Covid-19 Pandemic outbreak prediction and Designing a constraint satisfaction programming model to report PPE allocation

Abdul Irfan Mohammad[1], $ManojKumarBhuma$[2], $DR.ZubairFadlullah$[3]
Department of Computer Science[1,2]
Lakehead University[1,2]
amoham29@lakeheadu.ca[1], $mbhuma@lakeheadu.ca$[2], $Zubair.Fadlullah@lakeheadu.ca$[3]

Abstract- Officials around the world are using different disease modelling models for COVID-19 to make informed decisions and implement appropriate control measures. Simple epidemiological and mathematical models have been given further prominence by authorities among the standard models for COVID-19 global pandemic prediction, and are common in the media. Standard models have demonstrated poor performance for long-term estimation due to a high variation and a lack of critical data. While the research includes many attempts to address this problem, there is a need to strengthen the basic generalisation and robustness capabilities of current models. Optimization methods and Machine Learning are key for providing a comparative analysis to estimate the pandemic and distribution of personal protective equipment related to COVID-19. The software prophet, is one that performs forecasting on time series data and it is developed by Facebook.This software is known to give promising results. Each of the three forms of forecasts is based on Models, such as the number of newly diagnosed cases, deaths and regeneration figures in the current 60 days. The extremely dynamic complexity of COVID-19 outbreak and its variability from nation-to-nation, this research indicates machine learning as an appropriate modelling method for the outbreak. This paper further gives the solution to distribution of PPE(Personal Protective Equipment) to staff in various hospitals according to the severity level.

I. Introduction

Machine learning (ML) has proven to be an influential field Research over the last decade by the implementation of certain quite difficult But tricky real-world problems. The frameworks Of almost all real-world areas like healthcare, business applications, natural language processing (NLP), entertainment. One of the most important areas of ML is forecasting, interdisciplinary approaches ML algorithms were used to direct the likely course of action needed in many fields of application including weather forecasting, disease forecasting, stock market forecasting. The research focuses in particular on live monitoring of confirmed cases COVID-19 and the research also focuses on COVID-19 outbreak predictions and Rapid Reaction. This forecast mechanism can be very helpful in decision-making to tackle the current situation in order to direct early steps for more successful treatment of these diseases.

This research aims to provide an early predictive model for the transmission of novel coronavirus, also known as SARS-CoV-2, officially identified by the World Health Organization (WHO) as COVID-19. Actively COVID-19 is a very significant danger to human life worldwide. The virus was first detected in a city of China named Wuhan at the end of 2019 when symptoms emerged in vast numbers of people. Corona virus is majorly found to be predominantly initiated from animals. They can cause diseases in humans ranging from common cold to more serious diseases such as Extreme Acute Respiratory Syndrome (SARS) and Respiratory Syndrome of the Middle East (MERS). The disease arising from the new coronavirus was called COVID-19. The

world declared this as a pandemic response to the COVID-19 outbreak. Countries tend to be in intensive containment mode through a number of public health initiatives to avoid the transmission of the infection, including physical isolation, travel bans and business and school suspensions.

Contributing to the ongoing humanitarian disaster our effort in this analysis is to build a COVID-19 forecasting framework. Prediction is made for the three main disease factors for the next 10 days: 1) the number 0f New reported cases. 2) the number of deaths 3) the number of Recovery.

Implementing the optimal solution for PPE (Personal Protective Equipment) distribution due to the COVID-19 pandemic to staffs in various hospitals according to the severity level (1-5) of the patients and number of patients and staff present in a hospital., the world needs a solution to distribute PPE .

This research has some main findings described below:

- Analysing COVID-19 cases in India
- Number of Active Cases affected state wise
- Confirmed cases Vs Recovered figures
- How the cases are rising (plotting exponential because we have to flatten the curve)
- Identifying whether the dynamics is same as mostly affected countries like Wuhan/Italy.
- Insight on the world wide data
- Predicting total number of cases world-wide with Prophet.
- Implementing optimal solution for PPE (Personal Protective Equipment) distribution

II. Back Ground and Motivation

Many case studies are being implemented to tackle the global pandemic COVID-19, mostly the machine learning models shows advancement. Many machine learning models are being in development using approaches like Supervised Machine Learning Models to make predictions. And in this learning methodology the learning algorithm takes a data set of input instances along with their corresponding instances regressor to train the regression model. The results are not favourable. But Facebook developed a software called 'Prophet' which is an open source software which is available to download on CRAN and PyPI. This shows promising results for the prediction and using google OR tool the optimal solution is given using constraint satisfaction programming model for the distribution of the PPE to curb the pandemic.

III. METHODOLOGY

The prediction is done based on the analysis of COVID-19 cases, Number of Active Cases affected state wise, Confirmed cases Vs Recovered figures, How the cases are rising (plotting exponential because we have to flatten the curve), Identifying whether the dynamics is same as mostly affected countries like Wuhan/Italy, Insight on the world wide data and finally predicting total number of cases world-wide with Prophet. And the solution by implementing optimal solution for PPE (Personal Protective Equipment) distribution.

A. Data Set

The learning models have been trained using the COVID-19 patient stats released by official government website, time series data-set scraped provided by Johns Hopkins. https://www.mohfw.gov.in/
https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset

B. Prophet

Prophet is a technique of time series data prediction focused on an additive model wherein non-linear patterns are composed on a yearly, weekly and seasonal basis, plus outcomes for the holidays. This fits better for time series which have clear seasonal effects and historical evidence from many seasons. Prophet is resilient towards lost pattern data and transitions, and generally treats outliers well. Prophet is accurate and fast, fully automatic and has tunable forecasts.

When a forecasting model is not working as expected, we want to be able to change the method 's parameters with regard to the specific problem at hand. Tuning these approaches requires a detailed understanding of how the sequence model functions. Prophet Forecasting offers intuitive, efficient-to-tune parameters. And someone who lacks experience in forecasting models may use this

to make accurate forecasts in a business scenario for a variety of problems. This uses a compostable model of the time series with three key components of the model. They are mixed into the equation below:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

Fig. 1.

Using time as a basis for regression, Prophet attempts to suit many linear and nonlinear time functions as components. The same method followed by exponential smoothing in the Holt-Winters technique is the modeling of seasonality as an additive variable.

C. Procedure

The input to Prophet, with 95percent prediction interval is created as a base model which takes a dateframe with two columns which is a datestamp and value. The value must always be quantitative and represents the metric we would like to estimate, as presented through the case here. Next we will generate a forecast for the 60 days ahead of confirmed cases of new cases using Prophet. we use a method to predict in which we assign each row in future as a predicted value named as yhat. The forecast entity used is a new data frame that contains a column yhat with the prediction and columns for intervals of uncertainty. The tolerance lies between yhat lower and yhat upper.

D. designing a constraint satisfaction programming model

constraint satisfaction problem (CSP) is a problem that requires its solution within some limitations or conditions also known as constraints. It consists of the following:

- A finite set of variables which stores the solution (V = V1, V2, V3,....., Vn)
- A set of discrete values known as domain from which the solution is picked (D = D1, D2, D3,.....,Dn)
- A finite set of constraints (C = C1, C2, C3,....., Cn)

The approach follows by

- Creating a variable set
- Creating a domain set
- Creating a constraint set with variables and domains (if possible) after considering the constraints
- Finding an optimal solution to distribute the PPE to the staff in the hospitals

E. RESULTS AND DISCUSSIONS

Using the data visualization methods the forecast is shown and predicted results are shown which are found to be accurate by the data published using the Prophet open source software.
predicting the future with date, and upper and lower limit of y value in the below image.

	ds	yhat	yhat_lower	yhat_upper
216	2020-08-25	1.798858e+07	1.643078e+07	1.939385e+07
217	2020-08-26	1.811946e+07	1.655003e+07	1.956646e+07
218	2020-08-27	1.826311e+07	1.663526e+07	1.971450e+07
219	2020-08-28	1.840845e+07	1.675735e+07	1.994988e+07
220	2020-08-29	1.854895e+07	1.684840e+07	2.012529e+07

Fig. 2.

The plot of world wide covid cases is in the image below

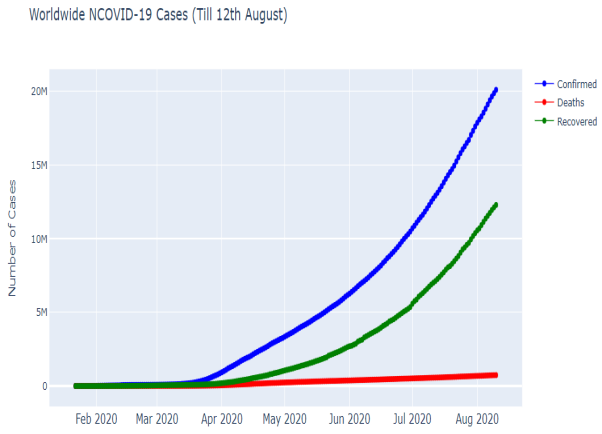

Fig. 3.

Below image is the forecast data frame using prophet plot

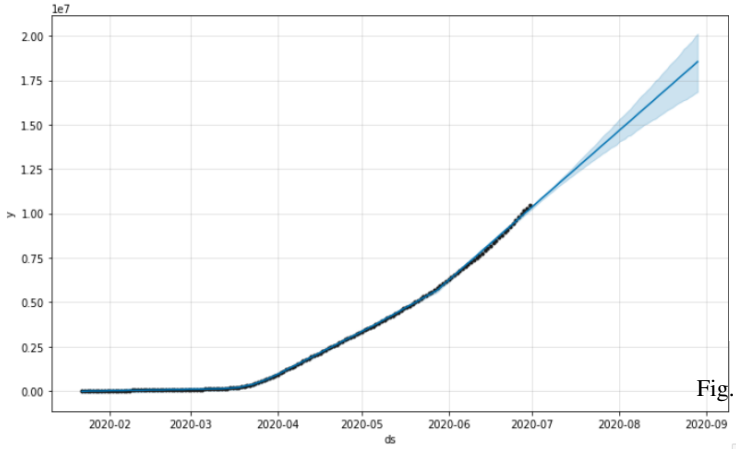

Fig. 4.

The plot of confirmed cases forecast is in the image below

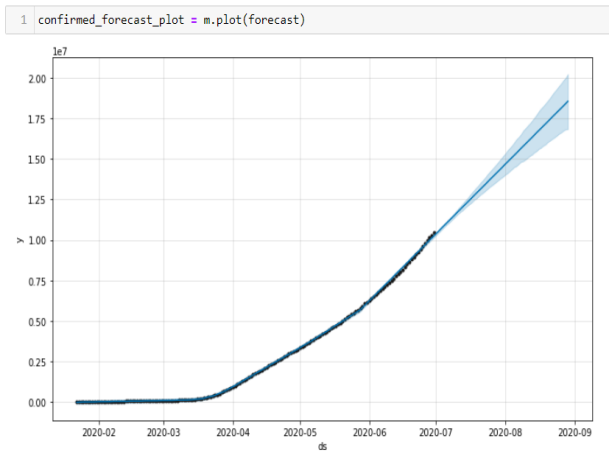

Fig. 5.

The plot of deaths cases forecast is in the image below

The plot of recovered cases forecast is in the image below

The plot of final analysis of actual, forecast, confidence is in the image below

The most important question which we have set to answer, inference from the analysis and forecast done India can become the leading country and home to COVID-19 if the same trends are seen when necessary steps are not taken as seen in the graph in upcomming 7 days.

- Very Uncertain Results about Death
- Most Recoveries are reported near Weekends

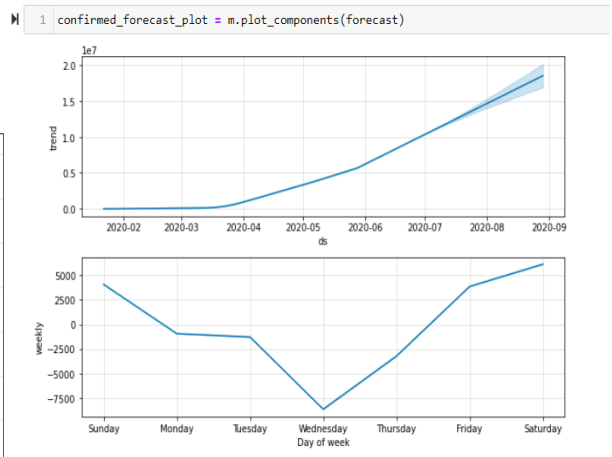

Fig. 6.

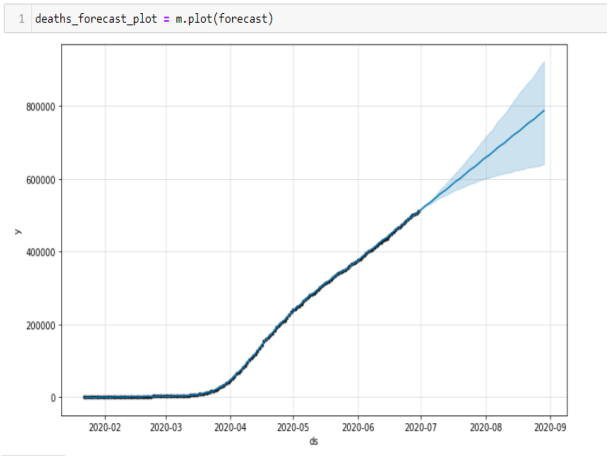

Fig. 7.

- Confirmed Results Exceed the Forecasted Results, Meaning that situations are worse than those predicted with 95 Percent confidence interval expectations.
- Confirmed Deaths are significantly lower than Forecasted Deaths(Exponential in Nature), means that infection is less lethal. The below image shows the optimal solution to distribute the PPE to the staff in hospitals in the emergency situation of the pandemic considering the severity level as well of the patients. The following parameters are considered for framing the problem and getting an optimal solution. Conditions complexity that are included to give solution keeping in mind about the real world scenarios possible
  - All staff should be given at least one PPE.

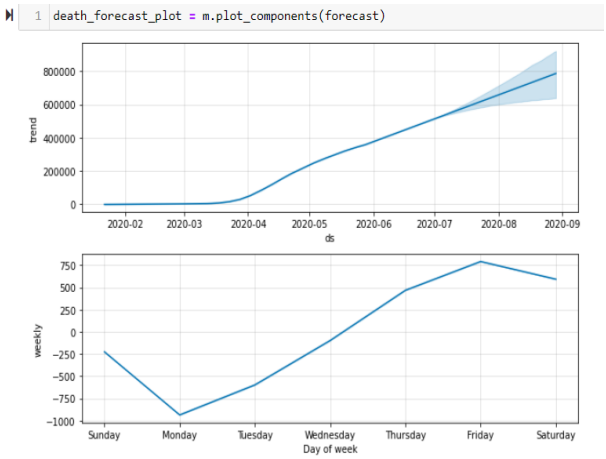

Fig. 8.

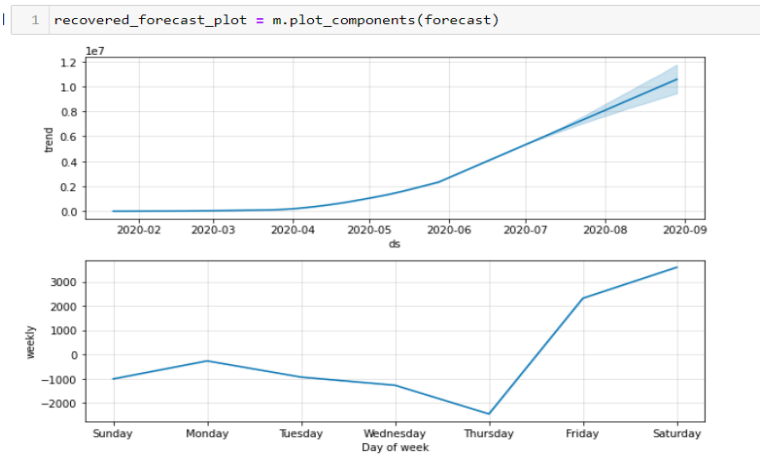

Fig. 10.

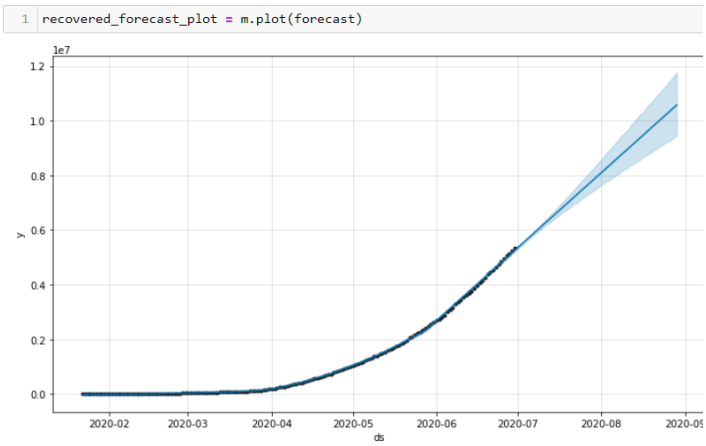

Fig. 9.

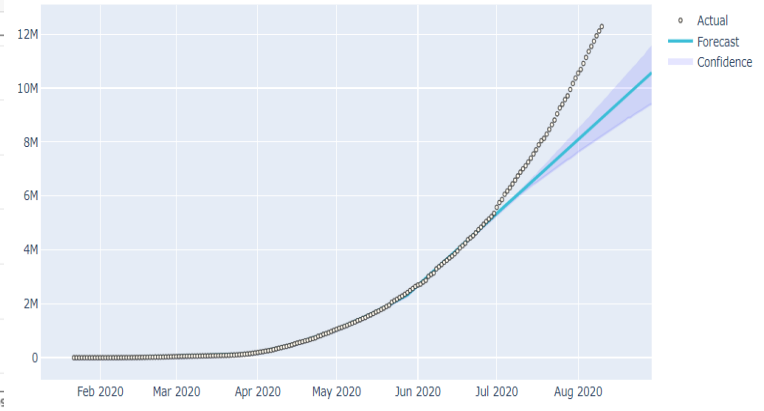

Fig. 11.

- No two staff can share a single PPE.
- No PPE should remain unallocated.
- The hospital dealing with the most severe patients will have more PPEs allocated to there.
- The hospital with the higher number of staffs will have more PPEs allocated to there.

The optimal solution is published below keeping the hard and soft constraints.

IV. CONCLUSION

COVID-19 pandemic's hazard will cause a major global crisis. Many researchers and government agencies around the world are worried that the pandemic can impact a large proportion of the world 's population. An ML-based prediction framework for predicting the likelihood of the COVID-19 outbreak globally has been proposed in this report. The program analyses the data set containing the actual day-to-day past results, and uses machine learning algorithms to forecast the prediction of the near future. Generally, we conclude that model forecasts are accurate according to the present scenario, which could be useful in knowing the situation that is yet to come. Therefore, the study forecasts can also be of great help to the authorities in taking appropriate action and in making decisions to tackle the COVID-19 crisis. Within the future prospects, this analysis will be constantly improved, next we expect to explore the prediction approach using the revised data set, using the most reliable and effective ML methods for forecasting. Live forecasting in real-time will be one of the key focus of our future research. The constraint based programming model is developed which gives an optimal solution for the main problem to distribute PPE (Personal Protective

Parameter	Values
Total PPE	700
Hospital	5
Patients	Hospital #1= 30 Hospital #2= 20 Hospital #3= 35 Hospital #4= 34 Hospital #5= 15
Staff	Hospital #1= 50 Hospital #2= 55 Hospital #3= 65 Hospital #4= 72 Hospital #5= 45
Severity Level	Initialize severity level for each patients randomly in the range of 1-5

Fig. 12.

```
hospitals = range(5)
patients = [30,20,35,34,15]
staff = [50,55,65,72,45]
severity = [sum(randint(1,5) for i in range(patients[j])) for j in hospitals]
ppeCount = 700
```

```
solve3(hospitals, patients, staff, severity, ppeCount)
```

```
1 : 50
2 : 55
3 : 65
4 : 485
5 : 45
```

Fig. 13.

Equipment) to staffs in various hospitals according to the severity level (1-5) of the patients and number of patients and staff present in a hospital.

V. Contribution and Novelty

- Project Motivation (Abdul Irfan) : I had hands-on on optimization problems, for the course Optimization Methodology which had problems mostly related to health care. From the exposure, developed a constraint programming model for PPE distribution for the staff of Hospital based on patient severity level and first analysed the situation of COVID-19 cases in India as it is leading now.
- Literature Review (Manoj Kumar) : Referred various approaches in IEEE papers regarding ML techniques and compared results.
- Programming (Abdul Irfan and Manoj Kumar) : Code written using various libraries to visualize the data set (time series) taken from John Hopkins and data from government website.
- Constraint Programming Model (Abdul Irfan) : Based on the course Optimization Methodology undertaken, using soft and hard constraints by using Google OR Tool.
- Novelty : The constraint programming model is designed on own and the outbreak prediction is made by referring various sources like ”COVID-19 Future Forecasting Using Supervised Machine Learning Models” and Edureka, online platform learning.
- Research Paper (Abdul Irfan and Manoj Kumar) : Overall contributed to all the sections discussing and giving insights.

References

[1] FURQAN RUSTAM , AIJAZ AHMAD RESHI , (Member, IEEE), ARIF MEHMOOD , SALEEM ULLAH , BYUNG-WON ON , WAQAR ASLAM , (Member, IEEE), AND GYU SANG CHOI ,”COVID-19 Future Forecasting Using Supervised Machine Learning Models”, VOLUME 8, 2020.

[2] https://facebook.github.io/prophet/

[3] https://www.edureka.co/post-graduate/machine-learning-and-ai

[4] G. Bontempi, S. B. Taieb, and Y.-A. Le Borgne, “Machine learning strategies for time series forecasting,” in Proc. Eur. Bus. Intell. Summer School. Berlin, Germany: Springer, 2012, pp. 62–77

[5] C. P. E. R. E. Novel, “The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (Covid-19) in China,” Zhonghua Liu Xing Bing Xue Za Zhi= Zhonghua Liuxingbingxue Zazhi, vol. 41, no. 2, p. 145, 2020.

[6] Johns Hopkins University Data Repository. Cssegisanddata. Accessed: Mar. 27, 2020. [Online]. Available: https://github.com/CSSEGISandData

[7] https://study.com/academy/lesson/constraint-satisfaction-problems-definition-examples.html

[8] https://www.mohfw.gov.in/

[9] https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset