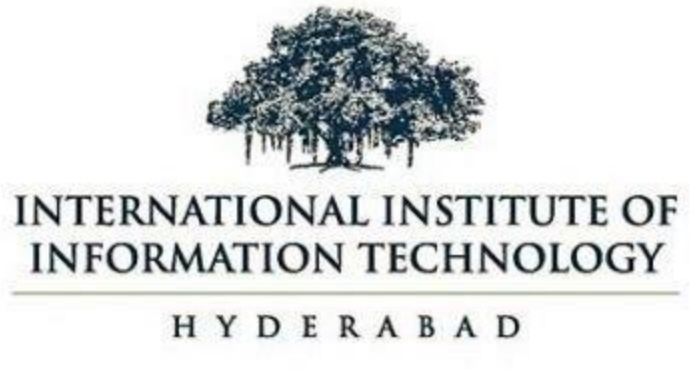
Analyzing COVID-19 Pandemic using SIR model in India and the effect of lockdown

CSE 486: *INTRODUCTION TO NEURAL AND COGNITIVE MODELING*



Project By: -

Arushi Agrawal (2019201015)

Gaurav Chaudhari (2019201045)

Shubhankar Saha (2019201097)

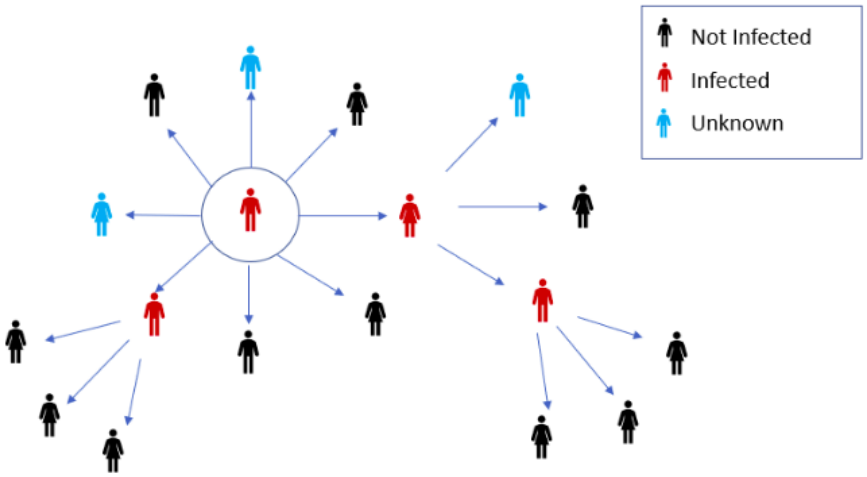
1. Introduction

The focus of the project is to study how social distancing impacts the spread of the corona virus. We have **modelled the covid-19 pandemic using the SIR model** and made several predictions based on that. The project aims on predicting the effects of lockdown in countries like India and US and at the state level. While making these predictions can be tricky, we have considered many factors while performing analysis.

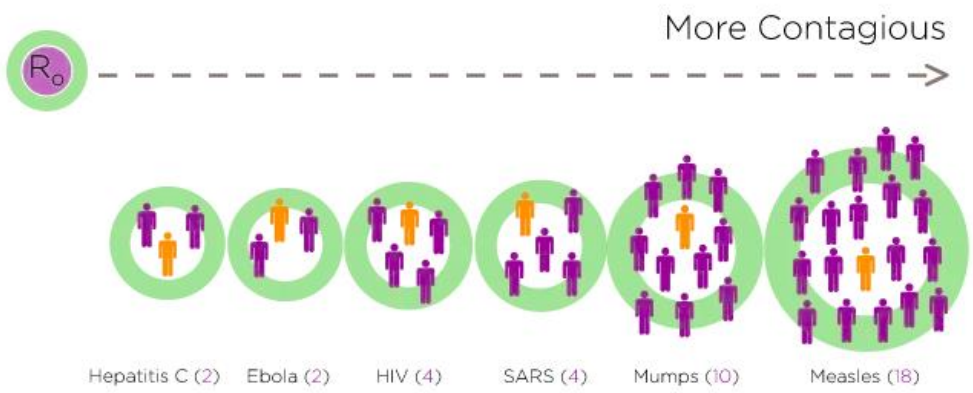
The datasets we have considered for the analysis are fetched from standard resources like government websites and other trusted unofficial sources and have considered as many practical scenarios as possible.

The predictions are based on varying the basic reproduction number Ro and simulating its impact on the spread of the virus using an Epidemic model called SIR.

Exponential disease spread:



The number Ro is a mathematical term that indicates how contagious and infectious a disease is. Say, a disease has a Ro of 6, an infected person will transmit the disease to an average of 6 other people, if no one has been vaccinated against it or is already immune to it in their community.



2. Previous Work

A number of groups have estimated R0 for this new coronavirus. The Imperial College group has estimated R0 to be somewhere between 1.5 and 3.5.

These differences are not surprising; there's uncertainty about many of the factors that go into estimating R0, such as in estimating the number of cases, especially early on in an outbreak.

Based on these current estimates, projections of the future number of cases of coronavirus are fraught with high levels of uncertainty and will likely be somewhat inaccurate.

The difficulties arise for several reasons:

- First, the basic properties of this viral pathogen – like the infectious period – are yet unknown.
- Second, researchers don't know how many mild cases or infections that don't result in symptoms have been missed by surveillance but nevertheless are spreading the disease.
- Third, the majority of people who come down with this new coronavirus do recover and are likely then immune to coming down with it again. As the virus moves into new regions and communities, it encounters people with varying health conditions that affect their susceptibility to disease, as well as different social structures, both of which affect its transmissibility.
- Finally, and likely the most important reason, no one knows the future impacts of current disease control measures. Epidemiologists' current estimates of R0 say nothing about how measures such as travel restrictions, social distancing and self-quarantine efforts will influence the virus's continued spread.

3. Methods

1. What we are going to do is estimate the β and γ to fit the SIR model, hence the Reproduction number R_0 with the actual confirmed cases (the number of infectious people). To solve the ordinary differential equation like the SIR model, we have used solve_ivp function in scipy module.
2. The loss function used in the optimization process was the root mean squared error (RMSE).

3. Deciding the initial values of compartments: SIR model expects the susceptible to be homogeneous, well-mixed, and accessible to each other. Setting the whole population in the country is not realistic for sure.
    a. Since Europe is currently recovering as cases have started to gone down and the same in the Asian countries, we expect 0.25 to go upto 0.30 at max.
    b. Max: 0.003135 Crore; 0.3% = 0.003 Min: 0.0008135 Crore; 0.08% = 0.0008. Also, note that Europe & USA are the countries contributing maximum to the total no. of cases globally so we have tried to keep the figures on the higher side.
    c. The maximum value range turns out to be 4.05 million which turns declaration of lockdown, percentage of people tested positive being low as compared to other countries, the mean age group in India turns out to be pretty much unrealistic right now based on the current stage that we are in, hence we'll go ahead with the manipulation on the minimum value first, so base value that we have is 8% of 135 crore, i.e. 1.08 million. The population of India is ~200 per sq. km. while that of Europe is ~143 per sq. km.
    d. Now, considering multiple factors like the earlier count being much lesser than that of Europe. We assume that only 50% of the current population size can actually get affected, i.e. to 0.54 million. Some factors might be debatable, like people being travel history being the first ones being affected, and then covid-19 gradually spread through community transmission.
    e. People from many under-developed parts of the country don't have the privilege to step outside their place. Also, no one from outside visits their places as there is no industrial benefits available. Mass gatherings are banned at airports, movie theatres, shopping malls or party halls etc.
    f. Also, there's also a theory that Indians might have a stronger immunity towards this virus as they've had slightly more episodes of diseases in general that coutries like UK, USA, other European countries etc. Also, as per the Indian Govt. almost 30% of the existing cases have generated from the religious event that happened in Delhi post lockdown and except that there has been no record of community transmission happening anywhere else in India.

Based on the above considerations and the existing current trend in the number of cases I select my population that could have actually come in contact with corona:

i) Best case scenario: 75000

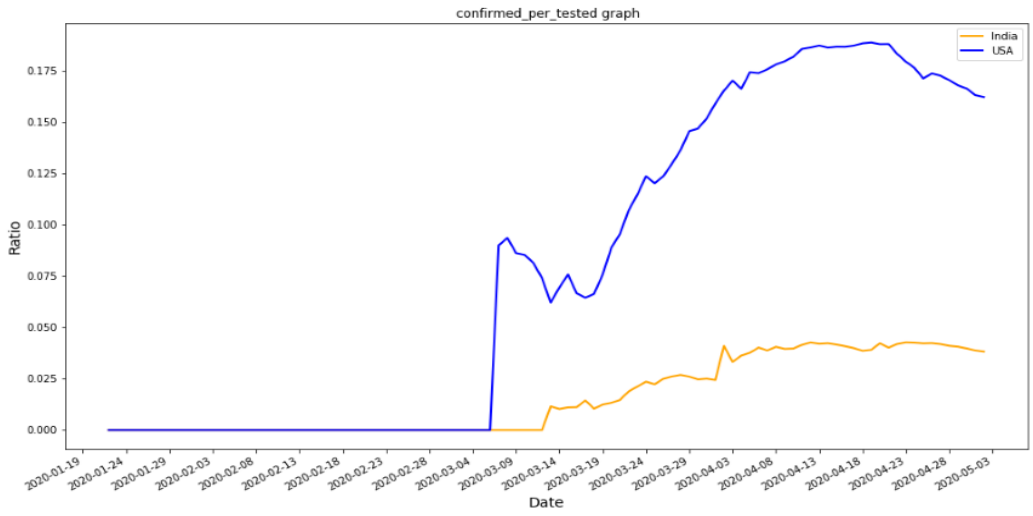
ii) Worst case scenario: 1,50,000

4. Result

1. Comparison between India (strict lockdown country) and USA (late lockdown)

Since testing rates of USA and India varied greatly, hence the number of confirmed cases varied greatly, tough population of India is much higher than USA.

So, to compare them, we used confirmed cases per test i.e. No. of confirmed cases/number of test that day as it'll give us a good idea about the difference in COVID-19 spread across the two nations. Below is the result:



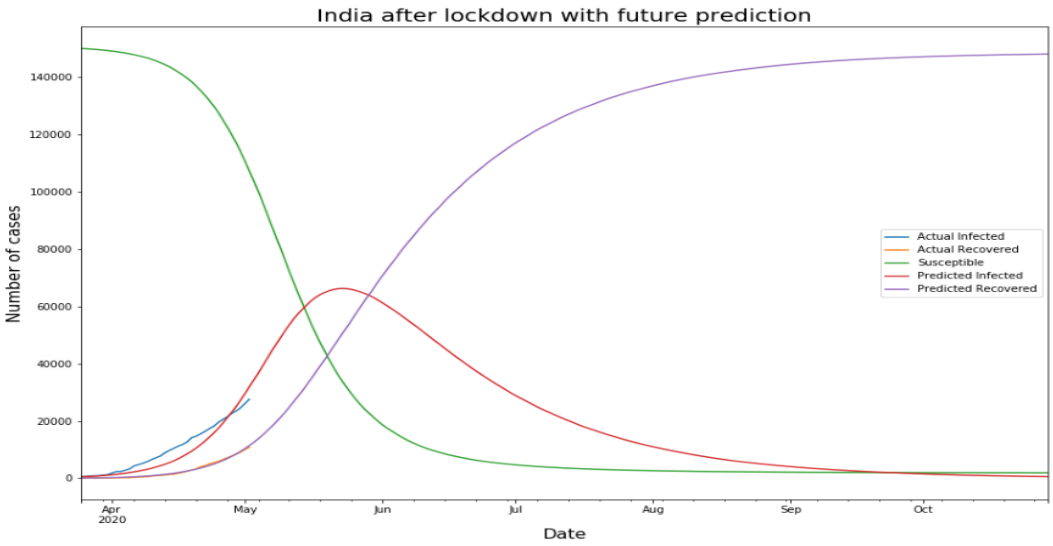
As we can see that the ratio is now going down in US, so is in India after the ascend.

But, due to strict Lockdowns India can keep the graph flatter and hence didn't allow the number of cases to rise too big, resulting in relieving the healthcare system.

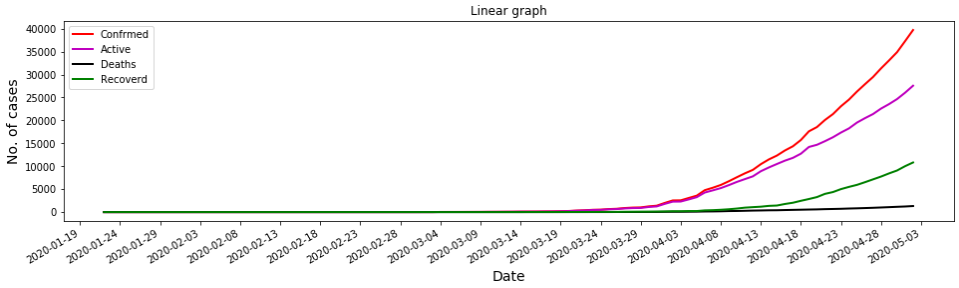
2. After feeding the current data to our SIR model, it fits the data and gives the below parameters.

$\beta = 0.15311306$, $\gamma = 0.03464741$, R0: 4.41917766

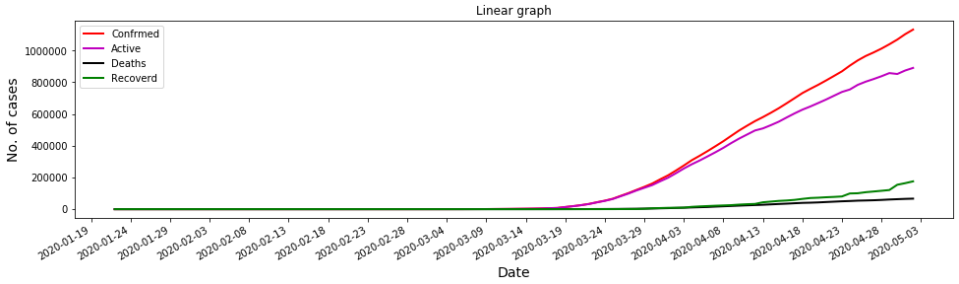
Also, on prediction on future date range, we get the below graph:



Following plot shows growth of covid-19 in India



Following plot shows growth of covid-19 in US



2. Comparing Covid-19 against previous flu-like virus episodes:

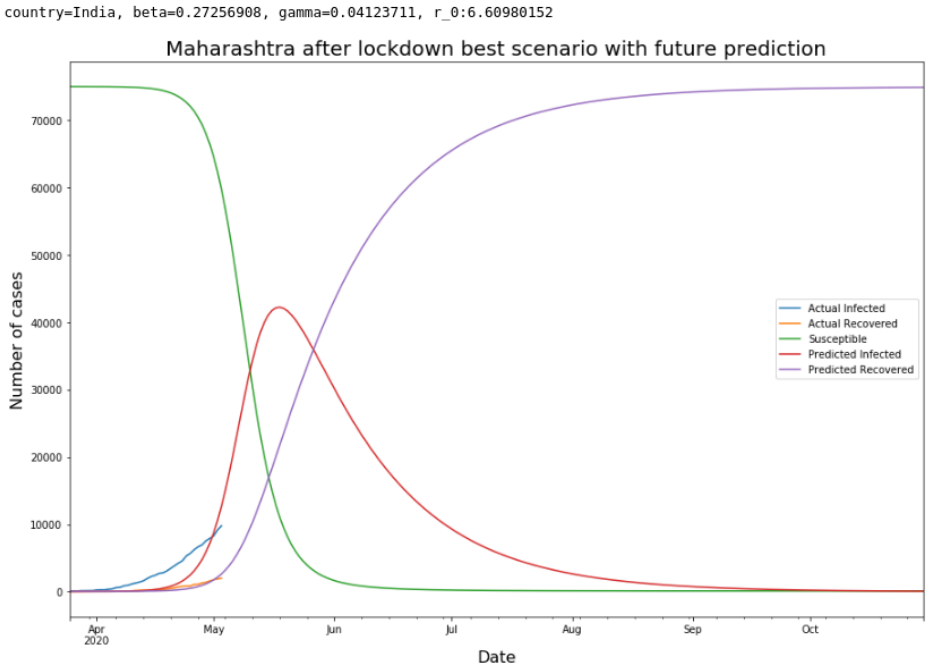
1. Swine flu or H1N1 virus from 2009 had a Ro value of ~ 1.5. The impact was not that much because of vaccines and antiviral drugs. Coronavirus has R_0 of 5.7, which is predicted to reduce to ~3 as an effect of lockdown.
2. The swine flu (H1N1 influenza) pandemic of 2009 had resulted between 1,51,700 and 5,75,400 deaths in the first year according to medical journal "The Lancet". As for Covid-19, the death toll as of 4th may, 2020 is 2,48,000, and the numbers are still growing every day.

3. Covid-19 has affected almost all the countries in varying degrees. According to "worldometer", the outbreak has affected 210 countries and territories around the world. The swine flu pandemic had struck 179 countries and territories.
4. However, unlike coronavirus, whose vaccine is still being worked upon, while H1N1 has a flu shot to prevent its spread.

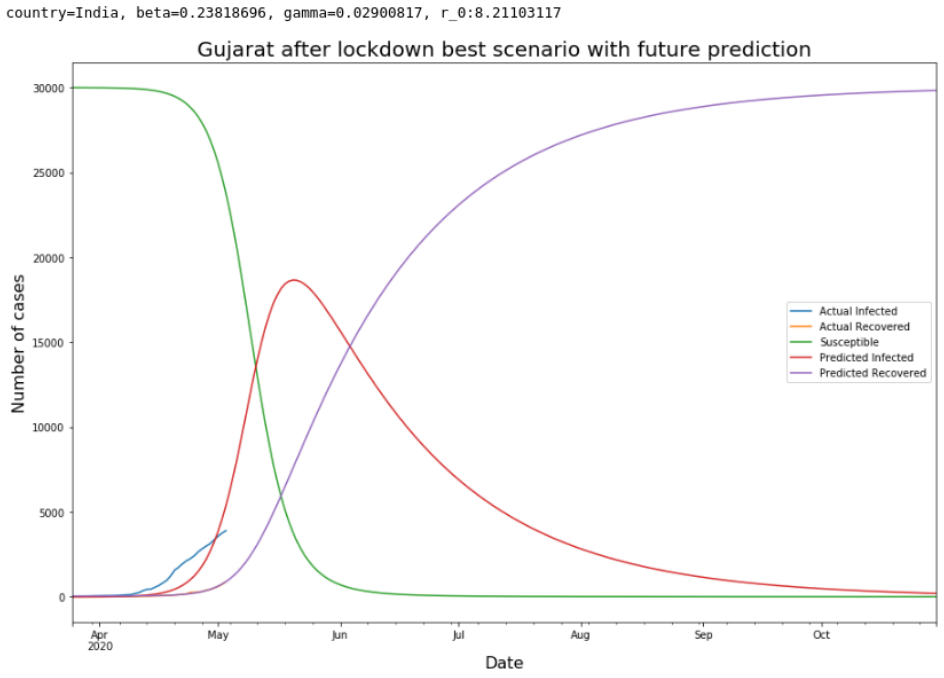
Disease	Reproduction number R0
Ebola, 2014	1.51 to 2.53
H1N1 Influenza, 2009	1.46 to 1.48
Seasonal Influenza	0.9 to 2.1
Measles	12 to 18
MERS	around 1
Polio	5 to 7
SARS	<1 to 2.75
Smallpox	5 to 7
SARS-CoV-2 (causes COVID-19)	1.5 to 3.5

3. Analyzing effects of lockdown state-wise (Maharashtra and Gujarat)

Following is the plot for Maharashtra State:



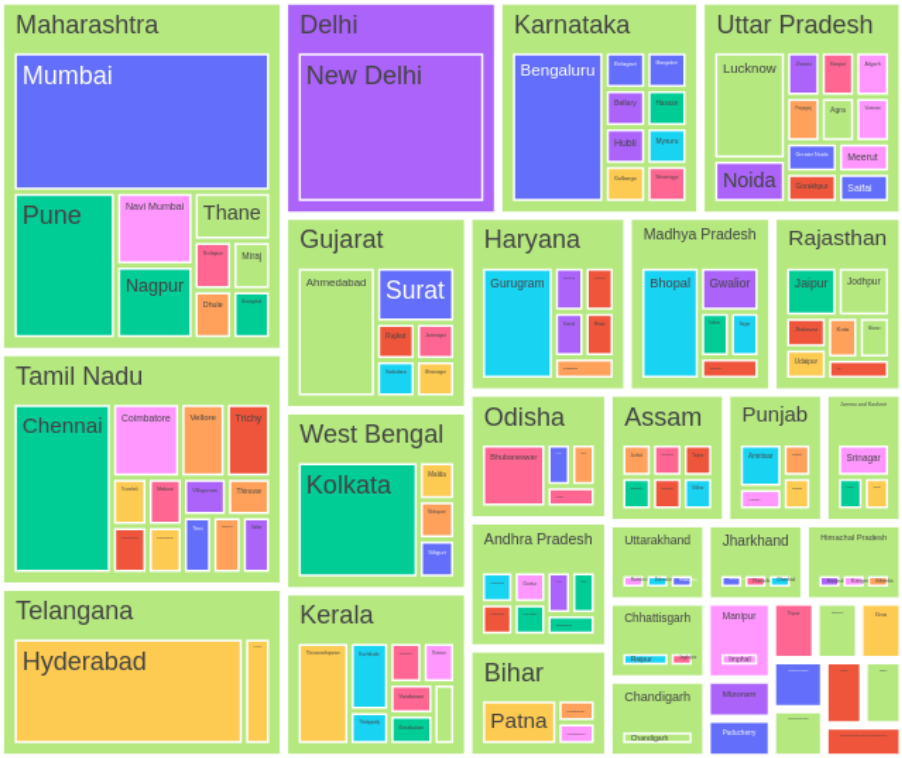
Following are the plots for Gujarat State:



After analyzing the results for various states, we concluded that states like Maharashtra and Gujarat, despite having highest number of Covid-19 cases, can handle the situation quite well. After some analysis of healthcare systems data, we got a supporting evidence for this conclusion, as the health-care system in these states are better developed than other states.

4. Status of health-care system

Following plot shows the summary of the status of health-care system in various states. Bigger the size of the box, better the healthcare system.



5. Conclusion

Since India has been self-aware about its limited resources in healthcare like shortage of beds/ventilators (0.55 bed per 1000 people), testing capacity, PPE kits, etc as compared to the European countries, US and China, Prime Minister of India Mr. Narendra Modi had announced strict national lockdown to delay/slowdown the spread of the virus, as social distancing is the best bet as of yet.

Also, Indian's have mostly been careful in maintaining social distancing (self-isolating even) and staying at homes, unlike majority of the citizens of US and European countries. This timely measure has resulted in reducing the reproduction number R_0 by a great extent. Our model predicts a higher R_0 Value, but Lockdown has been effective and R_0 is way down.

The Delhi incident has accounted for more than 30% of the cases as per government. Had it been prevented, we could have flattened the curve even better.

Prediction by the Singapore site is given below, they are very encouraging, and predicts that the pandemic will end 97% by the beginning of June.

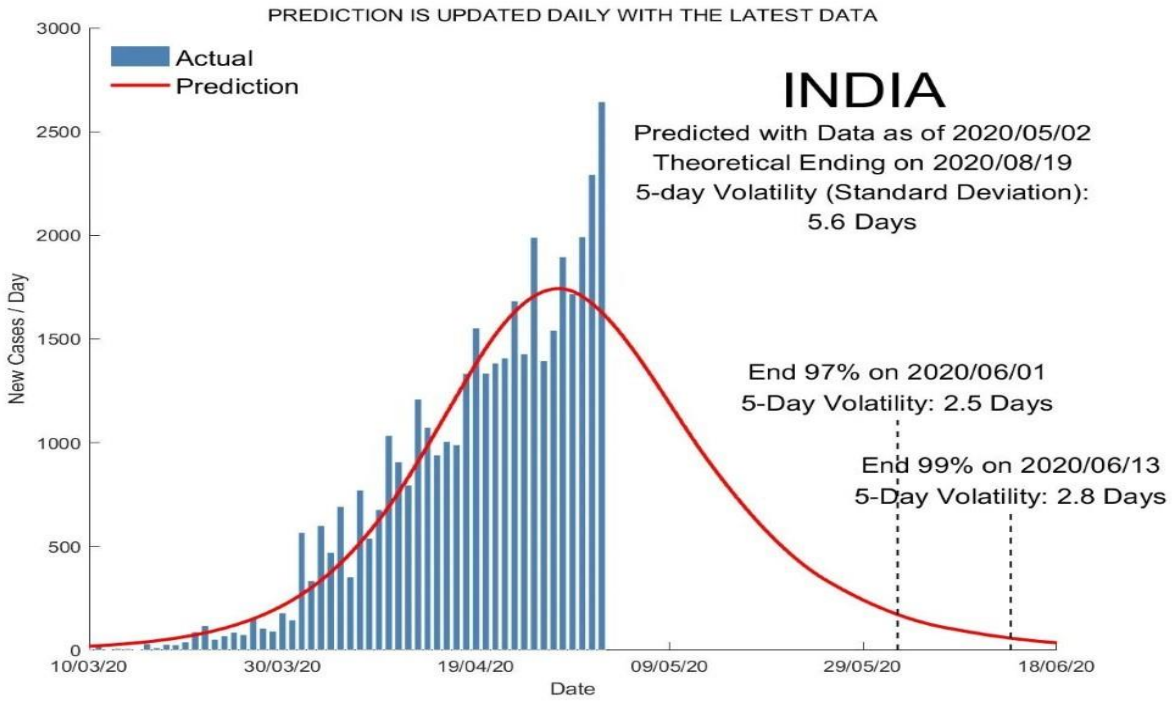
But our model predicts that it would be August for this pandemic to minimize.

Few observations:

1. The Singapore site uses a Gaussian distribution type curve to predict the End of Pandemic, although the SIR model creates a curve which resembles Poisson distribution.
2. Our model gives a slightly high value of R0, which results in the curve ending late.

Due to the many uncertainties associated with this virus and its unpredictable nature, the social structure of a location, government measures, healthcare facilities, policies, etc., there is no concrete value of R0, and its predictions lie in a range of values

Owing to the above analysis, we think that the spread of the Covid-19 virus strain will be restrained by 97% in the mid of July, considering the flaws in both the models.



6. References

1. https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology
2. https://www.lewuathe.com/covid-19-dynamics-with-sir-model.html
3. https://ddi.sutd.edu.sg/
4. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/
5. https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases

Datasets taken from:

6. https://covid.ourworldindata.org/data/owid-covid-data.csv
7. https://pomber.github.io/covid19/timeseries.json
8. https://raw.githubusercontent.com/covid19india/api/gh-pages/raw_data.json
9. https://api.rootnet.in/covid19-in/hospitals/medical-colleges
10. https://api.rootnet.in/covid19-in/hospitals/beds