A Data Analysis of Global Pandemic Coronavirus using Machine Learning Algorithms

Sakshi Jain

B.E, Department of Computer Science, P.D.M University, Haryana, India

Abstract: The world is experiencing so many disturbing thing at social, political, natural and health level. The outbreak of COVID-19 Coronavirus, namely SARS-CoV-2, has become a global issue and declared as pandemic by world health organization. Many countries have implemented restrictions on population movement to slow the spread of severe acute respiratory syndrome coronavirus 2 and prevent health systems from becoming overwhelmed; some countries have instituted full or partial lockdowns. There are several outbreak prediction models for COVID-19 are being used around the world to make informed decisions and enforce relevant control measures. We try to concentrate on COVID-19 predictions using Machine Learning Techniques. The covid-19 virus second mutate has been found which increase infection rate more rapidly and increase the number of patient. The Data such as the number of persons are infected by the virus in the world, total number of persons infected by this virus in India, number of persons infected by this virus per day, the number of persons deaths, number of persons per day, number of persons recover due to this virus is used as an input and these features are modeled for analysis and prediction.

Keywords: Forex Reserve, Foreign currency, Machine Learning, Arima, Predictive Analysis, Fbprophet, Forecasting Model, Central bank (RBI)

1. Introduction

Access to accurate outbreak prediction models is essential to obtain insights into the likely spread and consequences of infectious diseases.

After the first case of coronavirus (Covid-19) come out from Wuhan, China, Sudan, Africa suspected two caes of coronavirus infection and started coming from everywhere from the world.

World Health Organization has been declared COVID-19 as a Public Health Emergency of International Concern (PHEIC) on 30 January 2020.

People face various health issues but recently we found that the coronavirus has start mutating itself and become more complected to inspect through RT-PCR test.

The first case of coronavirus reported:

According to unpublished reports from government of china, the first case of coronavirus traced back to 17 December 2019, a 55 year old Hubei province citizen. There were four men and five women reported to be infected in November, but none of them were "patient zero".

2. Why it is called COVID 19?

Coronaviruses are a family of viruses that can cause illnesses such as the common cold, severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS).

‘CO’ stands for corona, ‘VI’ for virus, and ‘D’ for disease. Viruses are named based on their genetic structure to facilitate the development of diagnostic tests, vaccines and medicines.

Coronaviruses, named for the crown-like spikes on their surfaces, are a large family of viruses that are common in people and many different species of animals, including camels, cattle, cats, and bats.

Animal coronaviruses rarely infect people and then spread between people. This occurred with two earlier coronaviruses, MERS-CoV and SARS-CoV.

SARS-CoV-2 virus is a beta coronavirus, like MERS-CoV and SARS-CoV. These viruses have their origins in bats.

The virus itself is named SARS-CoV-2. Each part of the name is an abbreviation: SARS stands for “severe acute respiratory syndrome”

3. Symptoms

Signs and symptoms of coronavirus disease 2019 (COVID-19) may appear two to 14 days after exposure.

Common signs and symptoms can include:

i. Fever
ii. Cough
iii. Tiredness

Other symptoms can include:

Shortness of breath or difficulty breathing
i. Muscle aches
ii. Chills
iii. iv. Sore throat
iv. v. Runny nose
v. Headache
vi. Chest pain
vii. Pink eye (conjunctivitis)

This list is not all inclusive. Children have similar symptoms to adults and generally have mild illness.
4. Literature Review


Author: Sina F. Ardabili, Amir Mosavi, PedramGhamisi, Filip Ferdinand, Annamaria R. Varkonyi-Kocz, Uwe Reuter, TimonRabczuk, Peter M. Atkinson

Objective: analysis of machine learning and soft computing models to predict the COVID-19 outbreak

Conclusion: Although the most difficult prediction is to estimate the maximum number of infected patients, estimation of the individual mortality rate \((n(\text{deaths})/n(\text{infected}))\) is also essential. The mortality rate is particularly important to accurately estimate the number of patients and the required beds in intensive care unit.

2) Title Paper: Machine learning-based prediction of COVID-19 diagnosis based on symptoms

Author: Yazeed Zoabi, Shira Deri-Rozov and Noam Shomron

Objective: Prediction models that combine several features to estimate the risk of infection have been developed.

These aim to assist medical staff worldwide in triaging patients, especially in the context of limited healthcare resources.

Conclusion: Close contact with an individual confirmed to have COVID-19 was also an important feature, thus corroborating the disease’s high transmissibility15 and highlighting the importance of social distancing.

3) Title Paper: Intelligent system for COVID-19 prognosis: a state-of-the-art survey

Author: Janmenjoy Nayak & Bighnaraj Naik & Paidi Dinesh & Kanithi Vakula& B. Kameswara Rao

Objective: Analyzing the impact of data types and the nature of data along with challenges in processing the data for and to focus on some future challenges in COVID-19

Conclusion: The COVID-19 had a major impact on the global economy. These are extremely unsure times for the economic markets, with nations around the world anguish the threatening effects of the epidemic.

4) Title Paper: Predictive Analytics of COVID-19 using Information,Communication and Technologies

Author: Parikshit N. Mahalle, etal.

Objective: Analysis of various predictive analytics methods

Conclusion: Simulation result of this model shows that the confirmed COVID-19 infected cases would be 1.6 million and 2.3 million by end of May and June respectively.

5) Title Paper: Deep Learning applications for COVID-19

Author: Connor Shorten, Taghi M. Khoshgoftaar and BorkoFurht

Objective: Precision Diagnostics, Protein Structure Prediction, and Drug Repurposing and utilized in Spread Forecasting for Epidemiology.

Conclusion: Explored across data domains in Natural Language Processing, Computer Vision, Life Sciences, and Epidemiology.

5. Research Methodology

ARIMA: An autoregressive integrated moving average, or ARIMA, is very popular and widely used a statistical analysis model that uses time series data to either better understand the data set or to predict future trends.

Briefly, Autoregression (AR) refers to a model that shows a changing variable that regresses on its own lagged, or prior, values.

Integrated (I) represents the differencing of raw observations to allow for the time series to become stationary, i.e., data values are replaced by the difference between the data values and the previous values.

Moving average (MA) incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

A linear regression model is constructed including the specified number and type of terms, and the data is prepared by a degree of differencing in order to make it stationary, i.e. to remove trend and seasonal structures that negatively affect the regression model.

A value of 0 can be used for a parameter, which indicates to not use that element of the model. This way, the ARIMA model can be configured to perform the function of an ARMA model, and even a simple AR, I, or MA model.

Adopting an ARIMA model for a time series assumes that the underlying process that generated the observations is an ARIMA process. This may seem obvious, but helps to motivate the need to confirm the assumptions of the model in the raw observations and in the residual errors of forecasts from the model.

Prophet Model

Prophet is one of the newest buzzwords in the field of time series forecasting.

Prophet has many parameters that could be set, like growth, changepoints, holidays, interval_width and weekly_seasonality. Growth is used to specify a trend, linear or logistic. Other parameters are very intuitive.

Prophet has its implementation in Python and R.

How Prophet works:

At its core, the Prophet procedure is an additive regression
model with four main components:
A piecewise linear or logistic growth curve trend. Prophet automatically detects changes in trends by selecting change points from the data.

6. Limitations

The major limitations of the study are

- The study is made only in consideration with India and not applicable to any part of the globe.
- The study fully depends on the secondary data, which has its own limitations.

7. Data Analysis and Interpretation

Covid cases starting data

The Seven days average and previous six days data

World Covid start cases
Growth of confirmed, death and recovered

Covid second wave in states indicate the recovery of patients increased.

8. Observation

The number of cases is increase rapidly in the second wave of coronavirus and its become more dangerous due to its mutation. A big rise of confirmed cases from the month of april and it reach to its peak in mid of may. The mortality rate was its peak in between last april to min may after that the number of active patients will start decreasing which indicate the recovery of patients increased.

Top 5 Countries, highly affected by corona virus.
1) United State
2) India
3) Brazil
4) France
5) Turkey

<table>
<thead>
<tr>
<th>ds</th>
<th>yhat</th>
<th>yhat_lower</th>
<th>yhat_upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>190</td>
<td>2020-07-30</td>
<td>1.674593e+07</td>
<td>1.658745e+07</td>
</tr>
<tr>
<td>191</td>
<td>2020-07-31</td>
<td>1.695120e+07</td>
<td>1.679334e+07</td>
</tr>
<tr>
<td>192</td>
<td>2020-08-01</td>
<td>1.716899e+07</td>
<td>1.698455e+07</td>
</tr>
<tr>
<td>193</td>
<td>2020-08-02</td>
<td>1.736665e+07</td>
<td>1.718182e+07</td>
</tr>
<tr>
<td>194</td>
<td>2020-08-03</td>
<td>1.756134e+07</td>
<td>1.735239e+07</td>
</tr>
</tbody>
</table>

Number of cases with upper and lower limits
The prediction model the graph shows limit of increase.

The forecast of growth of cases

Volume 10 Issue 5, May 2021

www.ijsr.net

Licensed Under Creative Commons Attribution CC BY
9. Conclusion

At this time the Covid situation is very serious. Its start mutating and become more dangerous then before. It is important to remember that to recover from this pandemic situation we need to follow the guidelines very strictly. And vaccinate our self.

References


[34] Gao W, Mahajan S, Sulam J, Gray J. Deep learning in protein structural modeling and design