It is mainly used for data analysis and manipulation. |
| Matplotlib | It is a graph plot library. It gives an OOP-based API for insert plots into applications by using GUI [23]. |
| Seaborn | This library based on matplotlib. It is mainly used for statistical data visualization [24]. |
| Folium | We used this library to populate a geographical map. |

### Track Cases in Indian States/Territories

Now, we have name of states in India (Fig. 10.4).

### Find Total Confirm Cases

Now, we are finding Total No. of confirm cases (National + International) (Fig. 10.5).

So, as per statistics, as of 22nd March 2020 India has total **562** confirmed cases.

### Graphical Representation (Total Cases) (Fig. 10.6)

As per figure, the darker the red is in each of these cells the more the number of fatalities are. Actually here, we coloured each cell according to the fatality rate. As we can see, 'Karnataka', 'Kerala' and 'Maharashtra' have largest number of cases 41, 109, 101, respectively. Least Number of cases are in 'Chhattisgarh', 'Manipur' and 'Mizoram' with only one case each, as per 25th March statistics [26].

| | S. No. | Name of State / UT | Total Confirmed cases (Indian National) | Total Confirmed cases ( Foreign National ) | Cured | Death |
|---|---|---|---|---|---|---|
| 0 | 1 | Andhra Pradesh | 9 | 0 | 0 | 0 |
| 1 | 2 | Bihar | 3 | 0 | 0 | 1 |
| 2 | 3 | Chhattisgarh | 1 | 0 | 0 | 0 |
| 3 | 4 | Delhi | 30 | 1 | 6 | 1 |
| 4 | 5 | Gujarat | 32 | 1 | 0 | 1 |
| 5 | 6 | Haryana | 14 | 14 | 11 | 0 |
| 6 | 7 | Himachal Pradesh | 3 | 0 | 0 | 1 |
| 7 | 8 | Karnataka | 41 | 0 | 3 | 1 |
| 8 | 9 | Kerala | 101 | 8 | 4 | 0 |
| 9 | 10 | Madhya Pradesh | 9 | 0 | 0 | 0 |
| 10 | 11 | Maharashtra | 98 | 3 | 0 | 2 |
| 11 | 12 | Manipur | 1 | 0 | 0 | 0 |
| 12 | 13 | Mizoram | 1 | 0 | 0 | 0 |
| 13 | 14 | Odisha | 2 | 0 | 0 | 0 |
| 14 | 15 | Puducherry | 1 | 0 | 0 | 0 |
| 15 | 16 | Punjab | 29 | 0 | 0 | 1 |

**Fig. 10.4** COVID cases in India [2]. Total confirm cases (Indian National), Total confirm cases (Foreign National), cured cases and death cases

```
total_cases=df['Total cases'].sum()
print('Total No. of confirmed covid-19 cases till date[22/03/2020]:',total_cases)

Total No. of confirmed covid-19 cases till date[22/03/2020]: 562
```

**Fig. 10.5** Confirmed cases in India [2]

| | Name of State / UT | Total Confirmed cases (Indian National) | Total Confirmed cases ( Foreign National ) | Cured | Death | Total cases |
|---|---|---|---|---|---|---|
| 0 | Andhra Pradesh | 9 | 0 | 0 | 0 | 9 |
| 1 | Bihar | 3 | 0 | 0 | 1 | 3 |
| 2 | Chhattisgarh | 1 | 0 | 0 | 0 | 1 |
| 3 | Delhi | 30 | 1 | 6 | 1 | 31 |
| 4 | Gujarat | 32 | 1 | 0 | 1 | 33 |
| 5 | Haryana | 14 | 14 | 11 | 0 | 28 |
| 6 | Himachal Pradesh | 3 | 0 | 0 | 1 | 3 |
| 7 | Karnataka | 41 | 0 | 3 | 1 | 41 |
| 8 | Kerala | 101 | 8 | 4 | 0 | 109 |
| 9 | Madhya Pradesh | 9 | 0 | 0 | 0 | 9 |
| 10 | Maharashtra | 98 | 3 | 0 | 2 | 101 |
| 11 | Manipur | 1 | 0 | 0 | 0 | 1 |
| 12 | Mizoram | 1 | 0 | 0 | 0 | 1 |
| 13 | Odisha | 2 | 0 | 0 | 0 | 2 |
| 14 | Puducherry | 1 | 0 | 0 | 0 | 1 |
| 15 | Punjab | 29 | 0 | 0 | 1 | 29 |
| 16 | Rajasthan | 30 | 2 | 3 | 0 | 32 |
| 17 | Tamil Nadu | 16 | 2 | 1 | 0 | 18 |
| 18 | Telengana | 25 | 10 | 1 | 0 | 35 |
| 19 | Chandigarh | 7 | 0 | 0 | 0 | 7 |
| 20 | Jammu and Kashmir | 7 | 0 | 1 | 0 | 7 |
| 21 | Ladakh | 13 | 0 | 0 | 0 | 13 |
| 22 | Uttar Pradesh | 34 | 1 | 11 | 0 | 35 |
| 23 | Uttarakhand | 3 | 1 | 0 | 0 | 4 |
| 24 | West Bengal | 9 | 0 | 0 | 1 | 9 |

**Fig. 10.6** Graphical representation of confirmed cases in India [6]

### Find Total Active Cases

Now we have total death cases, total cured patients and sum of all. However, these data are not more relevant for our analysis.

What actually we are seeking is ***Active cases***. We only want to know the number of people that have been hospitalized at that moment (Fig. 10.7).

```
#Toatl cases= Number of death + Cured
df['Total Active'] = df['Total cases'] - (df['Death'] + df['Cured'])
total_active=df['Total Active'].sum()
print('Total number of active COVID-19 cases across India:', total_active)
Tot_Cases = df.groupby('Name of State / UT')['Total Active'].sum().sort_values(ascending=False).to_frame()
Tot_Cases.style.background_gradient(cmap='Reds')

Total number of active COVID-19 cases across India: 512
```

$$\text{Total Active Cases} = \text{Total cases} - \left(\text{Total Death} + \text{Total Cured}\right)$$

We can clearly see that, ‘Kerala’ and ‘Maharashtra’ have highest number of Active cases and combined cases in India have 512 of count [27].

Here, we have grouped states and union territories and further we sorted them by the value of their total active cases. Again, here we used same red-coloured gradient to visualize it better.

| Name of State / UT | Total Active |
|---|---|
| Kerala | 105 |
| Maharashtra | 99 |
| Karnataka | 37 |
| Telengana | 34 |
| Gujarat | 32 |
| Rajasthan | 29 |
| Punjab | 28 |
| Uttar Pradesh | 24 |
| Delhi | 24 |
| Tamil Nadu | 17 |
| Haryana | 17 |
| Ladakh | 13 |
| Madhya Pradesh | 9 |
| Andhra Pradesh | 9 |
| West Bengal | 8 |
| Chandigarh | 7 |
| Jammu and Kashmir | 6 |
| Uttarakhand | 4 |
| Himachal Pradesh | 2 |
| Odisha | 2 |
| Bihar | 2 |
| Manipur | 1 |
| Mizoram | 1 |
| Puducherry | 1 |
| Chhattisgarh | 1 |

**Fig. 10.7** Active cases in India

***groupby():***
This is a pivot function mostly we use in excel sheets. It actually turns wide table format into long table format.

**Location-based Tracking *(Total Cases)* (Fig. 10.8)**

Here, we used ***Folium*** library as *folium.map()* and we specified the location (Longitude and Latitude). We also use a red circle marker whose size depends on the number of cases in their particular region [28].

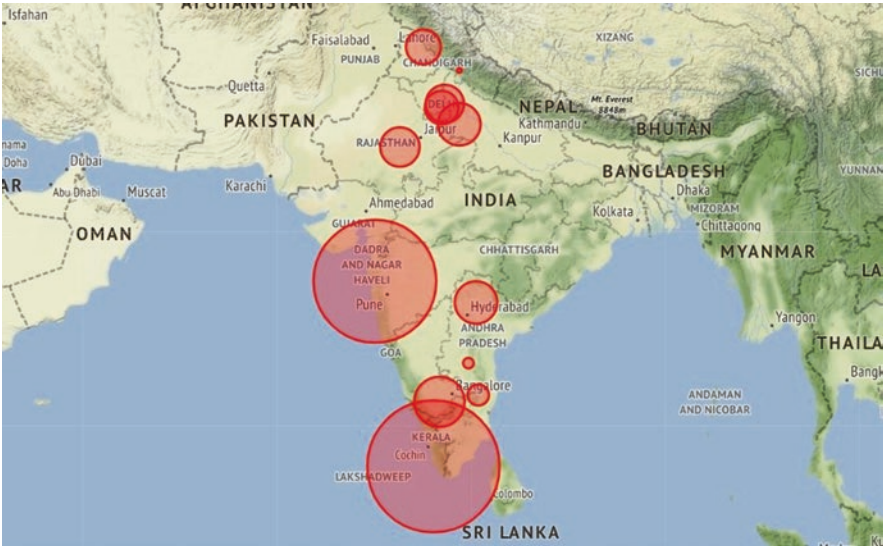


**Fig. 10.8** Visualizing the spread geographically

As we can see, 'Kerala' immediately followed by 'Maharashtra' both have very big red circles. In addition, north Indian side has some couple of circles and east side of India is less affected region.

### Confirmed Versus Recovered Cases (Fig. 10.9)

Here, we are basically using ***seaborn*** library for visualization where we are plotting a couple of bar graphs to showing and comprising total number of sure cases and total number of cured cases in Indian territories. Pink represents total number of cases similarly, cured cases are in green colour [29, 30].

As we can see, again 'Kerala', 'Karnataka' and 'Maharashtra' have highest number of cases and also, 'Haryana' and 'Uttar Pradesh' have good recovery.

If we compare 'Kerala' with 'Maharashtra', Kerala despites maximum number of cases and also maximum number of recovery than Maharashtra [31]. So, conclusion from above graph is that, the net percentage of affected people in Kerala is much lesser than Maharashtra.

### Rise of Coronavirus Cases (Fig. 10.10)

Here, we use a scatter plot and a line plus marker for a better understanding and visualization. This graph shows an actual rise of coronavirus cases in India.

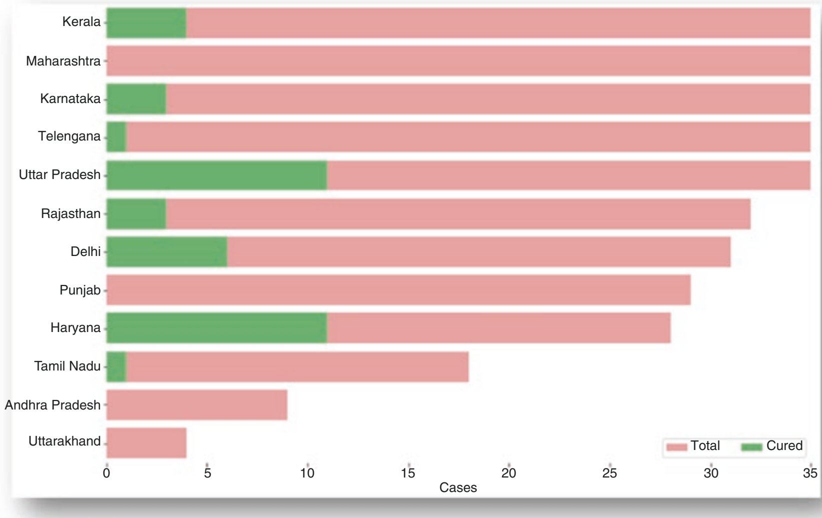


**Fig. 10.9** Total cases and recovered cases

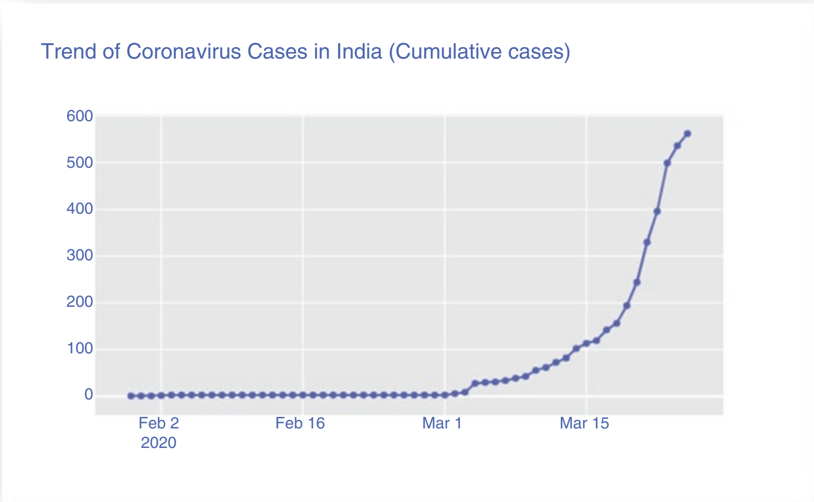


**Fig. 10.10** Trend of coronavirus cases in India [32]

In that above graph, X-axis intended to months similarly, Y-axis intended the cases rise in India. We can easily see how the graph takes a jump in March (Fig. 10.11).

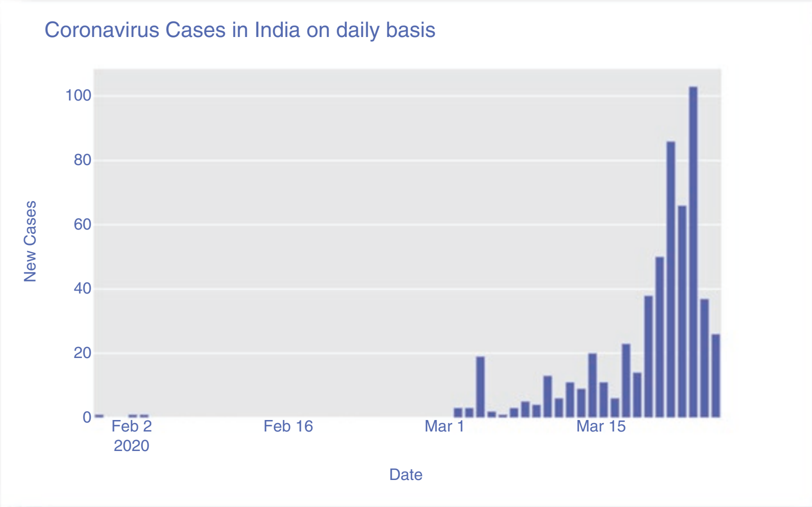


**Fig. 10.11** Daily basis analysis [33]

We also create a daily basis analysis for COVID-19 cases in India to get precise values and data for further analysis and accurate prediction. A SaaS approach for community Body Sensor Networks gives a more detailed view on body cloud [34].

This is a bar chart where our access is date and the values are new daily cases. After getting this output, we clearly see that up till 23rd of February, India has too little cases and after starting March we get more and more cases in India.On 23rd of March, India reached at a peak of 103 new cases on a particular day.

### *10.3.3 India Versus World* (Analyze Similar Trending Countries)

At this point, India had already crossed 500 cases. It is still very important to contain the situation in the upcoming days. The numbers of coronavirus patients had started doubling after many countries hit the 100 marks, and almost starting increasing exponentially.

Now, we have, all confirmed, recovered and death cases report and monthly-daily analysis. Up till here, we analyze about India only. Its time to compare and analyze India with few similar trending cases countries [35].

It is more important to analyze present condition of world. So, we are now finding some similar situations in other countries. These data will help us for better prediction and preparation [36].

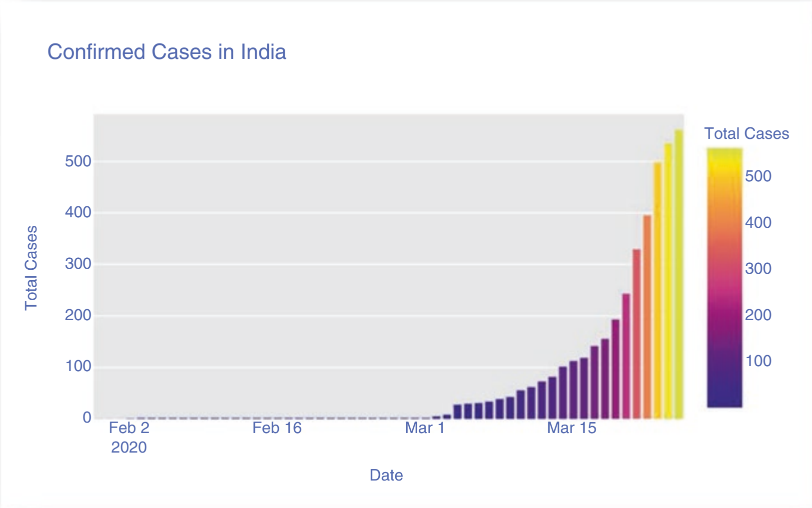


**Fig. 10.12** India

### India *(Confirmed Cases)* (Fig. 10.12)

For this type of graph, we imported column graph using plot Li and some colour gradient. The more dense the colour, the higher the confirm rate. As per our title, we are comparing India with world [37]. This is the graph for India. India has exponential growth in confirm cases and it has taken a hit since 3rd–4th of March.

### Italy *(Confirmed Cases)*

If we talk about Italy's condition, we can see a sharp and exponential increase in confirm case reports after the 3rd–4th of March. But, this graph looks so even and in flow, there are no breaks and kinks unlike India. At the end of March, it shows 69k cases from Italy that is much more than India (Figs. 10.13 and 10.14).

### South Korea *(Confirmed Cases)*

South Korea's graph has started to completely become a sigmoid curve since the 7th of March, the actual story behind is, South Korea had started the extensive testing. The government took very strong decision that anyone who got even mildest infection has been quarantined in this country [38].

This idea really works so that the curve is almost flattened in last few days of March. On 22nd–24th of March, the confirm cases seem so minimal (Fig. 10.15).

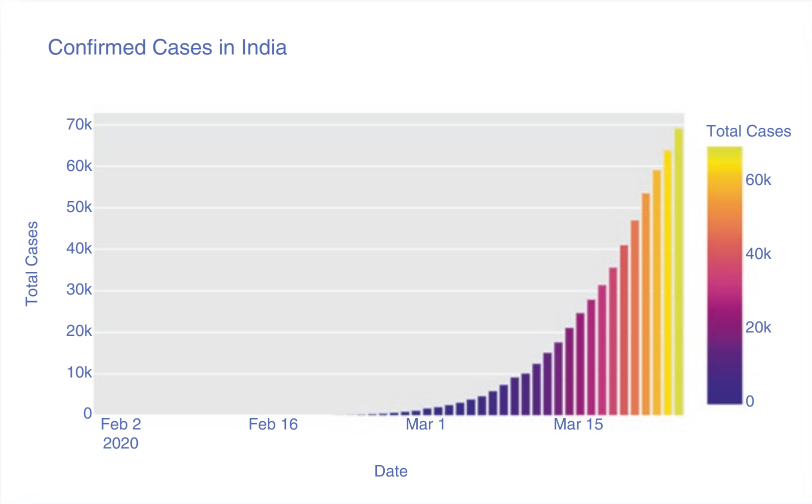


**Fig. 10.13** Confirmed cases in Italy

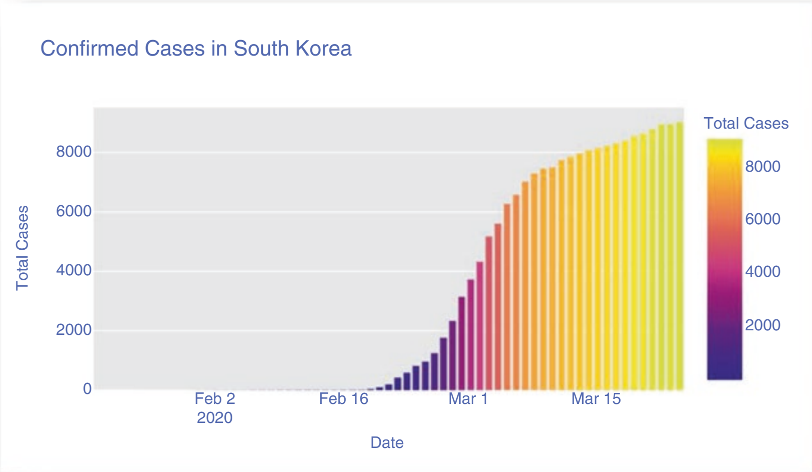


**Fig. 10.14** Confirmed cases in South Korea

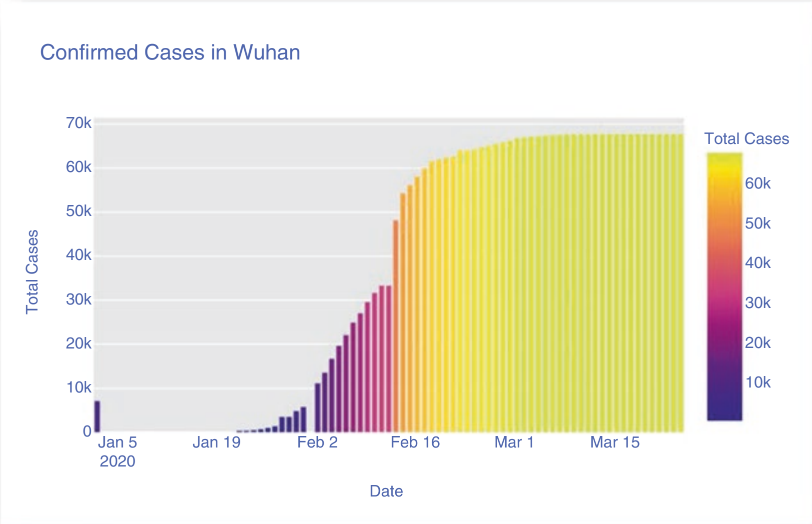


**Fig. 10.15** Confirm cases in Wuhan

### Wuhan

Wuhan's graph is also following sigmoid pattern. On 12th–13th February, there is steep rise. But it is successfully started to flatten out just after 66k–67k on 3rd of March.

China also has their own unique story; China has started to get help from some *artificial intelligent* models, which have helped them to diagnose people with flu on a very extensive scale. They used to scan bodies with heat mapping sensors, these sensors are much able to pick out people with even mild temperatures, which will help to quarantine and diagnose people and this project at its best in machine learning [39].

From the above visualization, one can infer the following:

- Confirmed cases in India are rising exponentially with no fixed pattern (Very less test in India).
- Confirmed cases in Italy are rising exponentially with a certain fixed pattern.
- Confirmed cases in South Korea are rising gradually.
- There have been almost negligible numbers of confirmed cases in Wuhan a week.

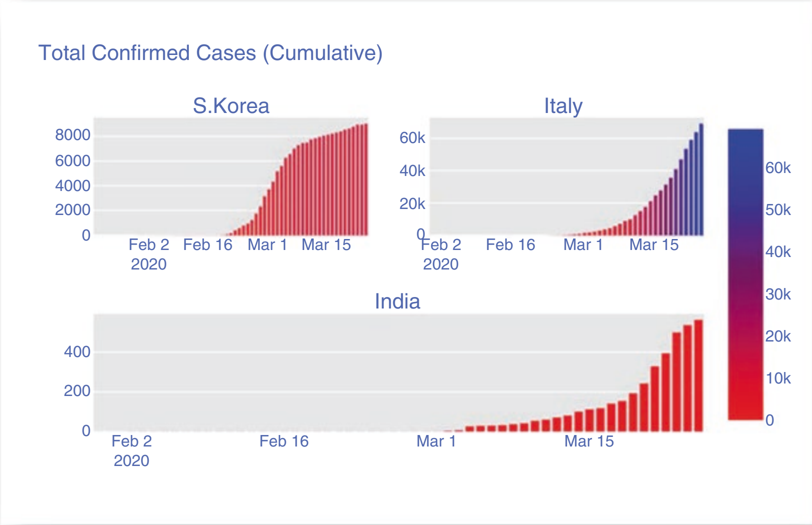


**Fig. 10.16** India: 562 Italy: 69 k South Korea: 9 k

### Overall Comparison (Fig. 10.16)

The more number of cumulative cases per day the bluer the graph becomes. So, we take some countries South Korea, Italy to compare. We put all graphs into a single canvas to differentiate them properly. This type of visualization will contribute more than the previous one [40].

We can see that, India has comparatively fewer cases than South Korea and Italy on the same date. At the same time, if we campier India with these countries, India has large population. What is the reason behind India has less cases? Let us figure it out.

### Trend after Crossing 100 Cases *(India, South Korea, Italy)* (Fig. 10.17)

As we can see, after crossing 100 cases the graph shows India has minimum number of cases and other both countries cross the mark of 5,600.

If we compare India with Italy and South Korea, India has low number in cases when it comes to in this pandemic. Why this is happening?

According to CNN reports, India is actually not testing people enough to find out whether the total number of reported cases are genuine or not and why is a highly populated country with billions of people testing so in less count [41].

By the experts, India has testing below scale because of being under resourced and an uneven public-health system.

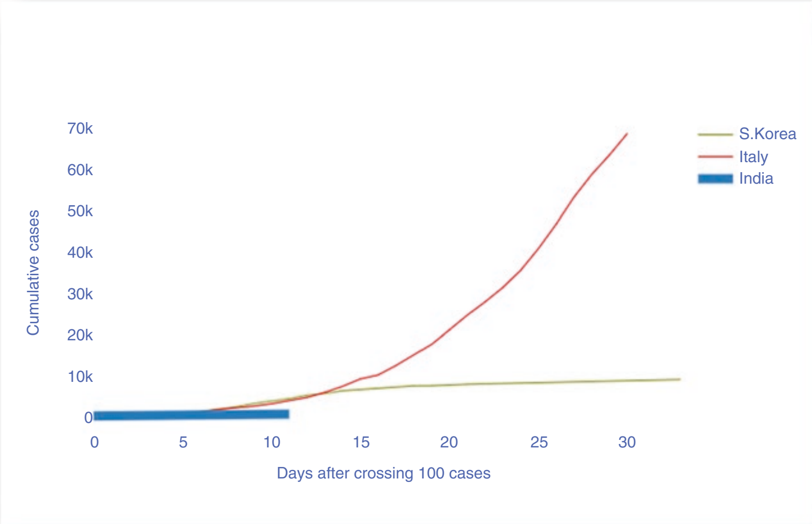


**Fig. 10.17** RED: Italy; Green: South Korea; Blue: India

### *10.3.4 Visualize Worldwide Data (Fig. 10.18)*

Here we are using coloured lines with markers. Blue lines show total number of confirmed cases around the world, Green lines show recovered patients and similarly red line intended to total number of death cases due to the coronavirus in world.

As we can see, between point A and B there is sudden rise in graph and not a perfect curve. Actually, in that particular day 12th of February, an organization came up with a unique method of counting affected people, but by the end of the day they realize that this is not a proper method to count fatalities. Hence, they came back to the original method of counting.

## 10.4 Forecasting/Prediction

### *10.4.1 Forecasting Total Number of Cases Worldwide*

For Forecasting and prediction here, we use an open source software called 'Prophet' which is developed by Facebook core data science team.

We actually use Prophet for forecasting in sort of time series results based data on an additive model where non-linear trends are suitable with yearly, weekly and daily basis.

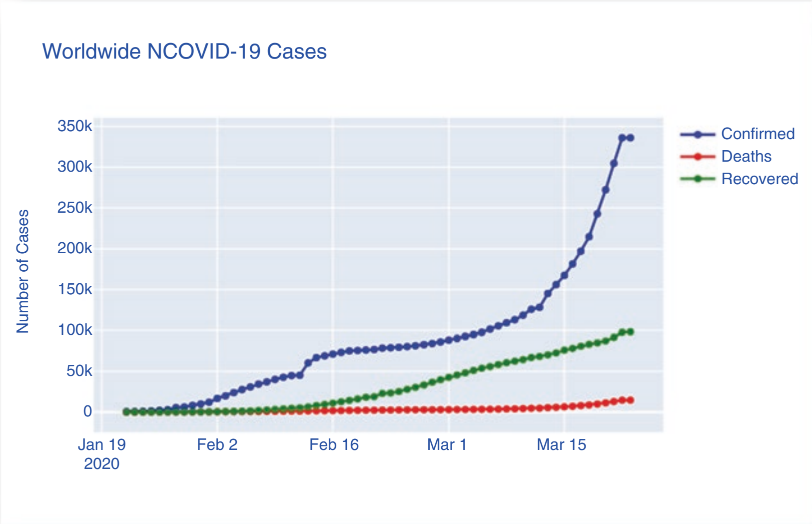


**Fig. 10.18** Worldwide graph

### Why Prophet?

- **Precise and quick:** Prophet is mainly utilized in different applications in Facebook for building authenticate and valid forecasts for goal setting. It is enough quick that you may get forecasts in a bit by using Stan module. Facebook finds it to execute better than any other approach.
- **Automated:** Get a reasonable forecast on messy data with no manual effort. Prophet is robust to outliers, missing data and dramatic changes in your time series.
- **Availability:** Facebook has introduced the Prophet Module procedure with support of Python and R programming language. Both languages share the Stan code. You can use any language that you are comfortable.

Confirm Cases Forecast

Now, it is time to predict upcoming coronavirus cases in the world. Here we are trying to find out a range within which the prediction is going to occur and in an addition to that we are finding upper limit and lower limit so that our prediction and values will not deflect so much (Fig. 10.19).

- yhat:- values which are predicted.
- yhat_lower:- It shows lower limit which is predicted.
- yhat_upper:- It shows how much high cases could go.

| | ds | yhat | yhat_lower | yhat_upper |
|---|---|---|---|---|
| 64 | 2020-03-26 | 355136.872975 | 334546.613119 | 374775.244231 |
| 65 | 2020-03-27 | 372235.326938 | 352367.910827 | 391712.469992 |
| 66 | 2020-03-28 | 388674.964143 | 367586.464833 | 410613.983488 |
| 67 | 2020-03-29 | 405307.954675 | 382990.550208 | 427082.101462 |
| 68 | 2020-03-30 | 418529.648466 | 394208.184567 | 439971.819186 |

**Fig. 10.19** Confirm cases prediction

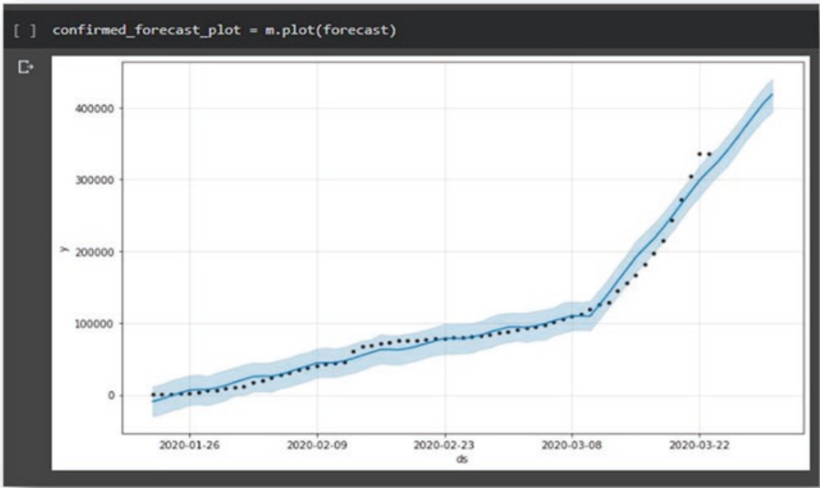


**Fig. 10.20** Confirm forecast

Tolerance lies between yhat_lower to yhat_upper (Fig. 10.20).

Here we visualize the data in graph by putting some plots. We use *prophet plot* Method to plot forecast by passing forecast frame.

As per graph, we can see the graph's line goes beyond 24th of March. Graph is raising constantly day by day (Fig. 10.21).

This graph actually focusses on a particular days of a week. As we can see, there is a dip on Tuesday to Wednesday [42]. These are because there is huge dip in the cases in the china in that particular day.

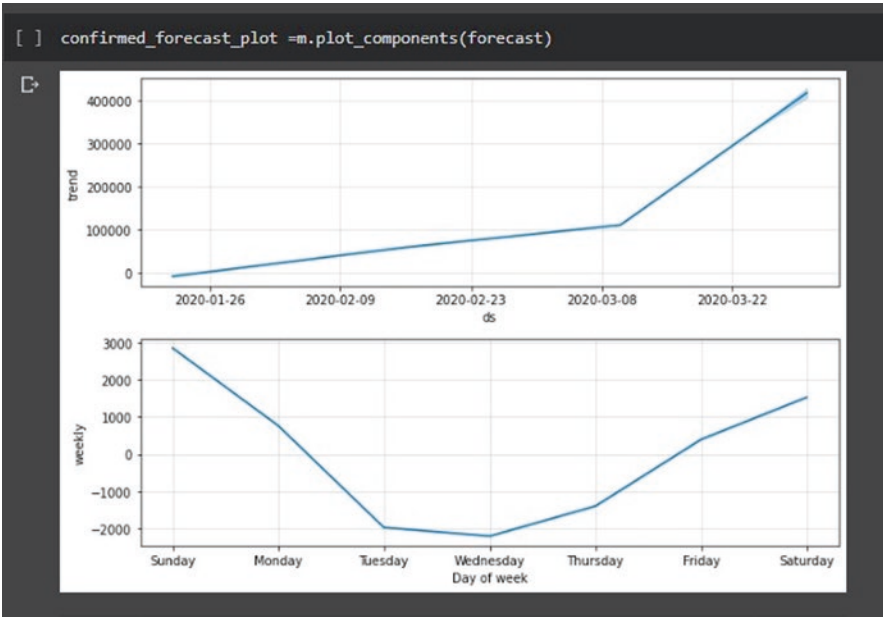


**Fig. 10.21** Confirm forecast *(weekly analysis)*

Death Cases Forecast (Fig. 10.22)

Here, in this graph dots represent the actual value and the blue line is representing the forecasting with upper and lower tolerance as we already calculated and as we can see in the beginning it is coinciding with each other but after 8th of march there is spike in death forecast (Fig. 10.23).

According to forecast, the number of deaths come down from Tuesday through Thursday because obviously the number of confirm cases are predicted to come down between those three days and after that it rises again.

## 10.5 Conclusion

Do not take your cough and cold lightly as you would. If you look at the data, the number of cases in India are rising just like in Italy, Wuhan, South Korea, Spain and the USA. We have crossed 100,000 cases already. Do not let lower awareness and fewer test numbers ruin the health of our world. Currently, India is a deadly and risky zone as there are very few COVID- 19 test centres available. Imagine how many infected people are still around you and are infecting others unknowingly. If the spread of coronavirus goes along with the forecast and as per our model then it would come up with big loss of lives as it presents the exponential growth of the transmission worldwide.

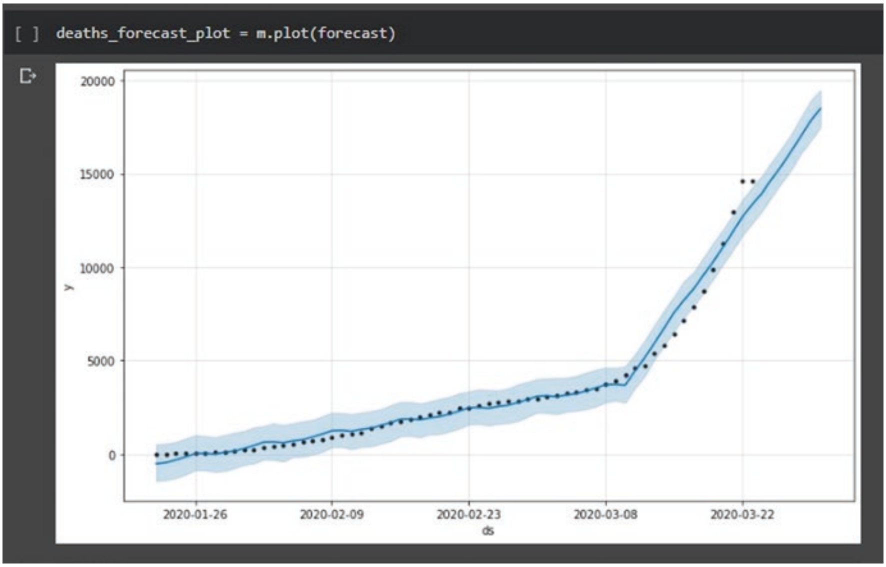


**Fig. 10.22** Death forecast

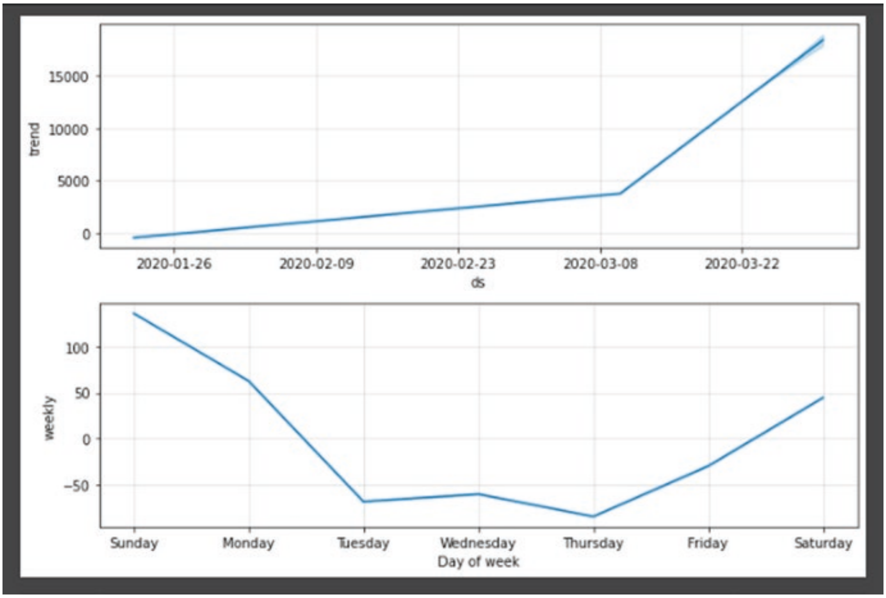


**Fig. 10.23** Death forecast *(weekly analysis)*

Let us give a hand in fighting this pandemic at least by quarantining ourselves by staying indoors and protecting others and ourselves around us. Take precautions, stay indoors.

## References


1. Basic definitions for outbreak epidemic and pandemic. https://www.webmd.com/cold-and-flu/what-are-epidemics-pandemics-outbreaks
2. Covid-2019 datasets from https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge
3. Machine learning information. https://en.wikipedia.org/wiki/Machine_learning
4. Types of Machine Learning. https://searchenterpriseai.techtarget.com/definition/machinelearningML
5. Deep Learning Information. https://www.tutorialspoint.com/tensorflow/tensorflow_machine_learning_deep_learning.htm
6. GitHub Repository. (2019). Novel coronavirus COVID-19 (2019-nCoV) data repository by Johns Hopkins CSSE. Retrieved March 31, 2020, from https://github.com/CSSEGISandData/COVID-19
7. Amini, A., Chen, W., Fortino, G., Li, Y., Pan, Y., & Wang, M. D. (2020). AI-driven informatics, sensing, imaging and big data analytics for fighting the COVID-19 pandemic. *IEEE Journal of Biomedical and Health Informatics, 24*(10), 2731–2732.
8. Navares, R., Díaz, J., Linares, C., & Aznarte, J. L. (2018). Comparing ARIMA and computational intelligence methods to forecast daily hospital admissions due to circulatory and respiratory causes in Madrid. *Stochastic Environmental Research and Risk Assessment, 32*(10), 2849–2859.
9. Barstugan, M., Ozkaya, U., & Ozturk, S. (2020). Coronavirus (covid-19) classification using ct images by machine learning methods. arXiv preprint arXiv:2003.09424.
10. Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*(2), 199–220.
11. Uschold, M. (2003). Where are the semantics in the semantic web? *AI Magazine, 24*(3), 25–36.
12. Ding, L., Finin, T., Joshi, A., Peng, Y., Pan, R., & Reddivari, P. (2005). Search on the semantic web. *IEEE Computer, 10*(38), 62–69.
13. Horrocks, I. (2007). Semantic web: the story so far. In *Proc. of the 2007 int. cross- disciplinary conf. on web accessibility (W4A)* (pp. 120–125). ACM.
14. Zeshan, F., & Mohamad, R. (2012). Medical ontology in the dynamic healthcare environment, the 3rd international conference on ambient systems, networks and technologies (ANT). *Procedia Computer Science, 10*, 340–348.
15. Kaur, P., & Khamparia, A. (2014). Review on medical care ontologies. *International Journal of Science and Research (IJSR), 3*(12), 677–680.
16. Jean-Mary, Y. R., Shironoshita, E. P., & Kabuka, M. R. (2009). Ontology matching with semantic verification. *Journal of Web Semantics, 7*(3), 235–251.
17. Susel Fernández, R., Velasco, J., Marsa-Maestre, I., Miguel, A., & Lopez-Carmona (2012). Fuzzy align: A fuzzy method for ontology alignment. In Proceedings of the international conference on knowledge engineering and ontology development.
18. Shitharth, S., Sangeetha, & Kumar, P. (2019). Integrated probability relevancy classification (IPRC) for IDS in SCADA. *Design Framework for wireless network, Lecture notes in network and systems, Springer, 82*(1), 41–64.
19. Piccialli, F., Somma, V. D., Giampaolo, F., Cuomo, S., & Fortino, G. (2021). A survey on deep learning in medicine: Why, how and when? *Information Fusion, 66*, 111–137.

20. Wong, K. K. L., Fortino, G., & Abbott, D. (2021). Deep learning-based cardiovascular image diagnosis: A promising challenge. *Future Generation Computer Systems, 110*, 802–811.
21. Ngo, D., & Bellahsene, Z. (2012). YAM++: A multi-strategy based approach for ontology matching task. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) LNAI* (Vol. 7603, pp. 421–425).
22. Bodenreider, O. (2013). *Disease ontology, encyclopedia of systems biology* (pp. 578–581). Springer.
23. Ibrahim, A. M., Hashi, H. A., & Mohamed, A. A. (2013). Ontology driven information retrieval for healthcare information system: A case study. *International Journal of Network Security & Its Applications (IJNSA), 5*(1).
24. Babcock, S., Lindsay, G., Cowell, J. B., & Smith, B. (2020). The infectious disease ontology in the age of COVID-19. In OSF preprints. Center for Open Science.
25. He, Y., et al. (2020). CIDO-a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis. http://www.nature.com/scientificdata.
26. Oyelade, O. N., & Ezugwu, A. E. (2020). COVID19: A natural language processing and ontology-oriented temporal case-based framework for early detection and diagnosis of novel coronavirus. Preprint www.preprints.org.
27. Shitharth, S., Shaik, M., Sirajudeen, & Sangeetha. (2019). Mining of intrusion attack in SCADA network using clustering and genetically seeded flora-based optimal classification algorithm. *Information Security, IET, 14*(1), 1–11.
28. Pirouz, B., Shaffiee Haghshenas, S., Shaffiee Haghshenas, S., & Piro, P. (2020). Investigating a serious challenge in the sustainable development process: Analysis of confirmed cases of COVID-19 (new type of coronavirus) through a binary classification using artificial intelligence and regression analysis. *Sustainability, 12*(6), 2427.
29. Mohammed, A., Al-qaness, Ewees, A. A., Fan, H., & El Aziz, M. A. (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. *Journal of Clinical Medicine, 9*(3), 674.
30. Gruninger, M., & Lee, J. (2002). Ontology – Applications and Design. *Communications of the ACM, 45*(2), 39–41.
31. Uschold, M., & Jasper, R. (1999). A framework for understanding and classifying ontology applications. In Proceedings of the IJCAI-99 workshop on ontologies and problem- solving methods (KRR5) (pp. 1–11). Stockholm, Sweden.
32. Maedche, A., & Staab, S. (2001). Ontology learning for the Semantic Web. *IEEE Intelligent Systems, 16*(2), 72–79.
33. Mao, M. (2008). Ontology mapping: Towards semantic interoperability in distributed and heterogeneous environments. Ph.D. dissertation, Pittsburgh University, Pittsburgh, PA.
34. Fortino, G., Parisi, D., Pirrone, V., & Di Fatta, G. (2014). BodyCloud: A SaaS approach for community body sensor networks. *Future Generation Computer Systems, 35*, 62–79.
35. Shvaiko, P., & Euzenat, J. (2013). Ontology matching: State of the art and future challenges. *IEEE Trans on Knowledge and Data Engineering, 25*(1), 158–176.
36. Niles, I., & Pease, A. (2001). Towards a standard upper ontology. In Proceedings of the 2nd international conference on formal ontology in information systems (FOIS- 2001) (Vol. 2001, pp. 2–9).
37. El-Sappagh, S., Franda, F., Ali, F., & Kwak, K.-S. (2018). SNOMED CT standard ontology based on the ontology for general medical science. *BMC Medical Informatics and Decision Making*. Article number: 76.
38. Buitelaar, P., Cimiano, P., Grobelnik, M., & Sintek, M. (2005, October). Ontology learning from text, tutorial at ECML/PKDD, Porto.
39. Aleksovski, Z., Klein, M., Ten Kate, W., & van Harmelen, F. (2006). Matching unstructured vocabularies using a background Ontology. In Proceedings of knowledge engineering and knowledge management (EKAW) (pp. 182–197).

40. Lopez, V., Uren, V., Motta, E., & Pasin, M. (2007). Aqua Log: An ontology-driven question answering system for organizational semantic intranets. *Web Semantics: Science, Services and Agents on the World Wide Web, 5*(2), 72–105.
41. Tablan, V., Damljanovi, D., & Bontcheva, K. (2008). A natural language query interface to structured information. In *The semantic web: Research and applications, LNCS 5021* (pp. 361–375). Springer.
42. Dasgupta, S., Patel, R., Padia, A., & Shah, K. (2013). Description logics-based formalization of Wh- queries, CoRRabs. 1312.6948.