CORONAVIRUS
PANDEMIC

Team Name: Symbians

Team Members :

1. Santosh Saxena

2. Gayatri Pingale

3. Nandita Jassal

4. Mohit Virmani

Institute Name:
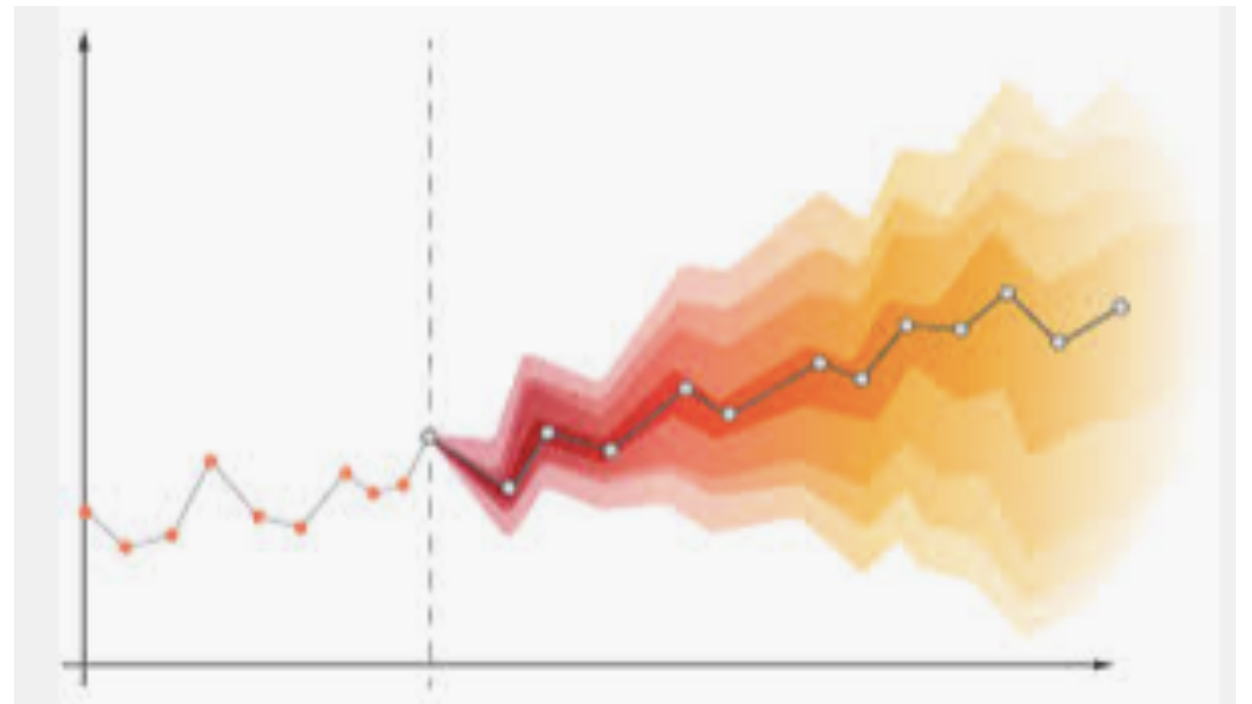
Symbiosis Institute of Technology

Topic:

Pandemic Forecasting

Qualification :

B-Tech 2nd year

Symbiosis Institute of Technology

Runtime Pandemic Forecasting

Pandemic Forecasting

Abstract **_: Todays world is suffering with one of the deadliest virus known as Corona Virus or n-COVID-19. This virus give major damage by its rate of infection. Hence for preparation we need some statistical information about this rate of infection so that we can plan for our next step. To apply machine learning, we don't have enough data with available data we can make wrong prediction also. Hence it is very important that we cover all the parameters that is required to predict the rate of infection of such deadliest pandemic._**

Introduction

Corona virus has some pattern , Which is seen in all the country. It has few stages such as stage 1, stage 2, stage 3, stage 4 , stage 5. In which each level indicates

stage1 infection is because of international citizens . In which citizens which came from outside of India brings infection and spreads infection to Indian citizens. This stage doesn't have any mortality rate

stage2 infection is person to person transmission which has some range of infection and can give some casualties.

stage3 infection is community to community transmission which cause wide range of infection and after that Level increases on the basis of casualties

The man part is that the infected person doesn't have ay knowledge about his infection under 14 age hence after 14 days all cases appears at the same time. Let's discuss about the architecture on which we are working.

Architecture

Taking input from user

Parameter

Process

Predicted output
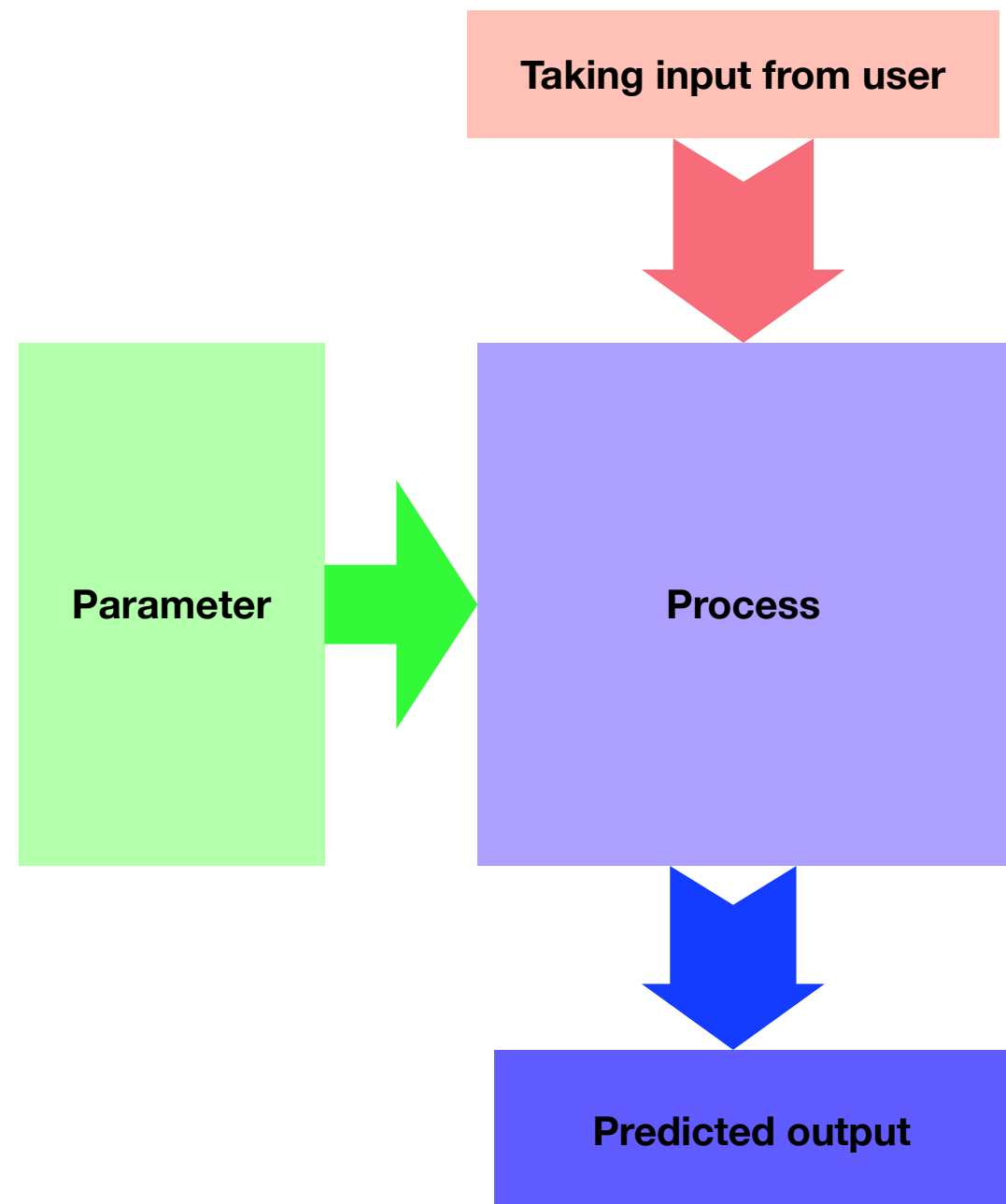
Methodology (Process)

The expected graph of rate of COVID-19 with respect to no of days will be
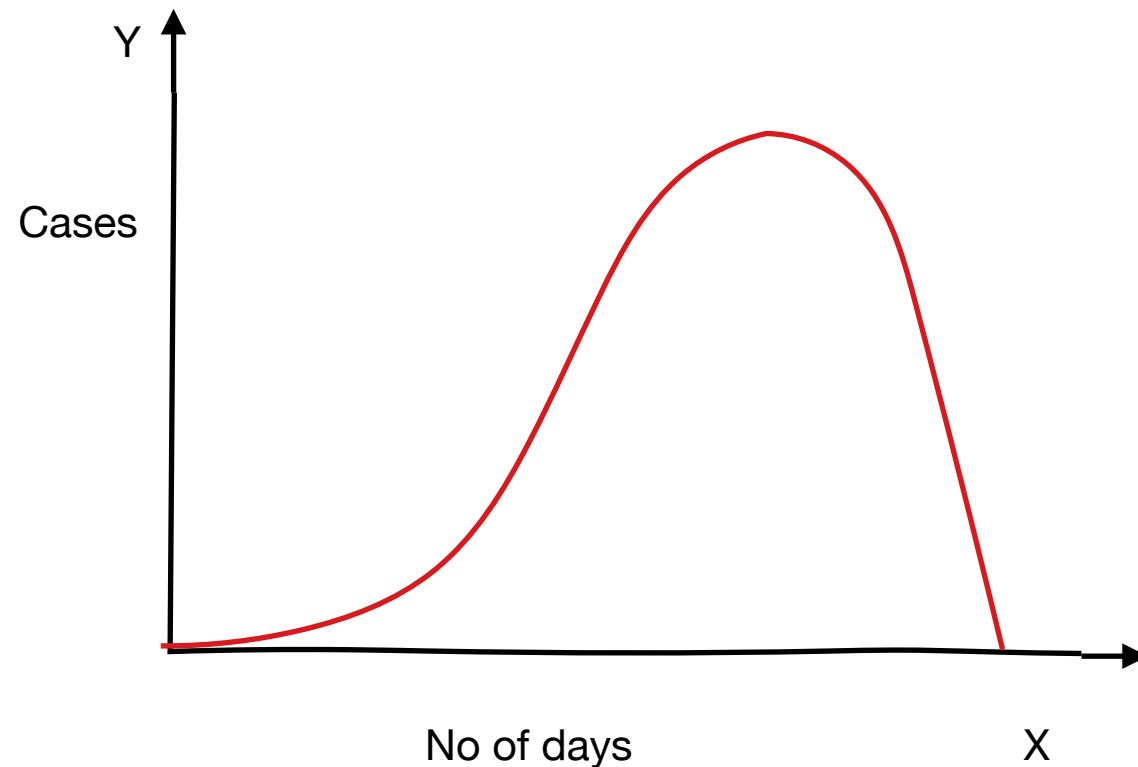
The reason behind this graph is that none of the country has completely won with this corona virus. Hence it is just infection period . The another phase is a prevention period in which the number of cases will go down and the recovery rate starts. This can be possible when the vaccination of the disease will invent. But the thing is the vaccination requires time and till that we have to avoid the growth of infection and try to maintain the figure which is present at the peak. Where the number of cases is minimum. To make plan for minimisation we need to check the idea. Which can be done with the help of this model.

Model

1) Probability

Let's consider 2 sets P(A) and P(B) to calculate the probability. Where P(A) stands for probability of Infected person and P(B) stands for probability of Not infected person. The Venn-diagram will look like

Venn-Diagram
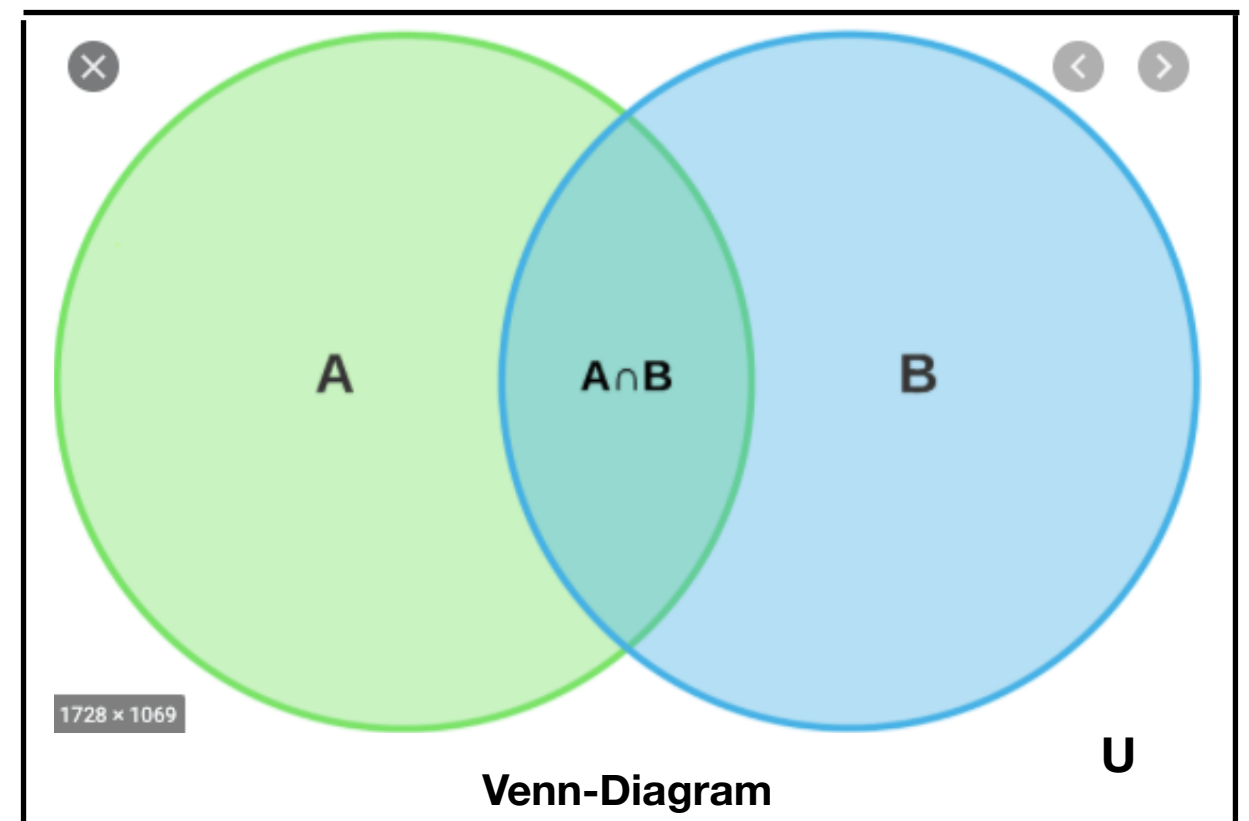
Here P(A ∩ B) are nothing but the unknown probability of infected citizens. This P(A ∩ B) will become P(A) after 14 days and   another   P(A ∩ B) will form. The recovered person will go from set A to B and the casualties will go in the Universe U.

Now set P(A) and set P(B)is known to us but P(A ∩ B) is

The Mathematical equation of probability of infection under the condition with P(A) will be

$P(B/A) = P(A \cap B)/P(A)$

Here,

$P(B/A) = P(A \cap B)/P(A)$

$P(A) = $ Given

$P(A \cap B) = $ Unknown.

To Calculate $P(A \cap B)$

  The first question to find $P(A \cap B)$ is, How many person can a single person can infect? The answer to this question is that it depend from person to person. Hence we need to make levels for that. Level1 indicates that the person is in home and the chances of infection is with their family members. Level2 indicates that the person has gone outside for some work. Level3 indicates that the person has gone in some crowded place . Hence the calculation for a single person will be

Level1 = X*1 (this value will be approximately equal to 4)

Level2 = X*1(the value of X will come from regression part)

Level3 = X*1(The value of X will come from regression part)

This calculation is for single person. For multiple person the equation will become

Level1 = X*m

Level2 = X*m

Level3 = X*m . And so on.

$A \cap B = \sum (\sum \text{Levels}) \ (i)$

$P(A \cap B) = (A \cap B) / \text{maximum chances of infection}$
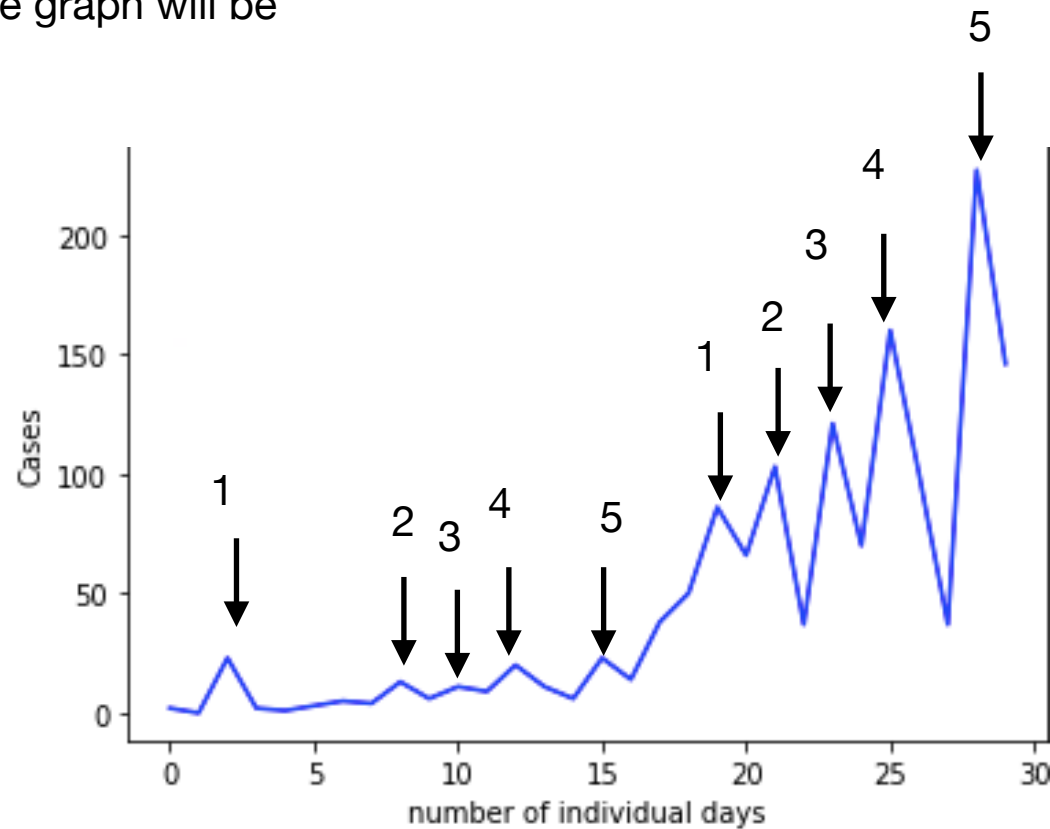
Hence ,

$P(B/A) = P(A \cap B)/P(A)$

2) Regression

  This is the data which we have up till now. These data has 3 stages as we discussed in the Introduction. We need to compute the parameters of stage 1 and we have to try to find the reaction between stage 2 and stage 3. But first we need to know what is the individual number of cases/day.

The graph will be

This graph has Peaks because the person who get infected on the first day has detected on the 14th day. Hence because of this continuous process the graph has peaks. We need to find the relation between the stage 1 peak and the rest of the stages peak.
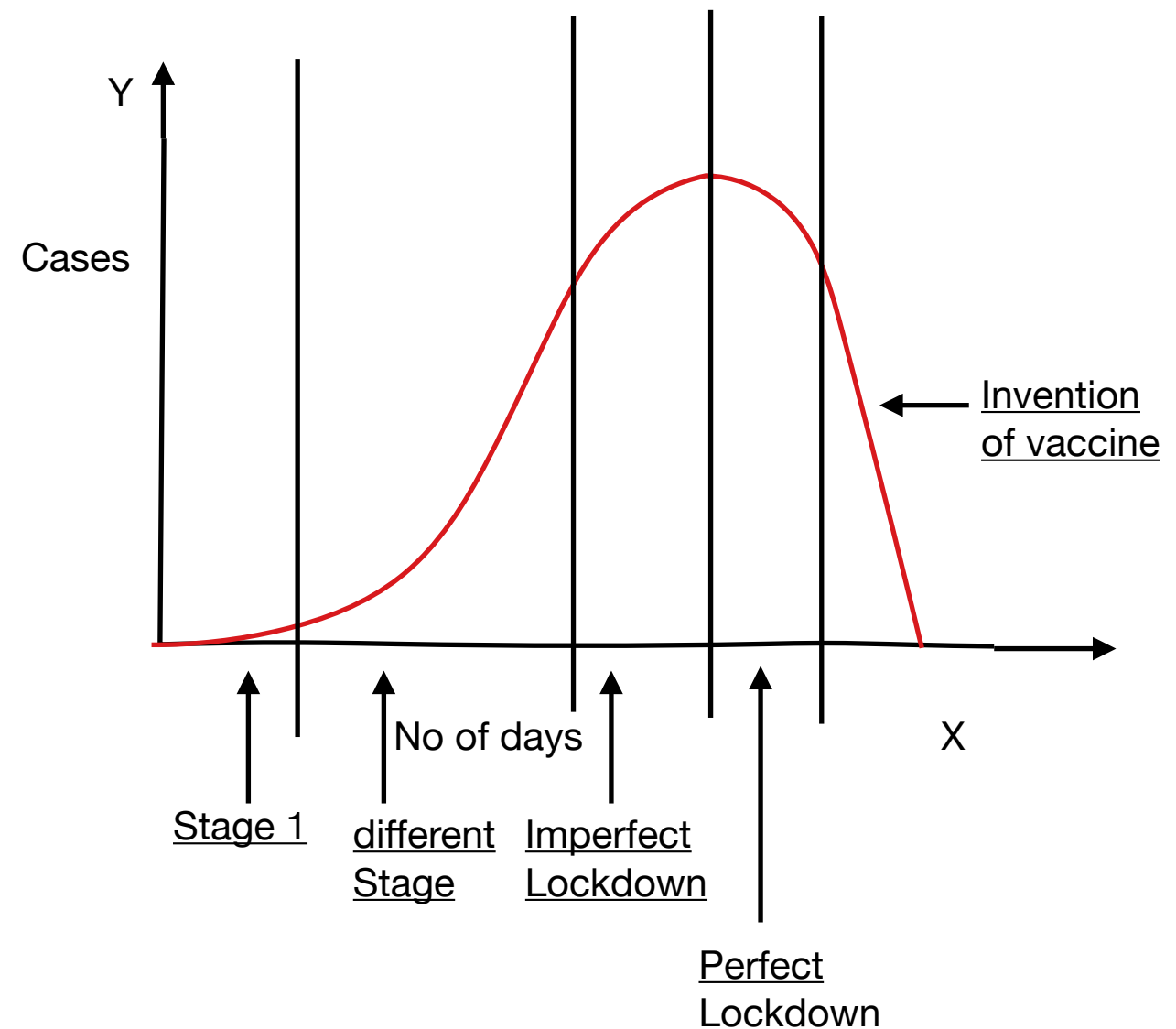
Computing of all the parameters will be

Number of cases = Q + (X*Stage1*P(A/B)) - recovered - dead

To Compute number of Recovery Rate.

P(Recovery rate) = (No of success/Total number of cases)

P(Death rate) = (No of death / Total number of cases)

Q will be calculated in the regression model i.e coefficient of particular stage

Our model will make run time decision. It will take the inputs from infected person about its activity. And according to that we can predict the number of cases after 14 days. Other than that If we want to predict for today , our model will take the data of previous 14 days and will process accordingly

Stage 1 : It is the initial stage from where the infection starts in the stage 1 the graph will be linear in nature. The number of cases will be limited to calculate P(A ∩ B) . Maximum cases will be of Level 1 and 2.

Different Stage: In this case the infection tries to spread with some wide extent and has some exponential graph. Actually our equation is A = A + (A ∩ B) in which A is still linear but A ∩ B converts the graph from linear to exponential. The calculation of A ∩ B will be all Level. Because there is no lockdown and infected people are spreading infection in the crowded places also. Hence all level will be useful.

A ∩ B = X * P(A ∩ B)

Where,

X is the maximum value of infection.

0 <= P(A ∩ B) <= 1

Imperfect Lock down: This is the period when the government has declared the lockdown. But the essential shops are open. Hence there is a scope of infection. Thus Level 1 and Level 2 will be there. At the same time Higher Level will also be present for the case where people are intentionally trying to spread the infection or illegal gathering and some time on the basis of fake news.

Perfect Lockdown: After government has come up with some great strategy which causes 100 % accuracy in terms of social distancing at that time number of cases will stop . The graph will go down because of recovery of confirmed patient. That value will also be generated in our model.

Invention of vaccination: We should not forget that the lockdown is a temporary process to avoid the casualties and doctors can work peacefully to invent the vaccine of this virus which will take almost a year. Thus after the invention of the vaccine the cases will fall and world will back to normal. Hence the last part of the graph is linear with negative slope.

Implementation

For Implementing ,we will use python language. Our first step will to classify the levels.

For eg: Suppose

Stage 1 = 1

Stage 2 = 100 * stage 1

Stage 3 = 1000 * stage 1

The value in this example i.e 1 , 100 , 1000 will be calculated in our model on the basis of statistics.

Then we will calculate A ∩ B on the basis of levels

For eg : one infected person has gone in a grocery shop and present at home. For this case the calculation will be

A ∩ B = 4*1 + 4*10 = 4 + 40 = 44

For multiple cases A ∩ B = ∑ (Level1 * n1) + (Level 2 * n2)+ …

Where Level 1 is 4 and Level 2 = 40 and n1 and n2 is no of infected person

The next step is to find P(A ∩ B)

P(A ∩ B) = (A ∩ B) / maximum number of cases

In eg P(A ∩ B) = 44 / 100 where maximum number of cases will be maximum number of the highest Level for that particular case.

After this we can fine P(B/A)

Then the next step is to find the parameters of the stage 1 hypothesis.

Suppose x is input , y is output and spread is predicted output

ypread = m * x (stage 1 is a straight line starting from 0)

$Min(m) = \sum (y - ypread)i^2 / 2$

Hence our final equation will be

ypread = (m*x) + (stage no * P(A/B)) - recovered - dead

Ypred = (m*x) + (X* stage 1 * P(A/B)) - recovered - dead

The outcomes

Firstly the authorised person will add the data into the machine of previous 14 days cases and their activities which is present in our model. The output which you will be the predicted number of cases of after 14 days. This is the case if you want todays prediction. At the same time by adding todays data you can get the number of cases of

after 14 days. Hence the range of prediction is up to 14 days. The reason behind this is that it is a runtime prediction. That is the only possible range because if you directly apply machine learning the graph will never come down and you will get wrong predictions. It also dependent with the actions taken by the government. We can actually verify our plan by seeing predicted results and actual results if machine is trustable.

Advantage

1. You can add each and every case. You need not prepare another model for any special case the levels are already given.

2. You can get run time prediction under the range of 14 days

3. This model can be add on if you also want sensors to detect the activity hence can become part of an IOT.

4. Government can check their plan before implementing it

Requirements

1. Anaconda environment for Python with 3.7 version

2. Libraries like numpy , pandas and matplotlib which is already available in anaconda environment.

3. We are just focused on the output part hence , we will get output on our notebook only.

Data Preprocessing

1. I took the data from covid19india.org. Hence at that time only the useful data is converted into .csv file and loaded into the code.

2. The data is converted from everyday data into data of individual day.

3. Then data is ready for regression and calculation of other parameters.

4. The data visualisation and other section is done in the coding part.

Reference

1. https://www.brookings.edu/blog/up-front/2020/03/27/covid-19-does-india-have-enough-doctors-an-analysis-of-growing-covid-19-patients-and-existing-medical-capacity/
2. https://www.covid19india.org/
3. https://economictimes.indiatimes.com/industry/healthcare/biotech/healthcare/586-covid-19-hospitals-with-1-lakh-isolation-beds-over-11k-icu-beds-across-country-health-ministry/articleshow/75097529.cms?from=mdr
4. https://www.worldometers.info/coronavirus/?
5. https://www.mygov.in/covid-19/?cbps=1
6. https://zeenews.india.com/
7. https://www.indiatvnews.com/

Thank you