Epidemic Computational Model using Machine Learning for COVID-19

Sanjeev J. Wagh[1]
Department of IT/CSE
Govt. College of Engineering,
Karad
sanjeev.wagh@gcekarad.ac.in

Chaitanya S. Wagh[2]
Department of Computer Engg.
SKN College of Engineering,
Pune cwagh12@gmail.com

Parikshit N. Mahalle[3]
Department of Computer Engg.
SKN College of Engineering,
Pune aalborg.pnm@gmail.com

Abstract:

The world is under tremendous pressure and lock down in major giant countries due to ongoing coronavirus COVID-19 originated from Wuhan city of China. Corona is the name of a virus group that is found in humans and animals. They range from simple cold coughs to coronary viruses to serious illnesses such as SARS or MERS. The corona virus found in Wuhan city , China (2019), is different from the six to seven viruses previously reported in humans. So it's called the Novel Corona virus (COVID-19).

The Module which is popular for pandemic prediction diseases Susceptible, Exposed, Infectious, Recovered i.e. SEIR is the modified model of SEI which is the simplest epidemiologic model presents the reality of epidemic growth with reference to the populated cases. In this paper, the SEIR computational model analyzed by considering the total recovered cases again get 10 % susceptible. Also the interventions of authority get flatten the susceptible population by 40 % due to remedial intervention of concern authority. The results achieved are predicting that the peak of Infectious population will be in median of June in India.

Keywords: Computational model, COVID-19, Epidemic disease, Prediction, etc

Introduction

On January 30, 2020 the first case of COVID-19 reported in Kerala, India. According to the Ministry of health as on date May 6, 2020 the country has recorder 49,436 positive cases across India [10]. Out of which 14,183 were recovered and the death toll is raise to 1,695. Out of total closed cases 15,878, the recovered and discharged cases are 14,183 (89%) and 1,695 (**11**%) death cases as of 6, May 2020.

By considering first reported case in 30 January 2020, designated as Day 0, Phase 2 started at around Day 45, as per the regular cases reported and now India is started the shaping peak graph as per the major affected states like Maharashtra and Gujarat, Delhi, etc and reaching in between stages 2 and 3.The major hotspot identified in Mumbai, Pune, Delhi and some cities in UP and other states.

The number of cases now increased to peak rise since the outbreak emerged in India around 3 months. As per the review from authority, its remarked that the virus transmission in the country is still in second stage. But as per the media and social media sources, the opinion is differ as incidence happen due to congregation (more than 2500) in New Delhi, and pointed to the big spreader of coronavirus disease. The member of mass congregation after attending the meeting, travelled to various parts of states and cities including rural places result to exponential spread of disease. Out of total positive cases detected, 50 % and more due to this incidence happen as per the media source.

Notion used here are susceptible population represents *S,* and exposed population represents *E*, infected population represents I which can further categorized into two groups, I_H is hospitalized and I_S , is free

living in society and not isolated. I_H represents those that are infected and isolated, and IF represents that are infected but not isolated i.e. unreported cases. Thus, there is no chance to spread the virus from IH but major chances to spread the viruses are through I_C.

Motivation

In beginning with limited testing kits and few laboratory was available. Now Indian government have given permission to private 500+ laboratories after accreditation from ICMR and now 2500+ test taking places daily. The proposed model focuses on the populations, infected hospitalized IH and populations that are voluntarily testing from society IS. There are two possibilities after Infections: either recovery or death. These populations are represented by R *i.e* Recovered and D *i.e.* Death. The most of the recovered population may acquire immunity to the virus but as per cases in China, 10% are the cases are again susceptible to viruses hence considered here as addition in newly susceptible population. The status of coronavirus in upto 15 in India shown in Figure 1,2,5-8 [10]
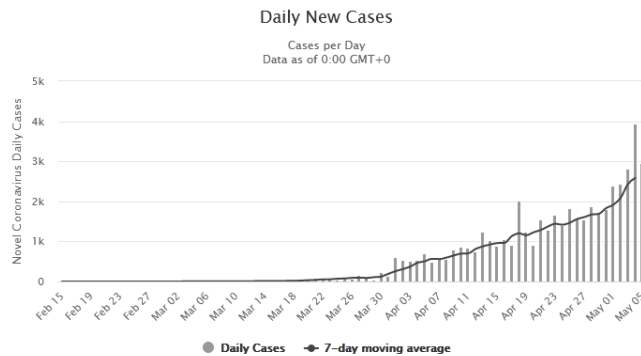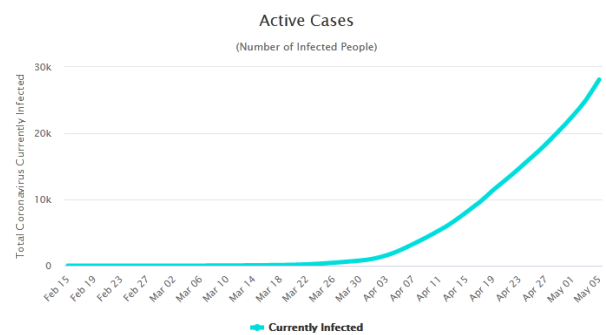

Figure 1: Daily reported cases

Figure 2: No. of infected active cases

Related Work:

The review by [1]E.C. PIRTLE and G.W. BERAN attempts to consolidate data from reported studies on the survival of viruses in the environment, also importance of research on numerous viruses survives in the environment. The paper focuses on variety of information about infectious agents with proper classifications based on natural behavior, and the need for more comprehensive studies in the future.

The various AI-driven tools are explored by KC Santosh in the web literature [4] to identify COVID-19 outbreaks and their prediction about spread across the world. The AI-driven tools will be considered to have more severe wider in scope for active learning-based cross-population training and testing. MERS-CoV and SARS-CoV are not occurred or dependent on a particular season, but infectious diseases human coronaviruses and Betacoronaviruses display seasonal patterns in their incidence [5]. The authors verified the data from different cities regarding community spread of COVID-19 using ERA-5 reanalysis. The data has been compared with the cities, areas which are affected or less affected or has no significance of disease spread

Li Yan 1 & Et L [2] work focus on use of machine learning method in COVID-19 prediction and diagnosis as per basis of need of the large scale explosive epidemic COVID-19 cases. More advancement

is needed to consider clinical clarity to avoid confusing factors and to increase the datasets for further improving the model. Cleo Anastassopoulou ,et.l [3] considered actual infected individuals with asymptomatic or mild courses and recovered cases in the population. At actual confirmed infected and recovered cases are considered instead of official numbers, the number of deaths were constant. The computations and analysis is based on a mean field of Susceptible, Infected, Recovered and Death (S-I-R-D) model. The concern claims about how successfully developing simulation model using Compartmental models in epidemiology for contagious diseases such as Ebola and Chikungunya is useful. Also SIR and SEIR compartmental epidemiological models successively simulated in previous study the spread of pathogens. [10, 11, 12]

Methodology

Since the popularity and comfortable of SEIR paradigmatic model for mathematical epidemiology, the computational model to simulate COVID-19 outbreak, SEIR model is used [11]. The advantages of SEIR variant models are that, the transparency and accuracy can be achieved and computed through mathematical simulation of biological processes regarding the epidemiological assumptions. By the comparisons of simulated model results and observed patterns, the disease epidemiology is tested.
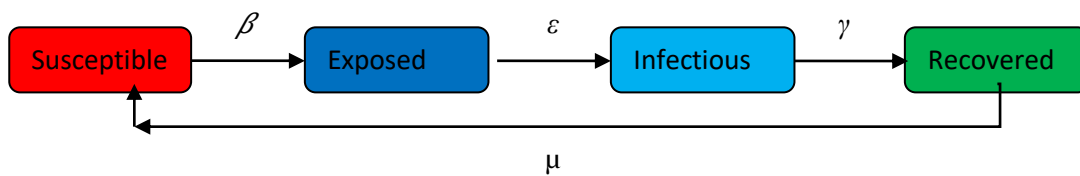

Figure 3: SEIR Model

The model is based on the behaviors of how an epidemic of contagious diseases occurs in the real world. The model composed of four categories of individuals. S-Susceptible – those who are susceptible to disease, E – Exposed – those who exposed to disease, I – Infectious – those who are infectious and may spread the disease, R – Recovered – those who have recovered from previous infection and may not spread or catch the disease in future.
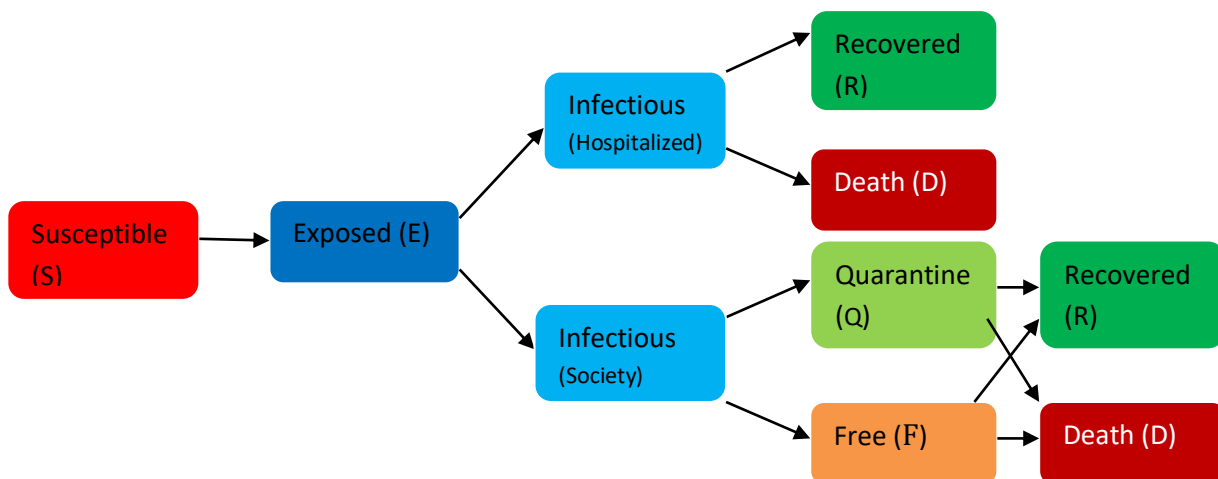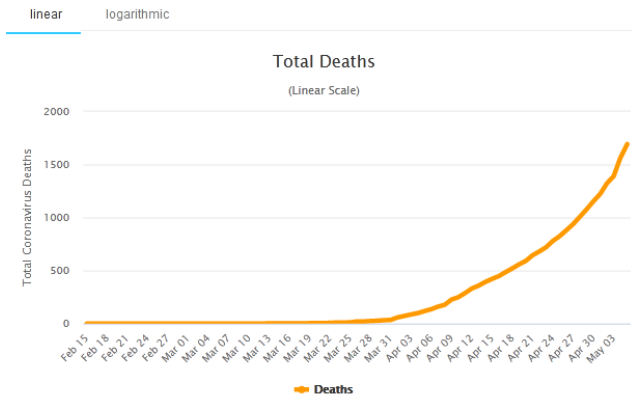
Figure 4: Modified SEIR model
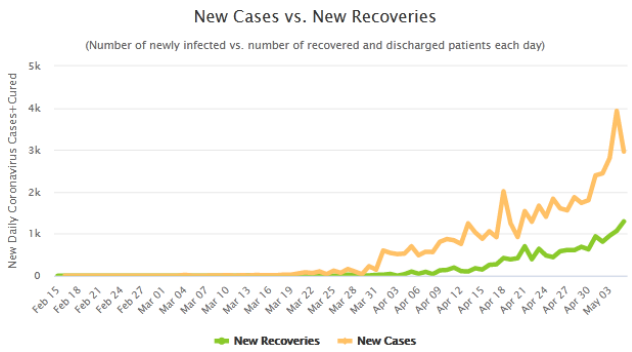

Figure 5: Total death cases

Figure 6: Total daily deaths

Figure 7: Newly infected cases Vs. Newly recoveries
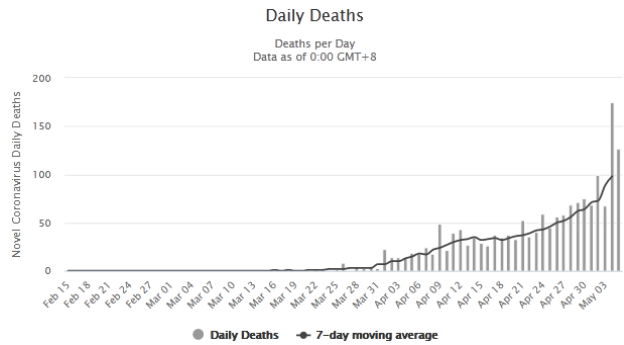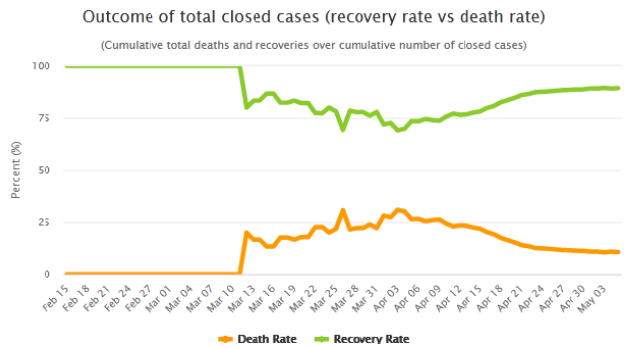
Figure 8:Outcome of total cases

To illustrate the contingency of computational model based on the available population datasets, differential equations are used [6]. The occurrence of disease transmission takes place as the susceptible come in contacts of infective neighbor(s) and outbreaks when infected population merged among susceptible. From some of the susciptables, some individual become infected and others may colligate the recovered population. By assuming the changes are continuous, the following equations is described as follows, Where N is total population,

The total population, $N = S + E + I + R$

$dS/dt = \beta*(S*I)/N$.................................(1)

$dE/dt = \beta* (S*I)/N - \varepsilon *E$ ....(2)

$dI/dt = \varepsilon E - \gamma I$ ....(3)

$$dR/dt= \gamma I \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(4)$$

The infectious rate , which controls the rate of spread i.e. chances of transmitting disease between a susceptible person and an infectious person. The incubation rate ε-, which is the rate of latent individuals becoming infectious i.e. average duration of incubation, is $1/\varepsilon$. Recovery rate γ, which is $1/D$, is compulsive by the average duration which is D of infection. The recovered individuals return μ to the susceptible statue due to loss of immunity.

Table 1: variables and values

Variable	Meaning	Values
α^{-1}	Incubation period	5.2
γ^{-1}	Infectious period	3 days
$R0$	Initial phase	1
β	Infectious rate	0.21
ε	Incubation rate	0.1
γ	Recovery rate	0.154
μ	Recovered returns	0.1

Now the total populations that are purpose for virus transmission will be

$$N=S+E+I_H+I_S+R_H+R_S+Q+F$$

$$dS/dt=-\beta I_C S/N \qquad dI_H/dt=\alpha\rho E-\gamma I_H \qquad dR_H/dt=(1-c)I_H \qquad dD_H/dt=\gamma c I_H$$
$$dE/dt=(I_C{*}S)/N -\alpha E \qquad dI_H dt=\alpha(1-\rho)E-\gamma I_S \qquad dR_S dt=\gamma(1-c)I_S \qquad dD_S dt=\gamma c I_S$$

Model Implementation:

In this work, we retrieved the number of reported coronavirus positive cases (dataset source: Ministry of health and family welfare India), the number of infected population considered till 15 April 2020, Mohfw [12]

The model estimated in three phases:

Initial Phase: No action on infected population

Preprocessing Phase: First aid action on infected population and quarantine with further remedial plan

Lockdown Phase: Most Schools, events, public area is closed for particular period.

Reference to Lipsitch et al design approach, [3] the population (N) is modeled of 1 000 000 individuals, and considering approx population to India $N = 13.80 \times 10^8$. That is, $\gamma(0) = pI(0) \times 13.80 \times 10^8 = 1$, Where $\gamma(t) = pI(t) \times 13.80 \times 10^{8.}$

The dynamics of transmission and epidemic progression is simulated considering assumptions such as, at the initial stage, includes all population as susceptible, all susceptible and infectious leads to infectious, the infectious and susceptible are similar, no policy to intervene in early stage of epidemic, Initially originated from single case, patients once confirmed are quarantined.

Considering R0 as contact rate, which is not possible to estimate at initial stage R0=1 indicating neutral production rate. We can use $R0$ and γ to estimate β using the relation $\beta=R0\gamma$. The initial infected population is not static over period. $R0<1$ represents spreading infected population less and $R0>1$, represents more spread.

$$P(t) = -\beta S(t)I(t)/N + \beta_t S_{is} + \mu \ldots\ldots\ldots\ldots\ldots\ldots(5)$$

Where $\mu = R/10$

$$I(t) = \beta * S(t) * I(t)/N - \gamma(t) * I(t) \ldots\ldots\ldots\ldots\ldots(6)$$

$$S(t) = -\beta S(t)I(t)/N \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(7)$$

$$R(t) = \gamma * I(t) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(8)$$

Based on SEIR model [4], the R0 can be calculated by

$I = I_H + I_S + Q$
Q is Quarantine cases at time (t) & I is infected population.

$$R = 1 + I/\gamma + \varepsilon(1-\varepsilon)(I/\gamma)^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots..(9)$$

Results and Discussion

On the basis of recent literature work and dataset released by Indian authorities and official media, the assumptions of the parameters are established, and the mathematical expression for COVID-19 based on the modified SEIR model is simulated.

In this model we haves considered susceptibility, Infected and recovered data from the available dataset from India up to 15 April 2020, the *β= 0.21, referred from {6], γ= No of days/recovered=0.154, t = 210.*
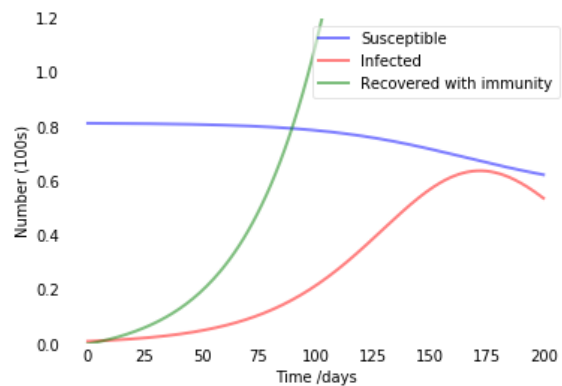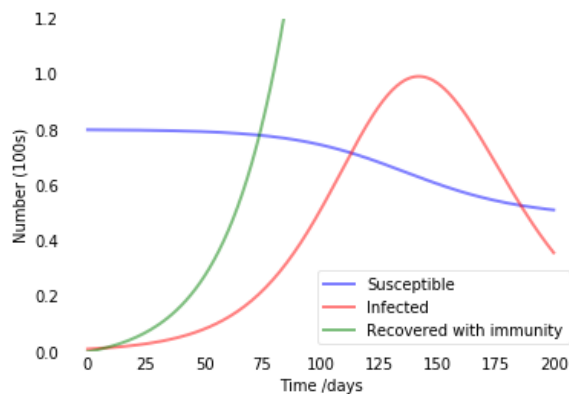
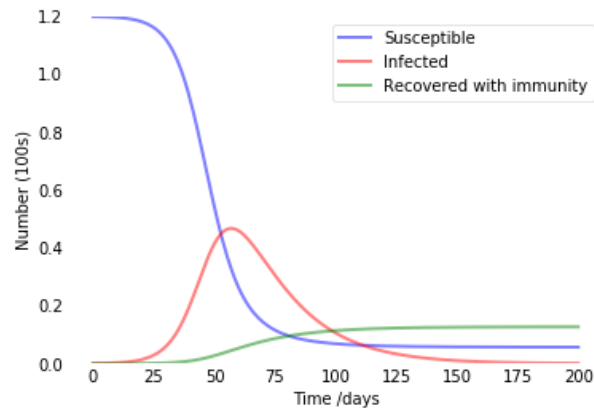*Figure 7: Current Population model*   *Figure 8: Predicted model after intervention of authority*

Figure 9: 10% recovered susceptible

Here we estimated that the rate of reproduction of infected population is $\mu = R/10$ and $I_0=1$, As per prediction the epidemic peak could possibly reach the middle of June by simulating the SEIR compartmental model to the daily reported cases of COVID-19 in India till May 6, 2020. The percentage of uncertainty can be the effective intervention by concern authority with respect to improving susceptible. The target aim is to fast recovery of infected population and may be improved by improving policies like seriously execution of lockdown, sanitization, social distancing and more testing measure.

The SEIR model that was used for our predictions can be easily applied to other contagious diseases with similar transmission patterns such as Whooping cough, Influenza, Ebola and Measles. Whooping cough or pertussis is a highly epidemic bacterium that lives in the mouth, nose, throat and it's life-threatening especially in infants. It can easily spread via air by cough and sneeze of an infected individual [7]. Influenza is one of the most common contagious illnesses caused by influenza viruses. Influenza spreads simply through coughs and sneezes and can cause life threatening diseases or even death fro the population, i.e. the elderly persons, young children, or those with counteracted immune systems or chronic illnesses.

Conclusion

The proposed model presents the prediction of spread of coronvirus disease considering additional parameters like Infected population hospitalized and un-hospitalized also the population quarantined and self quarantine. The concern authority intervention will help to reduce i.e reducing the gap between affected and recovered population. Also it is observed that 10% of the recovered cases are again foud to be susceptible concern to the geographical location. As per the serious precaution measures and intervention of authority, its predicted that 40 % reduction in final epidemic size.

The network graph theory and the epidemical disease studies are strongly linked. Such theory models are very useful to simulate and understand the exact scenario with simple simulation model. The SEIR can be used as first tool to predict the initial and future situation. The proposed model implemented will also applicable to the other contagious diseases like Whooping cough, Influenza, chekungunia predictions.

References:

[1]E.C. PIRTLE and G.W. BERAN , "Virus survival in the environment" *Rev. sci. tech. Off. int. Epiz.,* 1991, 10 (3), 733-748

[2] Li Yan 1 & Et L," Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan" medRxiv preprint doi: https://doi.org/10.1101/2020.02.27.20028027.

[3] Cleo Anastassopoulou, Lucia Russo, Athanasios Tsakris, Constantinos Siettos, "Data-Based Analysis, Modelling and forecasting of theCOVID-19 outbreak", medRxiv preprint doi: https://doi.org/10.1101/2020.02.11.20022186.

[4] KC Santosh, "AI-Driven Tools for Coronavirus Outbreak: Need of Active Learning and Cross-Population Train/Test Models on Multitudinal/Multimodal Data" Journal of Medical Systems volume 44, Article number: 93 (2020)

[5] Dureau J, Kalogeropoulos K, Baguelin M. Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. 2013. Biostatistics **14**:541–555.

[6] Anderson R. M. and May R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. New York: Oxford University Press.

[7] Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science. 2003;
300(5627):1966-1970.

[8] Wallinga J, Lipsitch M. How generation intervals shape the relationshipbetween growth rates and reproductive numbers. Proc R Soc B. 2007;274:599-604.

[9] Bridgen JRE, Cummings DAT, Ho A, Jewell CP. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. https://doi.org/10.1101/2020.01.23.20018549 medRxiv. 2020. Preprint online January 24, 2020

[10] https://www.worldometers.info/coronavirus/country/india/

[11] https://www.idmod.org/docs/hiv/model-seir.html

[12] Ministry of Health and Family welfares, https://www.mohfw.gov.in/