Abstract

<u>Prediction of Covid – 19 using Regression Model</u>

The outbreak of COVID-19 Coronavirus, namely SARS-CoV-2, has created a calamitous situation throughout the world. The cumulative incidence of COVID-19 is rapidly increasing day by day. Machine Learning (ML) can be deployed very effectively to track the disease, predict growth of the epidemic and design strategies and policies to manage its spread. This project applies an improved mathematical model to analyse and predict the growth of the epidemic.

An ML-based improved model has been applied to predict the potential threat of COVID-19 in countries worldwide.

The novel Coronavirus disease (COVID-19) was first reported on 31 December 2019 in Wuhan, Hubei Province, China. It started spreading rapidly across the world. The cumulative incidence of the causative virus (SARS-CoV-2) is rapidly increasing and has affected 196 countries and territories with USA, Spain, Italy, U.K. and France being the most affected. The World Health Organization (WHO) has declared the coronavirus outbreak a pandemic, while the virus continues to spread. As on 11 December 2020, a total of 30,031 confirmed new positive cases have been reported and 443 death counts. The major difference between the pandemic caused by CoV-2 and related viruses, like

Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS), is the ability of CoV-2 to spread rapidly through human contact and leave nearly 20% infected subjects as symptomless carriers.

Moreover, various studies reported that the disease caused by CoV-2 is more dangerous for people with weak immune systems. The elderly people and patients with life threatening diseases like cancer, diabetes, neurological conditions, coronary heart disease and HIV/AIDS are more vulnerable to severe effects of COVID-19. In the absence of any curative drug, the only solution is to slow down the spread by exercising "social distancing" to block the chain of spread of the virus. This behavior of CoV-2 requires developing a robust mathematical basis for tracking its spread and automation of the tracking tools for online dynamic decision making.

There is a need for innovative solutions to develop, manage and analyse big data on the growing network of infected subjects, patient details, their community movements, and integrate with clinical trials and pharmaceutical, genomic and public health data. Multiple sources of data including, text messages, online communications, social media and web articles can be very helpful in analyzing the growth of infection with community behaviour

In this project, prediction model is deployed to predict the possible cases of covid that can be expected to occur using a simple linear regression algorithm.

Regresion models help us describe a relationship between variables by fitting a line to the observed data. Linear regression models use a straight line, while logistic and nonlinear regression models use a curved line. Regression allows you to estimate how a dependent variable changes as the independent variable change.

Simple linear regression is used to estimate the relationship between two quantitative variables. Simple linear regression can be used in situations like -

1. How strong the relationship is between two variables (e.g. the relationship between rainfall and soil erosion).
2. The value of the dependent variable at a certain value of the independent variable (e.g. the amount of soil erosion at a certain level of rainfall).

Here, simple regression algorithm predicts the possibility of covid cases that can be confirmed using already available data from previous days using the existing Date and Confirmed columns. By doing so, suitable precautions could be taken from the government end to check for the stock of hospitality, medical equipments, doctors and much more in advance.

Rashika Suresh

+919483374137

rashikasuresh5758@gmail.com