International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.36, January 2021 – www.ijais.org

Predictive Modelling of Covid-19 Confirmed, Death related and Recovered Cases in Italy, Wuhan, South Korea and India

Ajinaja Micheal Olalekan
Department of Computer Science
Federal Polytechnic Ile-Oluji
Ondo State

Abiona Akeem Adekunle
Department of Computer Science
Federal Polytechnic Ile-Oluji
Ondo State

Ibiyomi Michael Adeniyi
Department of Computer Science
Federal Polytechnic Ile-Oluji
Ondo State

ABSTRACT
The coronavirus pandemic has been ravaging the world since December 2019.The pandemic started from Wuhan in China late 2019 and has greatly affected thousands of people with the impact felt in different economy. This required the need to observe, monitor and predict infected, recovered and death related COVID-19 cases for suitable control. Prophet was used to predict future values by creating a base model with no tweaking of seasonality-related parameters and additional regressors. The data was extracted from the John Hopkins state of art data centre website which spanned from 31st January 2020 to 25th Match 2020. We adopted the machine learning library (prophet) to predict a week data and also showed analysis using Plotly. Italy, Wuhan, South Korea and India were compared and analysed using the model. The results showed that confirmed cases in Italy continually kept rising exponentially with a certain fixed pattern, confirmed cases in India continually kept rising exponentially with no fixed pattern with the graph attaining sigmoid early but later increased exponentially with no particular fixed pattern which was as a result of very less test done in India, confirmed cases in South Korea was rising gradually with sigmoid attained at some level with the government able to reduce the spike in infected persons and in Wuhan, there have been almost a negligible number confirmed cases in a week. This prediction showed how government can prepare for any eventualities in the future and be able to combat future pandemic occurrence, Emphasis should be on testing and enforcing early lockdown

General Terms
Machine Learning, Predictive Analysis

Keywords
COVID-19, Prophet, recovered, deaths, confirmed, cases

1. INTRODUCTION
Numerous individuals who are informed in the analysis and information communication and technology (ICT) driven communities are now realizing the how data is hugely important in the industry. Many organizations has chosen to use to data despite them having no relation with storing and processing huge chunk of data or data involving enormous amount of computation. A lot of literatures have been written about data and its benefits in the industry for advance analysis in both the private and public sector. The ease of obtainability of fresh or new data sources coupled with the increase of complex analytical opportunities have warranted the need to reconsider the architecture of existing data to allow data analyst take advantage of big data.

Nowadays, data has caught the eye from industry, academia and government. Dataset whose size is beyond the processing capability of traditional databases or computers are referred to as big data. Data is generated constantly at an ever-increasing rate using different devices and platforms such as social media, mobile phones, computers and other imaging technologies. All these data generated are stored in a particular location and used for different purposes such as medical diagnosis, prediction of weather or spread of diseases among others. Sensors and devices routinely produce information that are required to be stored and processed in real time. Having those huge amount of data is extremely difficult to store in terms of cost and size but what is more challenging is analysing those huge amount of data especially if the data does not fit into the traditional concepts of data structure to ascertain meaningful pattern, extract useful information and use same data to predict certain outcome.

Quite a lot of industries have made headway in developing their capacity to collect and analyse data:

- Credit card companies observe each and every buy of their customers which can then use to identify fraudulent activities based on the same data of customers. This is done by identifying purchases who have irregular transactions or unusual patterns in transactions.
- Subscribers of mobile phone companies use data from her subscribers to analyse call patterns to determine preference of each customers and project promotions, call patterns to competitors, advertisement and subscribers.
- Huge companies like Instagram, whatsapp, LinkedIn and Facebook main products is data itself. The progress of these companies depends heavily on the amount of data they are able to gather which holds core values as the data grows.
- World health organization or medical companies now use data to predict ailments, pandemics and potential outbreak in densely populated area.

Data analysis can be divided into two forms which can be used for extracting models defining vital classes or used to predict future trends. This branch of data analysis is useful for E-Commerce companies to push product to customers. The world today in the case of corona virus use data analysis to predict disease outbreak in different communities and countries. Classification used to predict categorical class labels and prediction for continuous valued functions are two forms of data analysis.

Predictive analytics uses historical data gathered over time to predict future events or happenings. Basically most historical data after been analysed are used to build mathematical model that captures main trends. The predictive model derived is then used on present data to predict happenings and events in the future or provide suggestions to take for prime outcomes.

Predictive analytics over the years has received huge chunk of attention in recent years due to advances in technology predominantly in the region of machine learning, artificial intelligence and big data.

Predictive models are used for forecasting disease outbreak, weather and also energy consumption. Huge amount of data are made regularly available for model creation which can then be used for making decisions in the future. There are four steps as shown in figure 1 below in defining predictive analytics workflow from accessing the data to integrating it.

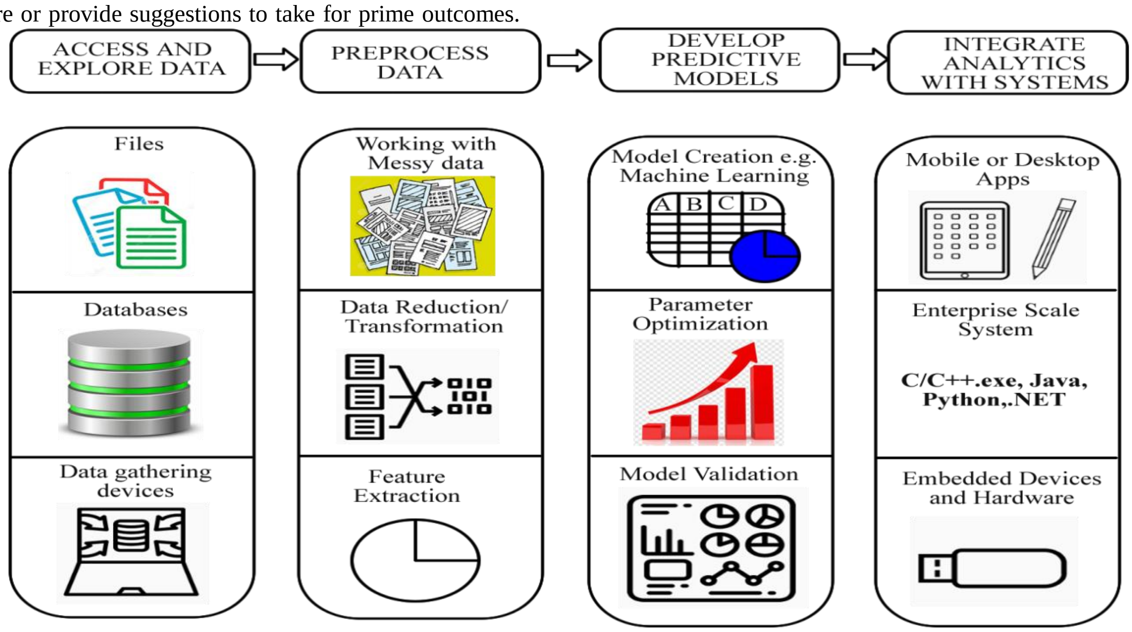

Figure 1 – Predictive Analytics Workflow

1.1 Domains of predictive analytics

Predictive analytics assist individuals and organization in the industries such as healthcare, finance, pharmaceuticals, aerospace, automotive and manufacturing.

- Developing driver assistance technology and autonomous vehicles using sensor data from connected vehicles – Automotive
- Monitoring aircraft engine health which helps in reducing cost of maintenance - Aerospace
- Forecasting the price of electricity and also demand – Energy Production
- Creating and developing risk models to predict high risk loans and credit risk – financial services
- Using pattern-detection algorithms to detect different diseases like covid-19 and typhoid

2. LITERATURE SURVEY

Zlatan Car et al, [1] used a multilayer perceptron to model the spread of Covid-19 infection. A total of 48384 ANNs were trained for each patient grouped deceased, recovered, and infected with each model evaluated using the coefficient of determination. The best models achieved consists of 4 hidden layers with 4 neurons in each of those layers, and use a ReLU activation function, with R2 scores of 0.98599 for confirmed, 0.99429 for deceased, and 0.97941 for recovered patient models.

Sina F. Ardabili et al, [2] worked on the comparative analysis on using machine learning techniques and soft computing models to predict Covid-19 disease outbreak. The work provided an initial benchmark to establish the potentials of machine learning in providing future analysis of outbreak in the medical field.

Sirage Zeynu A. [3] provided a detail and concise analysis of covid-19 outbreak in Ethiopia using Support Vector Machine (SVM) and Polynomial Regression (PR) models which were used to study and predict covid-19 aggressive risk. The model were used to predict the disease outbreak for the next 30 days. At the end of the research, the experiment showed the SVM performed overall better than PR

Gergo Pinter [4] recommended the Adaptive network-based fuzzy inference system (ANFIS) and Multi-layered perceptron-imperialist competitive algorithm (MLP-ICA) to predict time series of infected individuals and mortality rate of covid-19 patients in Hungary. After successively using the model, it predicted by late May 2020, the outbreak the total morality will drop substantially.

Durga M.M & Meet K.S [5] thesis was on using SVM (Support Vector Machines), RF (Random Forests) and ANN (Artificial Neural Network) machine learning techniques to predict the deadly covid-19 virus which helped increase the speed of disease identification resulting in reduced mortality rate. Evaluating the results acquired from experiments stated that Random Forest (RF) was performed better compared to other algorithms.

Richa T. & Sumeet M. [6] worked on using machine learning techniques to predict and visualize the outbreak of covid-19 by using Polynomial Regression and Support Vector Regression.

Aman Khakharia et al, [7] recommended using 9 different machine learning in densely populated countries to predict covid-19 outbreak. The proposed prediction models forecasted the count of new cases likely to arise for successive 5 days using 9 different machine learning algorithms. The highest accuracy of 99.93% was achieved for Ethiopia using Auto-Regressive Moving Average (ARMA) averaged over the next 5 days.

In China, Wei Feng & Ying-Hui Quan [8] predicted that the covid-19 epidemic to end after March 20, 2020 and cause 52,000–68,000 infections and about 2400 deaths. There data trends showed that the speedy and active approaches to decrease human exposure taken in China, such as restraint on population mobility and interpersonal contact rates, strict quarantine on migrants, have already had good impacts on control of the epidemic.

Zeynep Ceylan [9] stated there is an urgent need to observe and predict covid-19 prevalence to control this spread more effectively. In the study, Auto-Regressive Integrated Moving Average (ARIMA) models were used to predict the epidemiological trend of COVID-19 prevalence of Italy, Spain, and France, the most affected countries of Europe.

COVID-19 remains a major pandemic currently threatening all the countries of the world. Kayode Ayinde et al [10], provided a comparative analysis of models and estimators using Nigeria as case study. Quartic Linear Regression Model with an auto correlated error of order 1 (AR(1)) and found the Ordinary Least Squares, Cochrane Orcutt, Hildreth – Lu, and Prais-Winsten and Least Absolute Deviation (LAD) estimators useful to estimate the models' parameters. It was recommended that the daily cumulative forecast values of the LAD estimator for May and June 2020 with a 99% confidence level.

Parul A. et al [11] used deep learning models for predicting the number of novel coronavirus (COVID-19) using union territories of India as case study. Recurrent neural network (RNN) based long-short term memory (LSTM) variants such as Deep LSTM, Convolutional LSTM and Bidirectional LSTM were applied on the Indian dataset to predict the number of positive cases. LSTM model with minimum error was chosen for predicting daily and weekly cases. It was observed that the proposed method yields high accuracy for short term prediction with error less than 3% for daily predictions and less than 8% for weekly predictions.

Vikas C. & Saurabh P. [12] worked on using on the application of machine learning time series analysis (naïve method, simple average, moving average, single exponential smoothing, Holt linear trend method, Holt Winter method and ARIMA, for comparison) for prediction of Covid-19 Pandemic. Using the number of cases, deaths toll and recovery cases worldwide within a given period of time, the research was able to give predictions of spread of coronavirus.

Predicting mortality rate among confirmed CoVID-19 patients in South Korea using machine learning (logistic regression, support vector machine, K nearest neighbour, random forest and gradient boosting) and then deploying the best performing algorithm as an open-source online prediction tool for decision-making was worked on by Ashis K.D. et al [13]. The logistic regression algorithm was the best performer.

Ogundokun R.O. et al [14] worked on predictive modelling of COVID-19 confirmed cases in Nigeria. The ordinary least squares estimator was adopted to measure the impact of travelling history and contacts on the spread of COVID-19 in Nigeria and made a prediction. The prediction ensured that the government was right in enforcing early lockdown and temporal ban of travel agency which in turn prepared the government for re-opening.

Mostafa Salaheld and Abdelsalam Abotaleb [15] predicted Covid-19 cases using ARIMA and exponential growth model. The results of the research suggested that the exponential growth model is better than ARIMA models for forecasting the COVID-19 cases

3. METHODOLOGY

3.1 Knowledge Discovery

The non-trivial procedure of classifying valid, novel and potentially useful and ultimately understandable patterns in data is known as knowledge discovery [16]. It is a multi-step process for the creation of knowledge from structure and unstructured sources. The subsequent knowledge discovered needs to be in a machine readable and machine interpretable format. The result of the discovery must also represent knowledge in a manner that enables inferencing. The ultimate goal of knowledge discovery is to extract knowledge or information from inferior level data such as a database. The intent of knowledge discovery is to harvest huge amount of information by identifying and analysing patterns in raw data. Data mining is a step in the Knowledge Discovery process that entails finding patterns, anomalies and correlations within a large data sets to predict events.

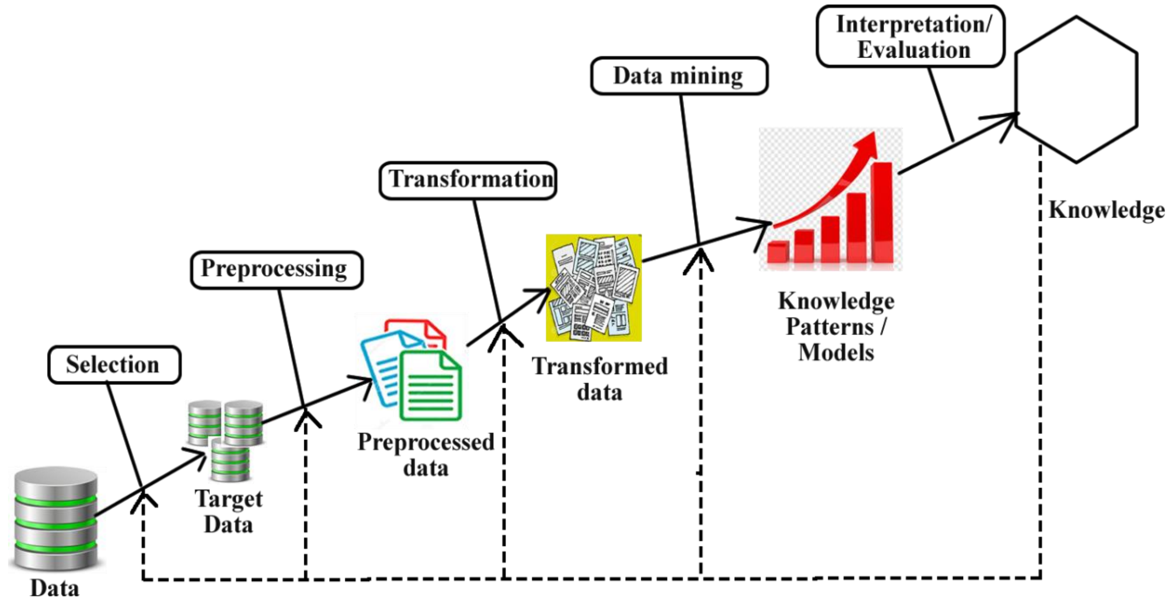

Figure 2 – The Knowledge Discovery Process

Knowledge discovery process is divided into five (5) different processes where raw data is passed. The idea of Selection is picking the target data that a researcher wants to mine or basically collecting data from multiple sources available. The data sources may include both structured and unstructured data which must then be made ready for production. The data pre-processing stage consumes 80% of the time of any data mining project. Data from different sources should be selected at the beginning, cleaning which is the process of cleaning the data by smoothing noisy data and filling in missing values, transformed (making data useful) to the format to which the result structured, formatted, anonymised and constructed if required. The next stage is the data transformation which involves smoothing, aggregation, generalization (replacing low-level data with high-level concepts), normalization (scaling data up and down into defined range) and attribute construction). The next stage is discovering knowledge patterns and models which are used to determine data patterns based on the objective of the project which requires a selected modelling techniques for the prepared dataset, creating a scenario to check and test the quality and validity of the model then finally deploy the model. The information of knowledge discovered during the data mining process should be made easy to comprehend for non-technical individuals.

3.2 Data Structure

Data structure is a technique of gathering and collecting in an organized form in a way that it can execute different operations on same data all done in an effective manner. It is a way of bringing together data in a computer so that it can be used effectively. Data grow accordingly and most times are increasingly unstructured. Numerous perceptions and insights can be extracted from the unstructured, quasi- or semi-structured data in the call center data.

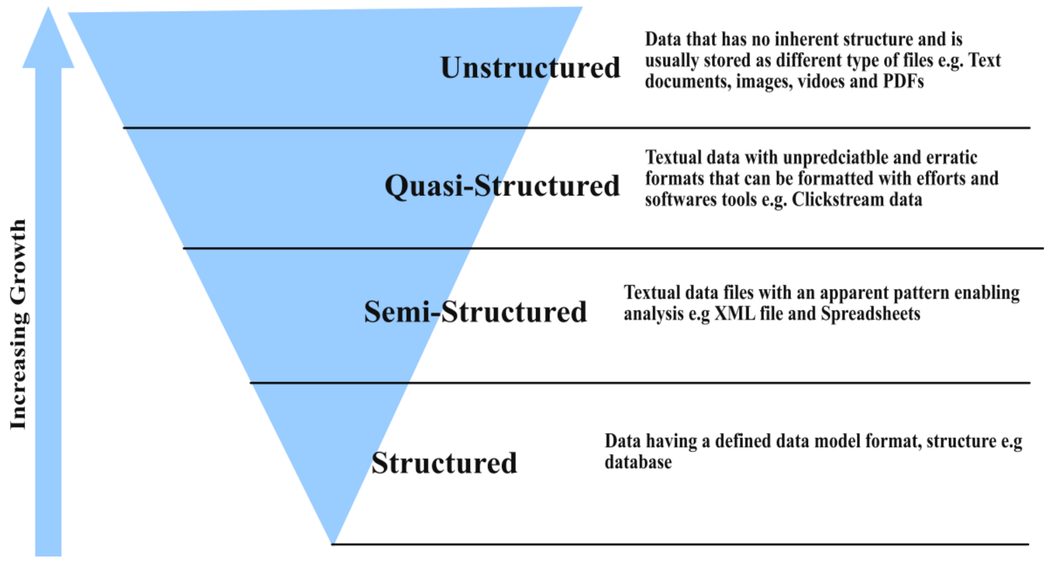

Figure 3 - Data Growth is increasingly unstructured

Types of data structures include;

- Structured data: contains data which have format, data type and structure (mostly like online analytical processing [OLAP] data cubes, CSV files, transaction data, traditional RDBMS and spreadsheets).

- Semi-structured data: is the data that do not conform to a data model but has some form of structure in it. They are also textual data files with a noticeable arrangement and pattern which enables parsing. Examples are Extensible Markup Language (XML) data files which describes itself and also defined by XML.
- Quasi-structured data: they are textual data with unpredictable and erratic data formats that can be formatted with tools, time and effort. Example is a web clickstream data that may contain inconsistencies in data values and formats
- Unstructured data: are data who are in different formats and has no integral arrangement and structure. Example are PDFs, text documents, video and images.

3.3 Exponential Growth

The exponential growth describes the way a quantity may increase over the time. It occurs when the instantaneous rate of change of a quantity with respect to time is proportional to the quantity. Basically it is a pattern that shows sharper increases over time. To give a better understanding on how a pandemic work using exponential growth, for example, if there is a glass slide held below a microscope which contains of explicit germs, this germ has a property to double every day. On the 1^{st} day, there is assumed to be one, on the succeeding day, it is two, on the third day there are four and the fourth day eight, and so on. At exactly the 60^{th} day, the slide is full. Unexpectedly, not until the 54^{th} day is the whole glass slide full which means that the slide goes from 1% to 100% in less than a week and hence, displays a property called exponential growth. This is how a Pandemic works in real life. The disease outbreak is fairly unnoticeable in the beginning, then, once it reaches a significant value, the growth to maxima is extremely quick as described in figure 4 below.

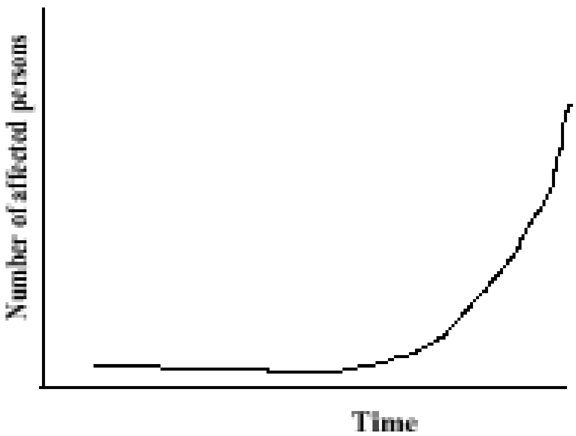

Figure 4 – Exponential Growth

A peculiar thing regarding a pandemic is that it cannot go on endlessly. The virus will ultimately stop locating people to infect and ultimately will slow down the count. This is known as logistic growth and the curve is known as a sigmoid as shown in figure 5. There would be a point where the available heath resources won’t be able to cater for all affected individuals in the country as infections spreads.

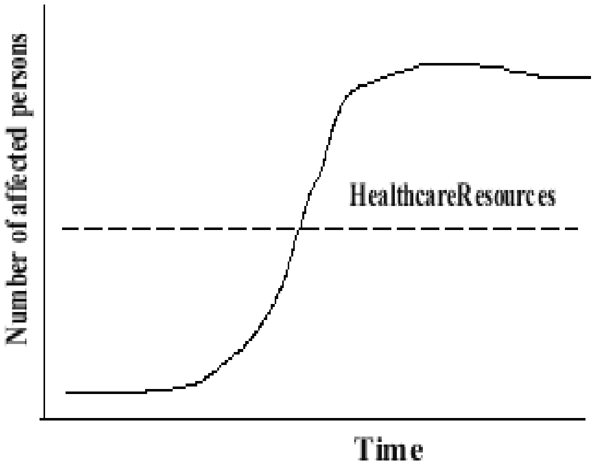

Figure 5 – Sigmoid level attained

At every point in the curve, the actual running total cases of the current day can be gotten. Digging into statistics, it is discovered that by plotting the slope of each day, the number of new cases per day can be gotten. At the very beginning, there are fewer new cases accompanied by a sharp rise afterwards.

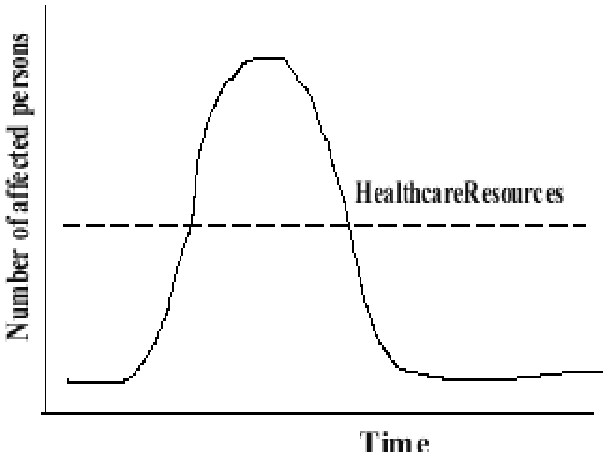

Figure 6 - The curve attained if you calculate the slope against each point of the sigmoid

In figure 6, the peak of the curve will definitely consume the healthcare systems which is the amount of resources available to the whole population of the country for taking care of affected individuals at any given point in time.

The total number of affected individuals overall cannot be taken care of by the health care system as available resources are scare, hence the best solution is to flatten the curve so as to drastically reduce the total number of cases at any given point in time as close to the healthcare line as possible as shown in figure 7.

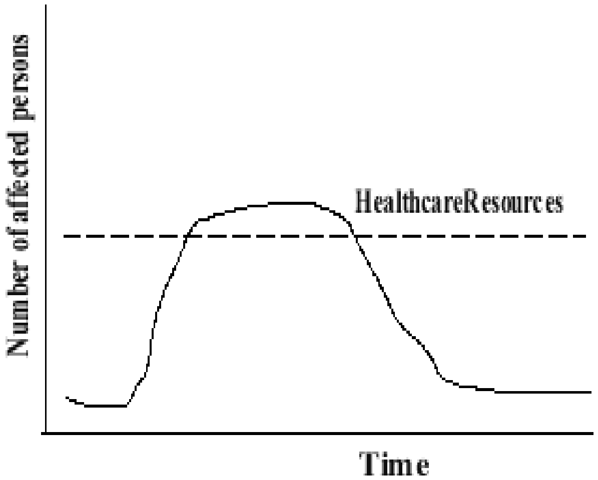

Figure 7 – Solution is to flatten the curve by getting the curve as close as possible to the health resources

This distributes the duration of the whole process for a while, but since the resources available in the healthcare system can only attain to a limited number of cases at a time, the fatalities are considerably lower.

Prophet is a process for forecasting time series data based on an addictive model where non-linear trends are fit with yearly, weekly and daily seasonality including holiday effects. The model works best with time series that has a strong seasonal effects and several seasons of historical data. It is robust to missing data and shifts in the trend and typically handles outlier well. It is an open source software released by Facebook's Core Data science team. Anaconda is an open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning, predictive analysis, data processing) with the aim to simplify package management and deployment. Reasons for using prophet includes the following:

- It is accurate and fast: It is used in many applications across Facebook for generating consistent and reliable forecasts for planning and goal setting. Facebook finds it to execute better than any other approach in the majority of cases. It fit models in Stan, so that it can forecasts in just a few seconds.
- It is fully automatic: It is used to get a practical and reasonable forecast on messy data with no manual effort. It is robust to outliers, missing data, and dramatic changes in time series.
- It provides tuneable forecasts - The Prophet library process comprises of countless possibilities for users to tweak and adjust forecasts. A human-interpretable parameters can be used to improve forecast by adding different domain knowledge.
- It is available in R or Python - Facebook has executed the Prophet Library procedure in both R and Python. Both share similar fundamental and underlying Stan code for fitting.

4. RESULTS ANALYSIS

The second wave of covid-19 virus has officially began in so many European countries such as Italy, India, Korea and Wuhan (city where the virus started). After the ease of lockdown, it was expected the many countries will experience the second wave of the virus. For the purpose of this study, Italy, Korea, India and Wuhan will be considered.

Considering data from John Hopkins data centre, the total number of new cases in Italy, Wuhan, Korea and India will be looked out for three months from the time the first case was identified. Data from 31st December 2019 to 22nd March 2020 was considered. Also, total number of confirmed global cases, deaths and recovered cases from 22nd January 2020 to 17th March 2020 (3months) was considered for analysis. The first stage is importing all the important libraries needed for the model such as Pandas which is an extremely fast and flexible data analysis and manipulation tool and allows to store and manipulate tabular data. Other visualisation libraries imported are matplotlib, seaborn and plotly through the command prompt in anaconda. Figure 8 shows the process of importing the required data and visualization libraries using Jupyter Notebook in anaconda.

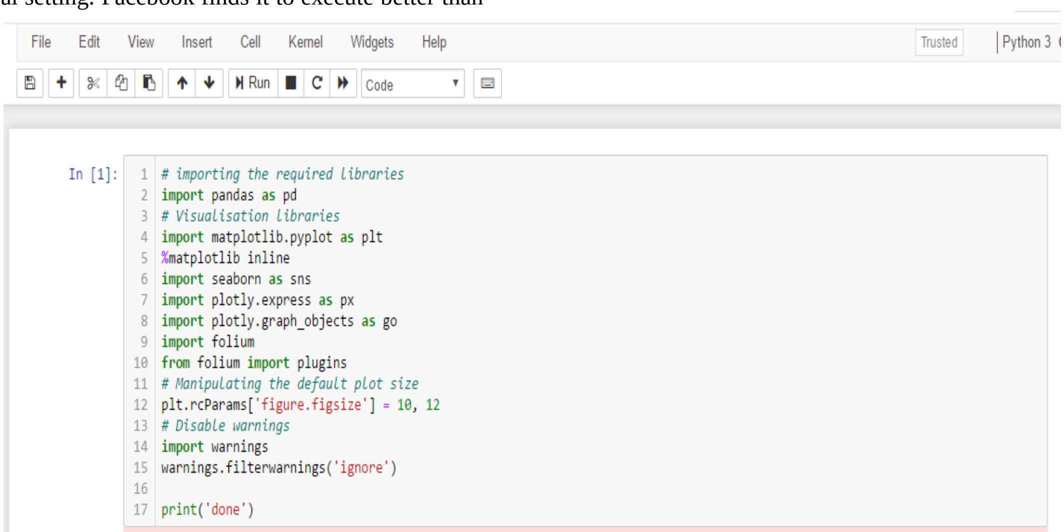

Figure 8 – Importing data and visualization libraries

The rise of covid-19 cases across India rose exponentially from March 5th 2020 and continued the upward trend.

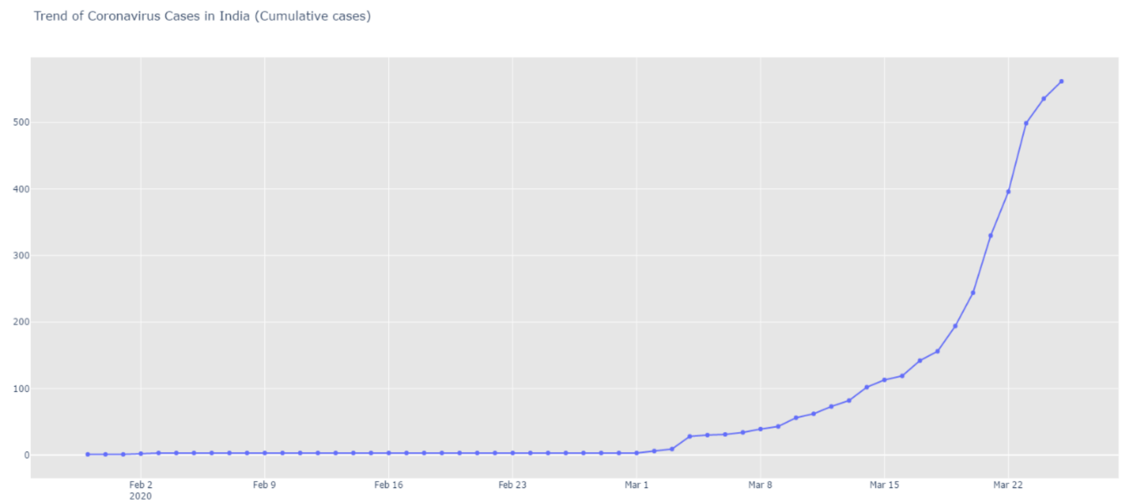

Figure 9 - Graphs depicting the trends of the rise of coronavirus cases across India.

Using Plotly library in Python to obtain graphs illustrating the trend of the rise of Covid-19 cases across India, it was noticed the trend continued even till 22nd March 2020. Figure 9 shows the graphs depicting the trends of the rise of coronavirus cases across India and figure 10 shows the daily basics cases in India.

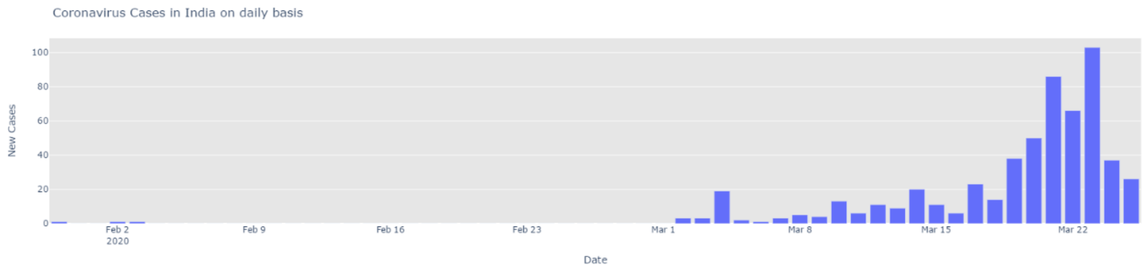

Figure 10 – Coronavirus cases showing the daily basics cases in India.

Noticing the trend in India, 23rd March signified the time India crossed the first 100 cases and check if the trend is similar to that available South Korea, Italy and the city of Wuhan. As India crossed the 500 cases barrier, it is important to be able to contain the situation in coming days. It was also noticed that as many countries hit the 100 cases mark, the number of coronavirus patients in the hospital had doubled exponentially. Figure 11, 12, 13 shows the confirmed cases in India, Italy, Wuhan and South Korea respectively.

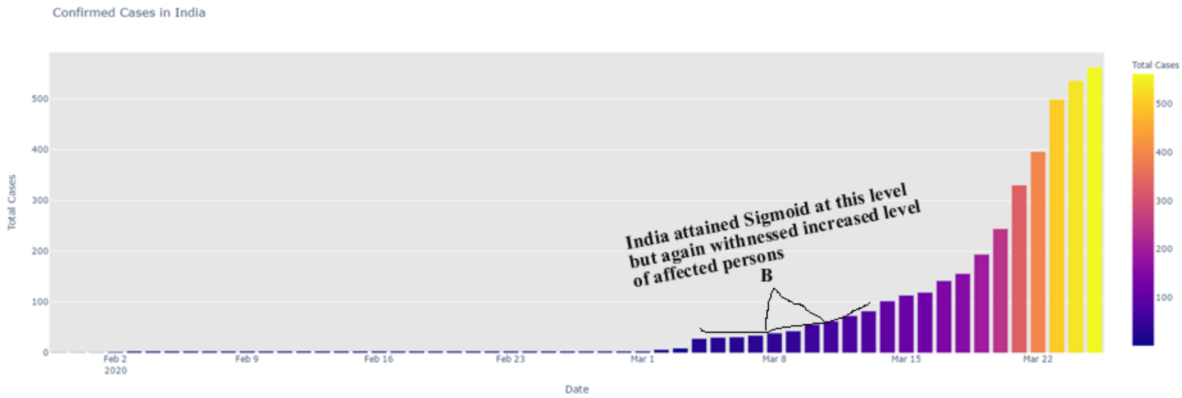

Fig 11 - Confirmed cases in India with Sigmoid level attained at point B. it didn't last as the country witnessed exponential growth

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.36, January 2021 – www.ijais.org

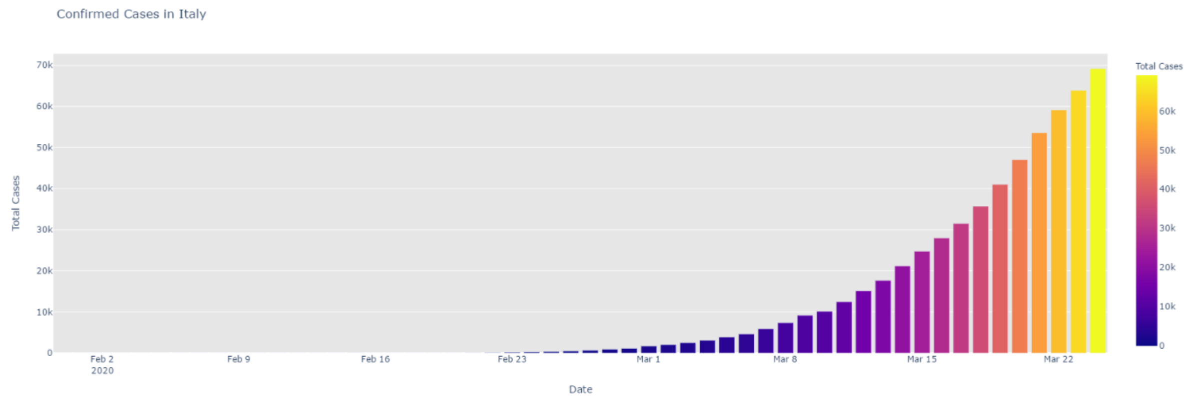

Fig 12 - Confirmed cases in Wuhan

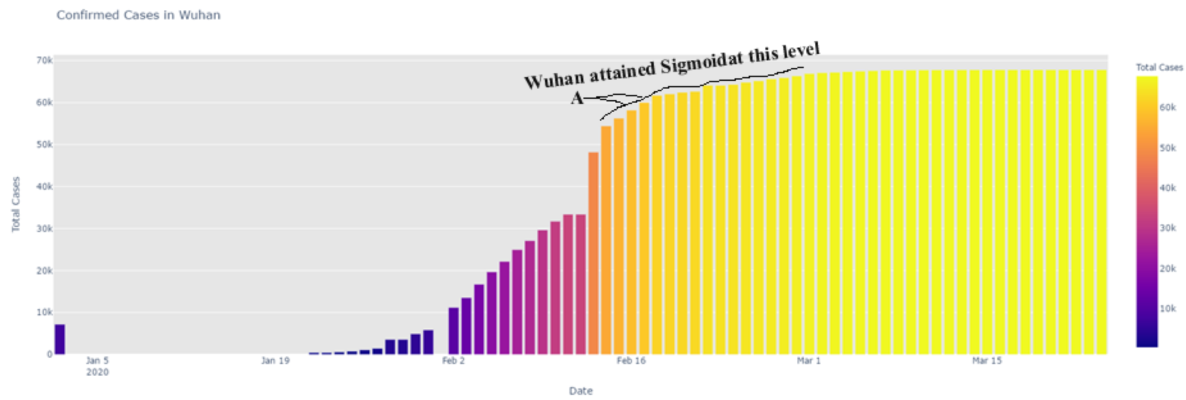

Fig 13 - Confirmed cases in Wuhan with sigmoid level attained at that point A

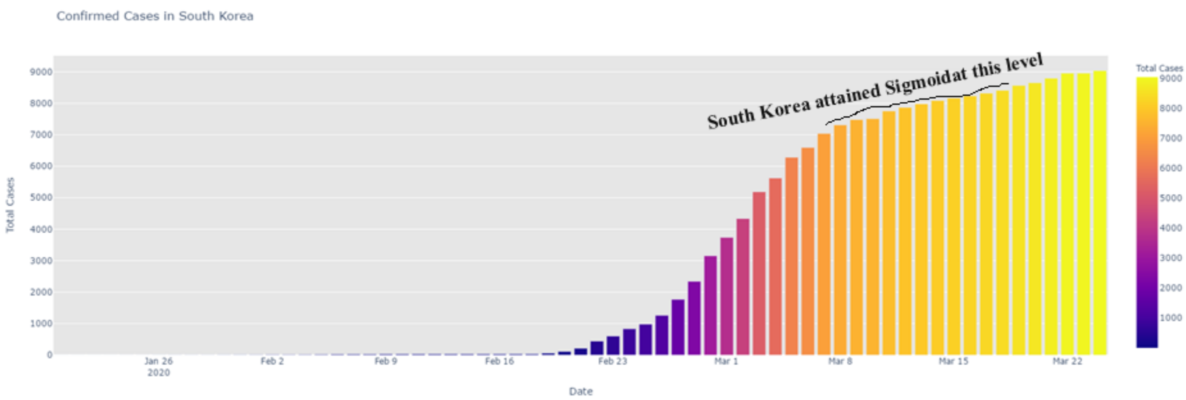

Fig 14 - Confirmed cases in South Korea with sigmoid level attained at the level stated

From the visualization above in figure 11 to 14, it can be concluded that:

- Confirmed cases in Italy continually kept rising exponentially with a certain fixed pattern
- Confirmed cases in India continually kept rising exponentially with no fixed pattern. The graph attained sigmoid early but later increased exponentially with no particular fixed pattern which was as a result of very less test done in India.
- Confirmed cases in South Korea is rising gradually with sigmoid attained at some level. The government was able to reduce the spike in infected persons.
- In Wuhan, there have been almost a negligible number confirmed cases in a week.

Figure 15 shows the comparison between the increase of cases in South Korea, Italy and India.

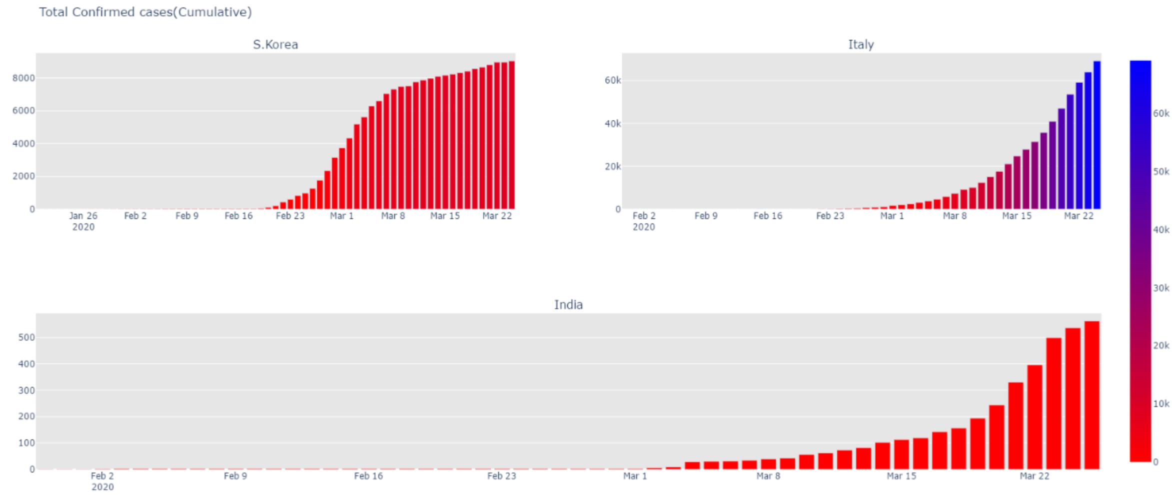

Figure 15 - Comparison between the rise of cases in South Korea, Italy and India

Countries all over the world witnessed increase in spike of infected people after crossing the 100 cases barrier with Italy having a huge spike after crossing 100 cases. This was probably due to that fact of high rate of infection and government imposing lockdown very late. Figure 16 show trend after crossing 100 cases in each country stated in the graph.

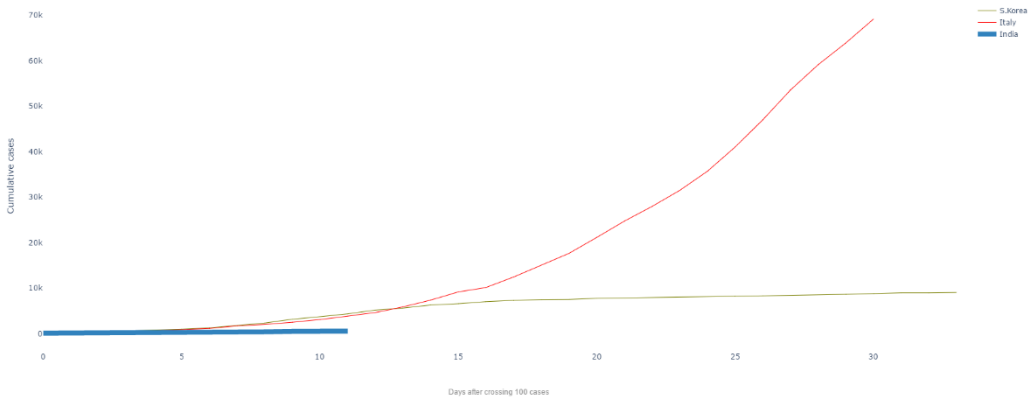

Figure 16 - Trend after crossing 100 cases

4.1 Forecasting Total Number of Confirmed COVID-19 Cases Worldwide with Prophet (Base model)

The next phrase of the study is generate a week ahead forecast of confirmed cases of COVID-19 using the Prophet library available on python with explicit prediction intervals by creating a base model both with and without tweaking of seasonality-related parameters and additional regressors.

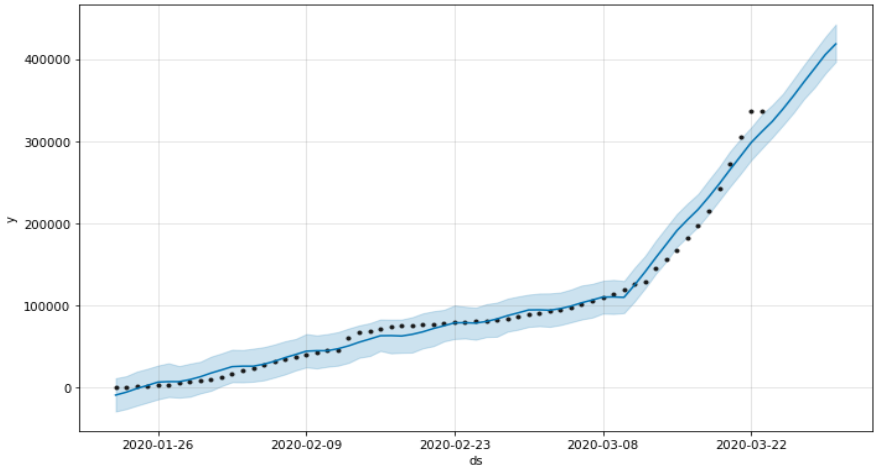

Figure 17 - Forecasting Confirmed COVID-19 Cases Worldwide with Prophet (Base model)

Figure 17 represents the degree of accuracy of the forecast made by the prophet library from 24th March 2020. This is denoted by the blue line running through the dotted black line on the graph which shown some level of accuracy since the blue painted path is within the dotted line. This means that the model fairly forecasted the number of confirmed covid-19 cases worldwide.

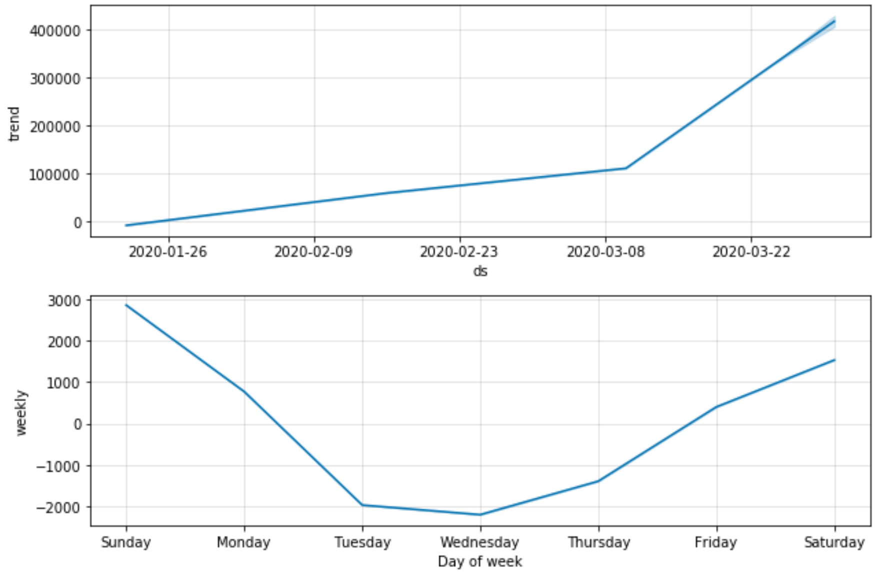

Figure 18 – Trend and Weekly forecast of confirmed covid-19 cases

Table 1 represents the future confirmed cases of covid-19 with date, and upper and lower limit of y value. The future confirmed values of confirmed covid-19 cases from 26th March 2020 to 30th March 2020 is represented in the table 1 below with the upper and lower limit value give for each day.

Table 1 – Future Confirmed cases of covid-19 with date, and upper and lower limit of y value from 26th to 30th March 2020

	ds	yhat	yhat_lower	yhat_upper
64	2020-03-26	355136.872940	334472.088335	375387.933812
65	2020-03-27	372235.326900	351491.294083	393066.921221
66	2020-03-28	388674.964099	365501.451378	409975.394466
67	2020-03-29	388674.964099	382371.807809	426969.504459
68	2020-03-30	405307.954628	396725.533594	442406.661622

4.2 Forecasting Worldwide Deaths using Prophet (Base model)

For the predicted values for number of deaths, a week ahead forecast of death related cases of COVID-19 is generated using the Machine Learning library available on python – Prophet, with 95% prediction interval by creating a base model with no tweaking of seasonality-related parameters and additional regressors.

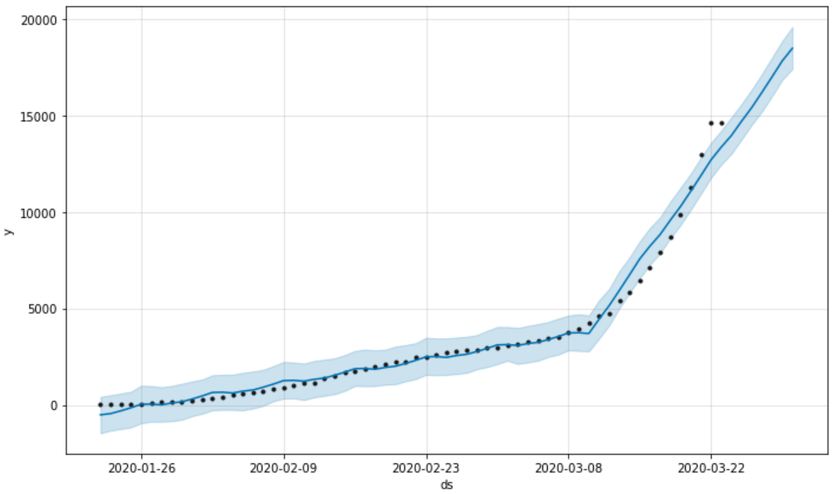

Figure 19 - Forecasting Death related COVID-19 Cases Worldwide with Prophet (Base model)

Figure 18 represents the degree of accuracy of the forecast made by the prophet library from 24th March 2020. This is denoted by the blue line running through the dotted black line on the graph which shown some level of accuracy since the blue painted path is within the dotted line. This means that the model fairly forecasted the number of deaths related covid-19 cases worldwide.

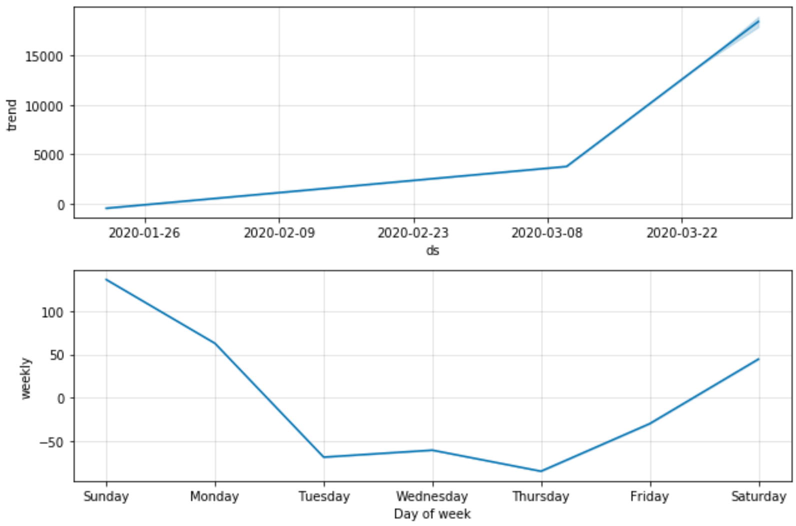

Figure 20 – Trend and Weekly forecast of death related covid-19 cases

Table 2 represents the future death cases of covid-19 with date, and upper and lower limit of y value. The future death values of confirmed covid-19 cases from 26th March 2020 to 30th March 2020 is represented in the table 2 below with the upper and lower limit value give for each day.

Table 2 - Future death related cases of covid-19 with date, and upper and lower limit of y value from 26th to 30th March 2020

	ds	yhat	yhat_lower	yhat_upper
64	2020-03-26	15411.877522	14523.268722	16350.848318
65	2020-03-27	16199.617603	15209.619681	17160.506848
66	2020-03-28	17006.904589	16008.876718	18017.911368
67	2020-03-29	17831.505383	16855.081118	18875.581810
68	2020-03-30	18490.986176	17418.107685	19602.994950

4.3 Forecasting Worldwide Recovered Cases with Prophet (Base model)

For the predicted values for number of recovered cases, a week ahead forecast of recovered cases of COVID-19 is generated using the Machine Learning library available on python – Prophet, with 95% prediction interval by creating a base model with no tweaking of seasonality-related parameters and additional regressors.

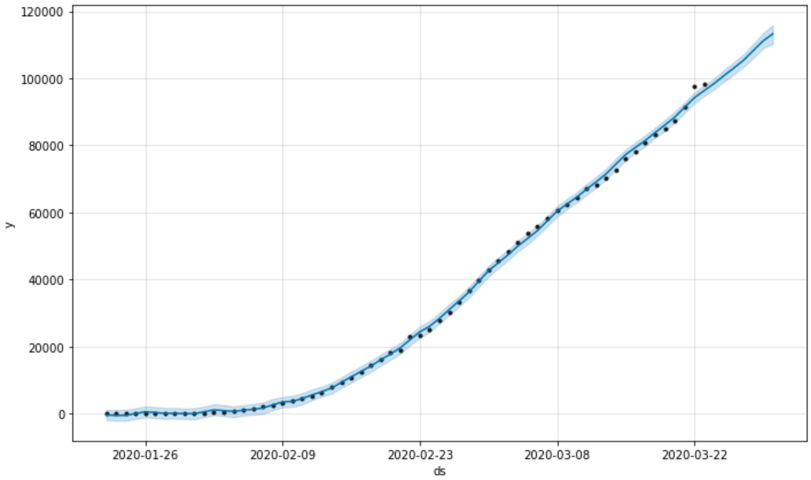

Figure 21 - Forecasting Recovered COVID-19 Cases Worldwide with Prophet (Base model)

Figure 21 represents the degree of accuracy of the forecast made by the prophet library from 24th March 2020. This is denoted by the blue line running through the dotted black line on the graph which shown some level of accuracy since the blue painted path is within the dotted line. This means that the model fairly forecasted the number of recovered covid-19 cases worldwide.

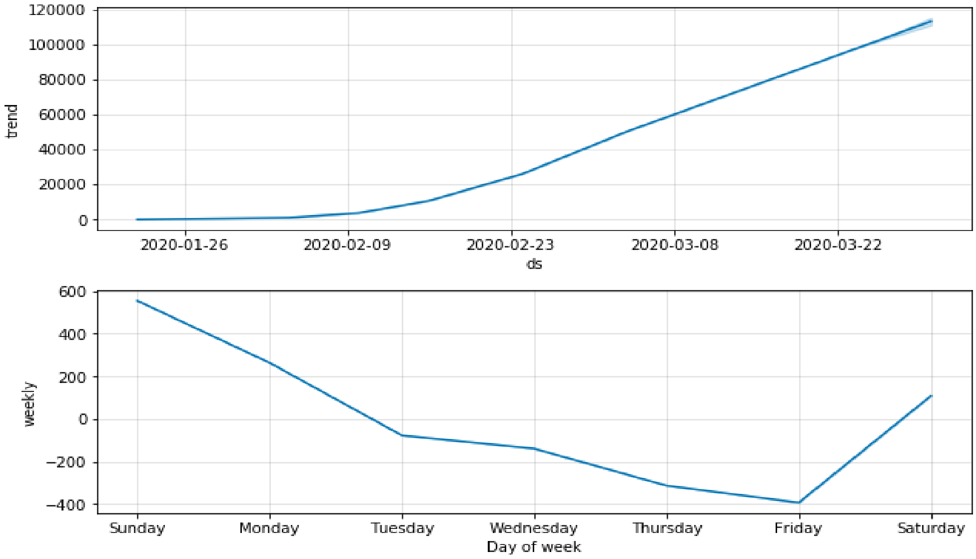

Figure 22 – Trend and Weekly forecast of recovered covid-19 cases

5. CONCLUSION

As past events and developments move forward, inventions and innovations leave huge amounts of data behind. The result of the proposed model are shown in the paper by considering datasets of Italy, Wuhan, South Korea and India from 31st January 2020 to 25th March 2020. This paper proposes the use of Machine learning library – (Prophet) used to predict future values by creating a base model with no tweaking of seasonality-related parameters and additional regressors. Observing the data in India for example, the number of infected cases is rising exponentially just as it is rising in Italy, South Korea and Wuhan. This is also available in most countries. Most countries in Europe have crossed the 100,000 cases of infected persons already. More test centres should be provided to combat the pandemic fully. In other for government to prepare for any future pandemic, the key to moderation is testing as early testing will shift the curve down to the health resources barrier meaning health centre would be able to combat the deadly disease when necessary.

6. REFERENCES

[1] Zlatan Car, Sandi Baressi Šegota , Nikola Anđelić, Ivan Lorencin, and Vedran Mrzljak, “Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron”, 2020 Computational and Mathematical Methods in Medicine, Volume 2020, Article ID 5714714, pg. 1 - 10. DOI://doi.org/10.1155/2020/5714714

[2] Sina F. Ardabili, Amir Mosavi, Pedram Ghamisi, Filip Ferdinand, Annamaria R. Varkonyi-Koczy, Uwe Reuter, Timon Rabczuk and Peter M. Atkinson, 2020 “COVID-19 Outbreak Prediction with Machine Learning Algorithms, 13, 249; pg. 1 – 36. doi: 10.3390/a13100249

[3] Sirage Zeynu Ahmed “Analysis and forecasting the outbreak of covid-19 in Ethiopia using Machine learning”. European Journal of Computer Science and Information Technology Vol.8, No.4, pp.1-13, August, 2020

[4] Gergo Pinter, Imre Felde, Amir Mosavi, Pedram Ghamisi and Richard Gloaguen, 2017 “COVID-19 Pandemic Prediction for Hungary; a Hybrid Machine Learning Approach” Hungarian State and the European Union.

[5] Durga Mahesh Matta and Meet Kumar Saraf, , May 2020 .“Prediction of COVID-19 using Machine Learning Techniques”. Bachelor of Science in Computer Science thesis, Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden.

[6] Richa Tamhane and Sumeet Mulge, “Prediction of COVID-19 Outbreak using Machine Learning”, May 2020. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 (p-ISSN: 2395-0072), Volume: 07 Issue: 05

[7] Aman Khakharia, Vruddhi Shah, Sankalp Jain, Jash Shah, Amanshu Tiwari, Prathamesh Daphal, Mahesh Warang and Ninad Mehendale, “Outbreak Prediction of COVID‑ 19 for Dense and Populated Countries Using Machine Learning”; Annals of Data Science, DOI: https://doi.org/10.1007/s40745-020-00314-9, October 2020

[8] Wei Feng and Ying-Hui Quan, “Trend and forecasting of the COVID-19 outbreak in China”, February 2020. Journal of Infection. Pg 472 – 472, vol: 80 DOI: https://doi.org/10.1016/j.jinf.2020.02.014

[9] Zeynep Ceylan, April 2020 , “Estimation of COVID-19 prevalence in Italy, Spain, and France”, Science of the Total Environment Journal,.

[10] Kayode Ayinde, Adewale F. Lukman, Rauf I. Rauf, Olusegun O. Alabi, Charles E. Okon and Opeyemi E. Ayinde, May 2020. “Modeling Nigerian Covid-19 cases: A comparative analysis of models and estimators”, Chaos, Solitons and Fractals Nonlinear Science, and Nonequilibrium and Complex Phenomena

[11] Parul Arora, Himanshu Kumar and Bijaya Ketan Panigrahi, June 2020. “Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India”: Chaos, Solitons and Fractals Nonlinear Science, and Nonequilibrium and Complex Phenomena, DOI: https://doi.org/10.1016/j.chaos.2020.110017. Pg. 1 – 9.

[12] Vikas Chaurasia and Saurabh Pal, 2020. “Covid-19 Pandemic: Application of Machine Learning Time Series Analysis for Prediction of Human Future”; Research Square Journal DOI: https://doi.org/10.21203/rs.3.rs-39149/v1; pg. 1 – 6,

[13] Das AK, Mishra S and Saraswathy Gopalan S. 2020. Predicting CoVID-19 community mortality risk using machine learning and development of an online prognostic tool. PeerJ 8:e10083: DOI 10.7717/peerj.10083

[14] Roseline O. Ogundokun, Adewale F. Lukman, Golam B.M. Kibria, Joseph B. Awotunde and Benedita B. Aladeitan, August 2020: “Predictive modelling of COVID-19 confirmed cases in Nigeria Infectious Disease Modelling” Infectious Disease Modelling 5: 543 – 548.

[15] Mostafa Salaheldin and Abdelsalam Abotaleb. April, 2020; “Predicting COVID-19 Cases Using Some Statistical Models: An Application to the Cases Reported in China Italy and USA”; Academic Journal of Applied Mathematical Sciences. ISSN(e): 2415-2188, ISSN(p): 2415-5225, Vol. 6, Issue. 4, pp: 32-40, URL: https://arpgweb.com/journal/journal/17, DOI: https://doi.org/10.32861/ajams.64.32.40.