

# PREDICTION OF DEATHS CAUSED BY COVID-19 USING MACHINE LEARNING

Kanaga Suba Raja.S<sup>1</sup>, Usha Kiruthika<sup>2</sup>, V.Balaji<sup>3</sup>, Anjana Sriram<sup>4</sup>

<sup>1,3,4</sup> Department of Information Technology, Easwari Engineering College, Chennai -600 089.

<sup>2</sup>Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur – 603 203

## ABSTRACT:

*Covid-19 prediction models are the most welcome and need of the hour in this current pandemic situation for paving way for decision making by Authorities. India due to its protective characteristics against COVID-19 is supposed to have the least death rate compared to nations having similar number of Covid cases. But the recent increase in the death toll has been a matter of deep concern. The SEIR models which were practiced earlier could not predict death rates accurately due to certain limitations in the procedure. Hence this paper is presented with a model which can efficiently predict the deaths due to covid-19 in every state in India. The aim of this paper is to present a model which predicts the number of fatalities caused by corona with maximum efficacy.*

**Keywords:** COVID-19, Prediction Models, SEIR model, Machine Learning Models

## 1. INTRODUCTION:

Corona Viruses is a group of Virus causing disease in Animals and Human. SARS virus strain which is also known as SARS – COV is an example and this spread rapidly in 2002 – 2003 resulting in 9.6% of the infected cases to fatality. MERS virus is one such virus belonging to the CORONA virus grouping. The outbreak of COVID -19 began in Wuhan in China and with first reporting in December 2019. Initially the seriousness was not realized and reports suggested the source to seafood and the Animal market in Wuhan as Corona Virus is common in animals like Camel & Cattle. But studies suggested that transmission from animals to humans is very rare. There were opinions that this new strain comes from bats though there are parallel views that Pangolins – the most trafficked mammal in the world may be the origin.

## Common Symptoms of COVID-19 :

Fever, difficulty in breathing, cough, sore throat, headache, muscular pain, loss of taste or smell etc. Under this eventuality One can not afford to neglect these symptoms though these may appear to be a day to day symptom in a common man under any circumstance. A person after exposing to the virus becomes symptomatic between 4 – 14 days and an average period of 14 days is required to assess the spread. It is opinioned that one in 6 people have a likely chance of experiencing severe symptoms like breathlessness which sometimes becomes fatal too.

How it spreads:

- Close proximity of less than 6 feet distance plays a major role in spreading the infectious disease.
- A symptomatic Covid – 19 carrier through their cough and saliva becomes high risk transmitters of the virus and hence their stage is highly contagious. Saliva or droplets can also land on nearby surface or objects making the chance of spreading the disease by mere touch. Studies suggests that an asymptomatic carrier who can be one among the daily acquaintances also carries 10% chance of transmitting the deadly virus.
- It is clinically advised that be it symptomatic or asymptomatic one should be quarantined for an estimated 14 days.

### Challenges & suggested preventive measures:

World Health Organization declared COVID-19 a pandemic in March 2020 and expressed that the risk factor could not be accurately assessed as the virus is spreading. Recently India's Corona Virus infected cases have crossed 2.5 million and almost all the states are becoming hot spots for the disease which a few months back was restricted to a few populated places. As population determines the spread of virus preventing the spread has become a global challenge. Though antiviral vaccination helps in preventing the spread of virus, controlling COVID-19 is becoming a threat and a challenge as developing a vaccine is still in stages and may take time to be opened for public distribution. Till such time the most awaited vaccine is developed as individuals the following stringent measures are to be followed :

1. Avoiding close contact with carriers of COVID – 19
2. Wearing Masks , washing hands regularly with soap for at least 20 seconds
3. There are apprehensions that wearing masks may reduce the oxygen levels or may pose difficulty in breathing . But researches have proved that it is not so.
4. A hand sanitizer can be used with at least 60% alcohol.
5. Avoid touching the face before washing the hands.

In the past SIR and SIER models were used to categorize infectious disease but with its own limitations. This paper is based on Machine learning models. Taking several factors into consideration the anticipated mortality rate is being predicted thus serving as a key factor for decision making by the Authorities.

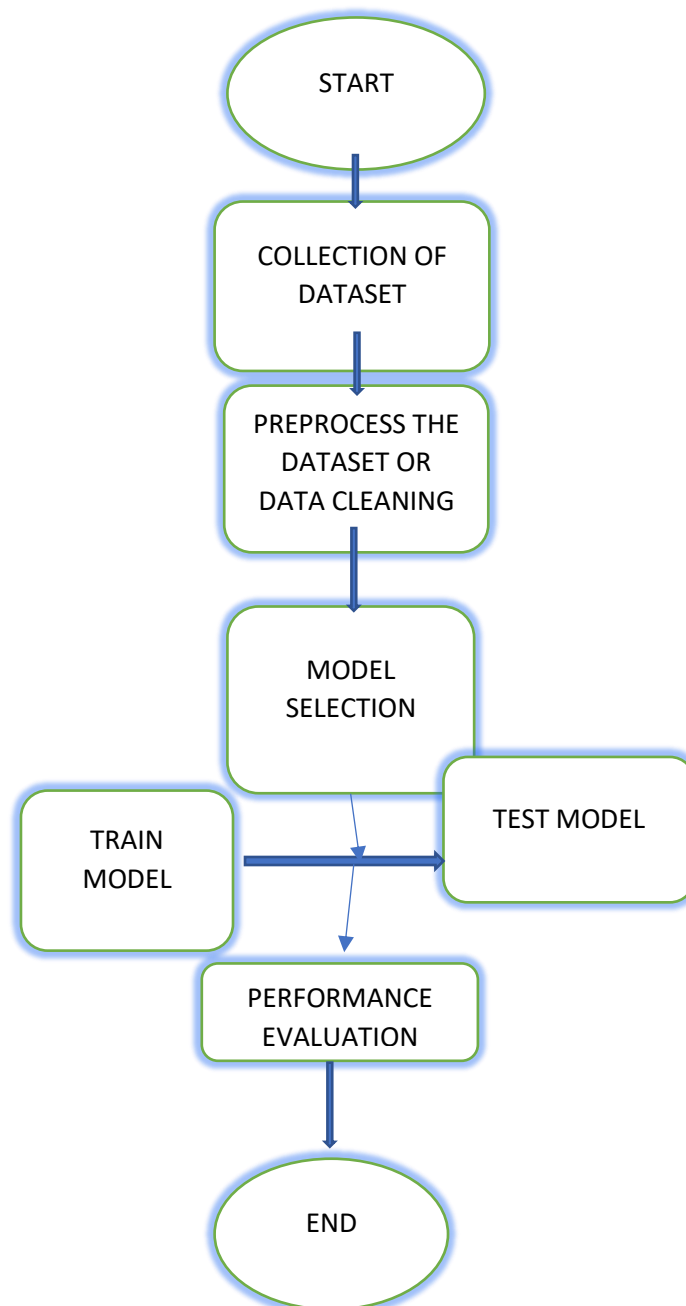
### **2. LITRATURE SURVEY:**

- A short term weather forecasting model was proposed by Dangi et al [1] .It was based on wavelet denoising and catboost algorithm which was used in the prediction of corona virus in India by using the correlations made between major cities in the world. This model included features like population density and temperatures recorded in each city to predict the covid-19 outbreak
- A model which predicts the high risk zones in the country was proposed by Sajadi et al. [2] This model used parameters such as weather ( weather modeling) and population density to predict the zones which are at high risk zones.
- Demonogeot et al. [3] has illustrated that the severity of coronavirus has a steep decrease in regions with hot and humid weather conditions. Hence the proposed model uses humidity and weather for prediction.
- A relationship between temperature and rate of speed of spread of the virus was proposed by Marvi and Arfeen .[4] These finding were very vital in developing other such relations.
- Xu et al conducted a study worldwide that uses datasets till April 2020.[5] The test was conducted at more than 3700 locations worldwide. This study uses statistical method . They believed that the result of their study would be vital in controlling the spread of the virus. They used features such as population density , delay in detection etc.
- Many studies including mathematical modelling [5] , analysis of epidemiology[6] and laboratory [7] point out that temperature and humidity play a very important role in the transmission of the decease.
- A model to check the effectiveness of public policies and social isolation in overcoming the trouble caused by the pandemic was proposed by Crokidakis[8]. This was done using the Susceptible-Infectious-Quarantined Recovered (SIQR) model.
- A study was conduction using 19 heterogeneous population to demonstrate and understand the spread of the virus. This study was conducted by Contreras et al[9]. Mohammed et al [10] studied different mathematical models and found that accurate results could be obtained by using fractional adams bashforth method (AB).
- A model proposed by Chakraborty and gosh[11] proposed approaches like forecasting and risk assessment for covid-19.

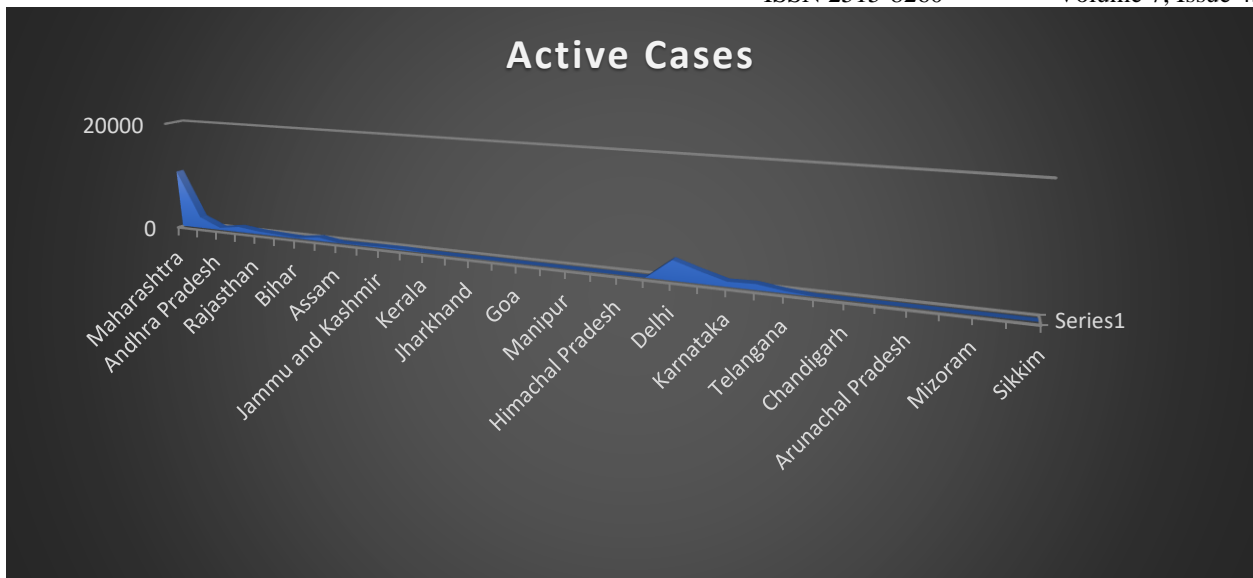
- A mathematical model was proposed by Mandel et al [12] which included governmental measures to prevent the transmission of the disease.
- A neural network model based on time series approach was proposed by Melin et al [13].

### 3. PROPOSED SYSTEM:

In this work we developed a Machine learning model to understand the correlation between confirmed and active cases in different states in the country to efficiently predict the number of deceased in each state in the country. A set of Linear models such as multiple linear regression model and polynomial linear regression is used. These models were selected to achieve high efficiency. These models were trained with features such as overall confirmed cases throughout the country, overall positive cases throughout the country, state wise confirmed cases and state wise positive cases.



Here we have analyzed the correlation between the different attributes and the death caused.



**4. MACHINE LEARNING MODELS USED:**

• **MULTIPLE LINEAR REGRESSION MODEL:**

Multiple linear regression is a machine learning regression model which can be used to predict the outcome of a variable by taking several attributes into consideration. In Multiple linear regression there is a linear relationship between the dependent and independent variables.

**FORMULA:**

$$y_i = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots \beta_ix_i + \epsilon$$

Where, yi= dependent variable

Xi=independent variable

B0,B1,B2...BI- features used.

efficiency achieved through this model is – 90.23%

**POLINOMIAL LINEAR REGRESSION MODEL:**

Polynomial linear regression is a machine learning regression model which is used to predict the outcome of a variable by taking several attributes into consideration. Though it is called linear , polynomial linear regression model forms a non linear relationship between attributes. For a dataset which has a high varied range, polynomial regression can yield more efficiency.

**FORMULA:**

$$y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \dots \beta_nx^n + \epsilon$$

where, y= dependent variable

x=independent variable

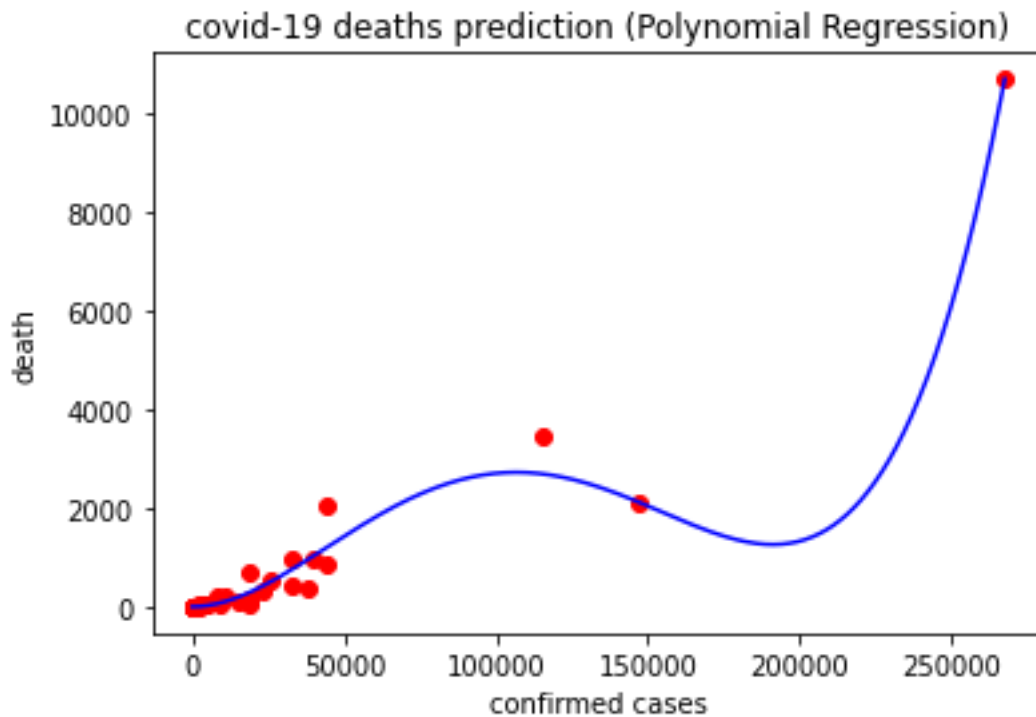
B0,B1 etc – features

Efficiency achieved through this model is -93%

**5. RESULT:**

Since polynomial linear regression showed better accuracy we take this model for prediction.

Accuracy achieved -93%



## 6. CONCLUSION:

Hence using polynomial regression we are able to accurately predict the rate of fatality or deaths caused by covid-19 by taking several factors into consideration. We were able to find correlations between attributes like confirmed cases and active cases and the number of fatal cases. We believe that this paper will serve officials in important decision making.

## 7. References:

- [1] R.R. Dangi, M. George Temperature, population and longitudinal analysis to predict potential spread for COVID-19
- [2] SSRN Electr J (2020), 10.2139/ssrn.3560786
- [3] Google Scholar
- [4] M. Sajadi, P. Habibzadeh, A. Vintzileos, S. Shokouhi, F. Miralles-Wilhelm, A. Amoroso Temperature and latitude analysis to predict potential spread and seasonality for COVID-19
- [5] SSRN Electr J (2020), 10.2139/ssrn.3550308
- [6] Google Scholar
- [7] J. Demongeot, Y. Flet-Berliac, H. Seligmann Temperature decreases spread parameters of the new covid-19 case dynamics
- [8] Biology, 9 (5) (2020), p. 94, 10.3390/biology905009
- [9] CrossRef Google Scholar
- [10] M. Marvi, A. Arfeen Demystifying a hidden trend: do temperature variations affect COVID-19 virus spread?
- [11] SSRN Electr J (2020), 10.2139/ssrn.3567084
- [12] Google Scholar
- [13] R. Xu, H. Rahmandad, M. Gupta, C. DiGennaro, N. Ghaffarzadegan, H. Amini, et al. Weather conditions and COVID-19 transmission: estimates and projections
- [14] SSRN Electr J (2020), 10.2139/ssrn.3593879
- [15] Google Scholar
- [16] A.C. Lowen, S. Mubareka, J. Steel, P. Palese Influenza virus transmission is dependent on relative humidity and temperature

- [17] PLoS Pathog, 3 (10) (2007), p. e151, 10.1371/journal.ppat.0030151
- [18] CrossRefView Record in Scopus
- [19] Google Scholar
- [20] A.I. Barreca, J.P. Shimshack Absolute humidity, temperature, and influenza mortality: 30 years of county-level evidence from the united states
- [21] Am J Epidemiol, 176 (suppl\_7) (2012), pp. S114-S122, 10.1093/aje/kws259
- [22] CrossRefView Record in Scopus
- [23] Google Scholar
- [24] T. Żuk, F. Rakowski, J.P. Radomski Probabilistic model of influenza virus transmissibility at various temperature and humidity conditions
- [25] Comput Biol Chem, 33 (4) (2009), pp. 339343, 10.1016/j.compbiolchem.2009.07.005
- [26] ArticleDownload PDFView Record in Scopus
- [27] Google Scholar
- [28] Crokidakis N.. COVID-19 spreading in rio de janeiro, brazil: do the policies of social isolation really work?2020;. 10.1101/2020.04.27.20081737
- [29] Google Scholar
- [30] S. Contreras, H.A. Villavicencio, D. Medina-Ortiz, J.P. Biron-Lattes, Á. Olivera-Nappa A multi-group SEIRA model for the spread of COVID-19 among heterogeneous populations
- [31] Chaos, Solitons & Fractals, 136 (2020), p. 109925, 10.1016/j.chaos.2020.109925
- [32] ArticleDownload PDF
- [33] Google Scholar
  
- [34] M.S. Abdo, K. Shah, H.A. Wahash, S.K. Panchal On a comprehensive model of the novel coronavirus (COVID-19) under mittag-leffler derivative
- [35] Chaos, Solitons & Fractals, 135 (2020), p. 109867, 10.1016/j.chaos.2020.109867
- [36] ArticleDownload PDF
- [37] Google Scholar
- [38] Chakraborty T., Ghosh I. Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis2020;. 10.1101/2020.04.09.20059311
- [39] Google Scholar
- [40] M. Mandal, S. Jana, S.K. Nandi, A. Khatua, S. Adak, T. Kar A model based study on the dynamics of COVID-19: prediction and control
- [41] Chaos, Solitons & Fractals, 136 (2020), p. 109889, 10.1016/j.chaos.2020.109889
- [42] ArticleDownload PDF
- [43] Google Scholar
- [44] P. Melin, J.C. Monica, D. Sanchez, O. Castillo Multiple ensemble neural network models with fuzzy response aggregation for predicting COVID-19 time series: the case of mexico
- [45] Healthcare, 8 (2) (2020), p. 181, 10.3390/healthcare8020181
- [46] CrossRefView Record in Scopus
- [47] Google Scholar