

IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

VOLUME : 5 ISSUE : 8 Print / Issue Publication Date: 18-Feb-2021

ISSN : 2455-2143

DOI : 10.33564/IJEAST.2020.v05i08.015

Indexed In

Google SCHOLAR OPEN ACCESS doi Crossref

WWW.IJEAST.COM

editor@ijeast.com

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 102-110
Published Online December 2020 in IJEAST (http://www.ijeast.com)

A COMPARATIVE ANALYSIS OF COVID–19 BETWEEN INDIA AND OTHERS USING DATA MINING AND MACHINE LEARNING APPROACH

Kanika Bhalla
Department of Electrical Engineering
National Taipei University of Technology, Taipei, Taiwan

Ashish Kumar
Department of Electrical Engineering
National Taipei University of Technology, Taipei, Taiwan

Abstract—Novel coronavirus has caused a global pandemic which leads to acute respiratory disorder in humans. In this study, analysis of the transmission of communicable COVID19 disease in India is done. The machine learning model presents the comparison of India with other countries during initial phase of virus spread in India. After that its comparison with initial hard-hit countries is also done. Finally, we also performed time series analysis for prediction using prophet for the next seven days showing confirmed, recovered and deaths that will happen.

Keywords—COVID-19, Time series analysis, Prophet, machine learning

I. Introduction And Related Work

The novel coronavirus is an ongoing global pandemic which is caused due to acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. It started in December 2019 originated from Wuhan, China [4][6].world health organization (WHO)proclaimed it as a public health emergency of international concerns in later January 2020 and a pandemic in early March 2020 [7][8].In the month of March COVID-19 pandemic spread across 190 countries and infected 4.2 lakh people where around 16,500 were killed of them. In India, as many as 562 COVID-19 cases were reported till 23rd March 2020. Of those, 37 recovered and 10 died. On March 24, 2020, the prime minister of India Narendra Modi declared a complete 21-day lockdown to combat the quickly spread of the contagious virus. This period was used to boost the healthcare infrastructure. On the other hand, cases of infections were rapidly rising in Germany, Italy, Spain, France, United Kingdom and the United States. Furthermore, led to the massive impact on the global economy and stock markets which can be seen till date. The severe outbreak of COVID-19 turned out to be a major international crisis which has adversely influenced the important aspects of daily life. Then Ministry of Civil Aviation decided to restrict the domestic or international flights in India till 21 day's lockdown period. Whereas, in highly affected areas people started to stock up on essential goods leading to shortage of essential stuff.

On 25th March, in the afternoon India reported its 9th death with 562 total confirmed cases due to COVID-19. Fresh cases from states like Manipur, Bihar, Gujrat and Madhya Pradesh were reported by the union ministry of health and family welfare (MoHFW). The coronavirus outbreak continued to spread quickly in tightly knit across all over the country. However, the question that Indians tried to answer was "Will India be able to tackle this pandemic or are they were going to witness another Italy, South Korea, Wuhan or even worse".

As the coronavirus cases are increasing day by day, so exponential growth will be observed during its graphical representation. This is because of the replication property of the virus which doubles with every day that passes. Every country is trying to achieve a sigmoid pattern which tell us that even if the country is not able to find a cure for the virus they have boosted their healthcare facilities so that the number of deaths could be less.

Many researchers have conducted research in the field of prediction and forecasting of virus and their spreads in case of epidemics and pandemics. Andrea et.al [6] formulated a mathematical expression to predict the H1N1 influenza pandemic spreading. It was found that the pathogen transmissibility, mobility features, transportation network, demographic profile and mixed virus patterns were the deciding factors. Furthermore, result classification was done based on history of travel.

Stephen et.al [7] produced a mathematical model to predict disease outbreak with the use of population-mobility data, actual census and land-use. Results were presented by the bipartite graph which maps the total no of visitors, number of locations visited and number of contacts. Then Randi et.al [8] presented the graphical representation based on eigenvector centralization index for the quickly spread of the infectious disease in a primate social network. Moreover, it was based on the interaction between an individual and a group. This model attained 72.6% accuracy. Anjan et.al [9] developed a machine learning model for the prediction for the mortality of heart diseases with the use of Naive Bayesian Classifier. In addition advanced encryption techniques were also adopted to secure the patient's data records. Classification accuracy of 84.07% and 98.2% for the patient's data security was achieved. Olivera et.al [10] designed the Bayesian Monte Carlo probabilistic theory to predict the quickly spread of the infectious diseases. However, this model has a drawback of prior selection and posterior

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 97-101
Published Online December 2020 in IJEAST (http://www.ijeast.com)

probabilities. Caroline et.al [11] presented a mathematical model for the prediction of infectious diseases. In this study, spread of the influenza transmission between birds and human were analyzed. Pragmatic form of validation was done to check the accuracy of the mathematical model. Run et.al [12] did the COVID19 disease transmission analysis and as a result predicted the scale of the pandemic, recovery rate along with fatality rate. Song et. al [13] utilized machine learning and statistical methods of which time series was also a part. In the past it was mainly implemented in the area of infectious diseases to predict the monthly incidence of influenza in China. Lee et al. [14] developed a machine learning based time-series model with the collection of weekly time-series flu-related tweet counts which was further utilized to assess the real-time spread of influenza. This model was able to predict the scale of fast spreading of the novel corona virus across India and different countries in the earlier stages. In the proposed work visualizations were done with the use of python high level application programming interface (API) libraries for varied scenarios to derive precise inferences. This model also helped in the prediction of the spread of corona virus using prophet for the next 7 days.

The remainder of this paper was structured as follows. In Section I methodology for model was described in detail. Section III provides the experimental analysis and discussion. Conclusions and future study directions were concluded in section IV.

II. METHODOLOGY

In this section detailed description of proposed method is provided. Initially data set was obtained from John Hopkins Centre of Systems Science and Engineering and Ministry of Health and Family Welfare. Then four tasks were performed as follows: analysis of present condition in India, similar trend to Italy, South Korea and Wuhan, worldwide data exploration and predicting worldwide cases using prophet. Fig. 1. represents the data flow diagram of the machine learning model.

A. Datasets

This research work utilized publically available datasets [15] [16]. The datasets have entries from 30th January 2020 to 25th March 2020.A brief description about them is mentioned in Table 1.

TABLE-1 Description of Dataset Used

Dataset Name	Entries	Attributes
Covid complete.csv	**19220**	**Province/State, Country/Region, Latitude, Longitude, Confirmed, Death and Recovered.**
Covid cases in India.xlsx	**25 states**	**S.No., Name of State/UT, Total Confirmed cases (Indian National), Total confirmed cases (Foreign National), Cured and Death**
IndianCoordinates.xlsx	**36 states/UT**	**Name of State/UT, Latitude and Longitude**
Per day cases.csv	**56**	**Date, Total case, New case and Days after surpassing 100 cases**
Time series confirmed global.csv	**242**	**67**
Time series deaths global.csv	**242**	**67**
Time series recovered global.csv	**242**	**67**

B. Analysis of Present Condition in India

For the analysis of present condition in India we start from very beginning when coronavirus first patient come to India. A time line was prepared for that for increasing number of patients when they crossed the 100, 300and 500 mark. For finding the total number of cases we combined the Indian and Foreign nationals together to find out total cases (confirmed) and highlighted the states with maximum patients with color codes using python commands.

Total cases = Indian national + Foreign national

To find out the total active cases were moved the total number of cases with number of death and cured patents.

Total Active = Total cases – (Death cases + Cured cases)

Visualization of spread was done geographically using special library named folium in python to show total cases on Indian map. For getting the confirmed and recovered cases per state we created a plot to get the total cases and cured cases per state. Lastly for finding rise in coronavirus cases in India we used plotly library to make interactive graph for finding the trend of increasing cases in India (cumulative) and coronavirus cases on daily basis.

C. Similar Trend to Italy, South Korea and Wuhan

It was observed that initially hard-hit countries at that time were Italy, South Korea and China (Wuhan). Therefore, comparison was done with India. For this purpose, first we found the cumulative cases in each country for which we used plotly. express which was a high level API for creating figures to check which country was closed to achieve the sigmoid pattern. Then we made comparison between the countries based on inferences drawn from trend in cumulative cases. Wuhan was left out because of the results we got in section III. Lastly, we found the trend after crossing the 100 patients mark for comparison of India with other major hit countries like Italy and South Korea.

D. Worldwide Data Exploration

We combined attributes from different datasets to make data of country with date and state name having confirmed, recovered and death cases. As the data has multiple entries with same country name but with different entries. After performing the prior function, we separated the data based on date, latitude, longitude, confirmed, recovered and death cases. Then created

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 102-110
Published Online December 2020 in IJEAST (http://www.ijeast.com)

a worldwide case plot with confirmed, recovered and death cases.

E. Predicting Future Worldwide Cases using Prophet

Prophet was used for time series analysis. First, we found out the next seven days for prediction then predicted values of confirmed, recovered and death cases were found out for those days. We plot the predicted values and subplots to check for the trend, similarity and comparisons of our results.

III. Results And Discussion

From our study of analysis of condition in India Prime Minister of India Narenda Modi speech on 24th March 2020 was a good decision at right time to boost the health care resources in his speech he mentioned the following points:

- Proclaimed a nationwide lockdown for three-week starting midnight Next day he explained that it was mandatory step for controlling the Covid-19 increasing infection cycle.
- Public was appealed to not panic buying stuffs because people began crowding markets to stock up before the midnight deadline.
- The government issued a circular in which it was mentioned that essential services will remain open, as before and all essential commodities and medicines would be available. Banks, ATMs, petrol pumps, hospitals and grocery shops will continue to function.
- All transportation services such as roadways, air and rail were suspended until April 14.

A. Results from Analysis of Present Condition in India

As the background gradient was used therefore the results have highlighted data for the most number of cases for every attribute in data as shown in Fig. 2. The results show that Kerala had the greatest number of patients. The visualization tells about the other important details:

- Manipur and Mizoram reported their first case.
- Kerala has crossed Maharashtra in terms of highest number of confirmed cases.
- Haryana and Telangana had the highest count of confirmed Foreign National count.
- Till 25th of March, 9 people were died in India.
- Kerala, Maharashtra and Karnataka were the currently top three states with maximum number of confirmed cases.

	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death	Total cases
0	Andhra Pradesh	9	0	0	0	9
1	Bihar	3	0	0	1	3
2	Chhattisgarh	1	0	0	0	1
3	Delhi	30	1	6	1	31
4	Gujarat	32	1	0	1	33
5	Haryana	14	14	11	0	28
6	Himachal Pradesh	3	0	0	1	3
7	Karnataka	41	0	3	1	41
8	Kerala	101	8	4	0	109
9	Madhya Pradesh	9	0	0	0	9
10	Maharashtra	98	3	0	2	101
11	Manipur	1	0	0	0	1
12	Mizoram	1	0	0	0	1
13	Odisha	2	0	0	0	2
14	Puducherry	1	0	0	0	1
15	Punjab	29	0	0	1	29
16	Rajasthan	30	2	3	0	32
17	Tamil Nadu	16	2	1	0	18
18	Telengana	25	10	1	0	35

Fig.2.States with maximum number of cases.

From Fig. 3 it was observed that Kerala was the most affected state in initial phase having less resources with total 105 active cases followed by Maharashtra which had 99 active cases. Whereas, the North Indian region had less number of cases.

Fig. 4 illustrates the visualization spread geographically in India. The result of visualizations was pretty clear which showed the total number of cases at every state. The bigger the circle the greater number of cases present in that region. The top left part of graphs has zoom in and zoom out options.

Fig. 5 shows the confirmed vs recovered cases in India. It was found that net percentage of affected people are less in Kerala than in Maharashtra. Also, it showed that Uttar Pradesh and Haryana healthcare services did well with small amount of 35 and 27 people infected in their states respectively.

Name of State / UT	Total Active
Kerala	105
Maharashtra	99
Karnataka	37
Telengana	34
Gujarat	32
Rajasthan	29
Punjab	28
Uttar Pradesh	24
Delhi	24
Tamil Nadu	17
Haryana	17
Ladakh	13
Madhya Pradesh	9
Andhra Pradesh	9
West Bengal	8
Chandigarh	7
Jammu and Kashmir	6
Uttarakhand	4
Himachal Pradesh	2

Fig. 3. Active cases in affected State/UT.

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 97-101
Published Online December 2020 in IJEAST (http://www.ijeast.com)



Fig. 4. Visualizing spread in India

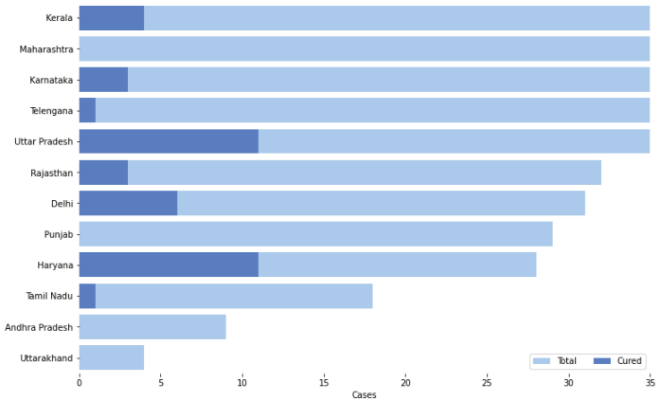


Fig. 5. Total and cured cases in India.

For finding the increasing number of corona virus cases, a scatter graph with liner and marker sort of visualization and bar graph for Indian scenario was created which showed India in early hit tackled the situation well as after 22nd March they started achieving the sigmoid pattern until the number of patients were low as represented in Fig. 6 and Fig. 7. The visualization has a hover and zoom effect as a particular section was highlighted. It was zoomed to draw more inferences. The zoomable option was available for even the timestamps between the dates.

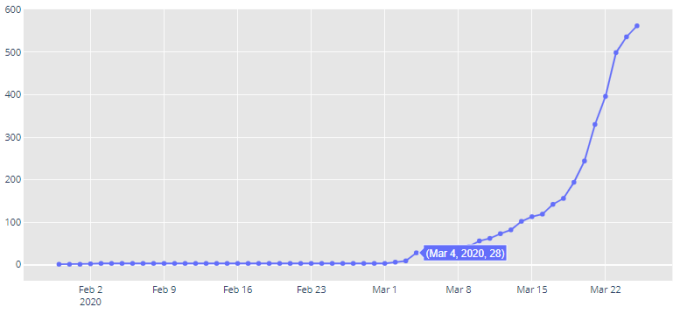


Fig. 6. Trend for coronavirus cases in India (cumulative cases).

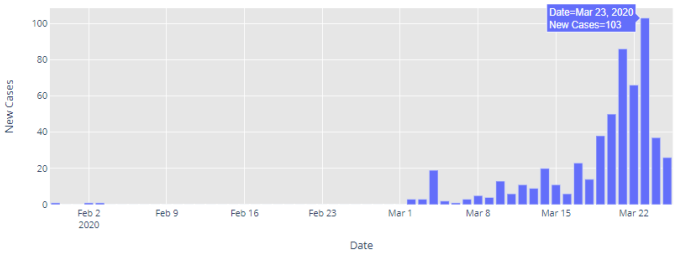


Fig. 7. Coronavirus cases in India on daily basis.

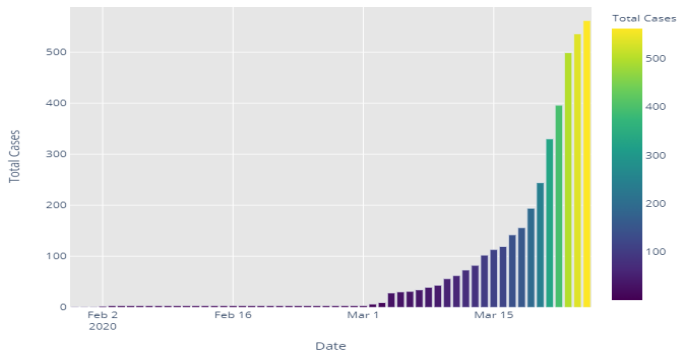
B. *Results from Similar Trend to Italy, South Korea and Wuhan*

It was observed that similar trend found in Wuhan, Italy and South Korea inferences from the visualizations. Analysis was described below:

- Fig. 8 showed that the confirmed cases in India raised exponentially with no fixed pattern.
- Fig. 9 provided the information that the confirmed cases in Italy raised exponentially with certain fixed pattern.
- Fig. 10 clearly illustrated that the confirmed cases in South Korea rose gradually.
- From Fig. 11 and Fig. 12 it was observed that there were only 3 confirmed cases in Wuhan in last week. The COVID-19 was almost controlled. The key reason behind the success of stopping the spread of this virus was that baidu's came up with thermal imaging cameras and it was deployed everywhere possible which helped in picking out people from streets whose temperature was high straight to the hospitals for treatment.

Also, the latest updates from Wuhan came on that day which made our inferences stronger as described below:

- China on Tuesday announced to lift the three-month lockdown on more than 56 million people living in the central province of Hubei.
- Bus services began in Wuhan for the first time since January 23 as a bus departed from its terminus at Hankou railway station at 5:25 am on Wednesday.
- The prolonged lockdown of Hubei's capital Wuhan will end on April 8, lifting the mass quarantine over the city with a population of over 11 million.



International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 102-110
Published Online December 2020 in IJEAST (http://www.ijeast.com)

Fig. 8. Confirmed cases in India.

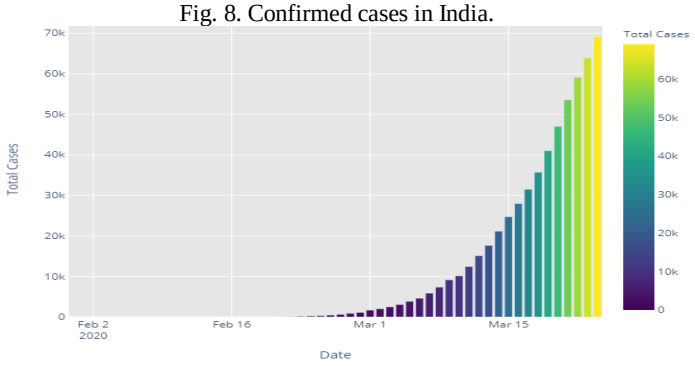


Fig. 9. Confirmed cases in Italy.

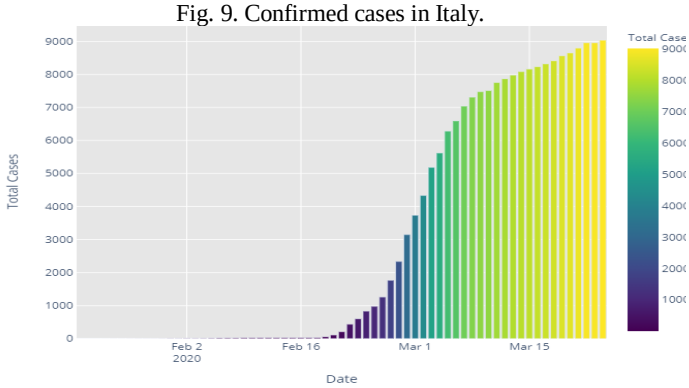


Fig. 10. Confirmed cases in South Korea.

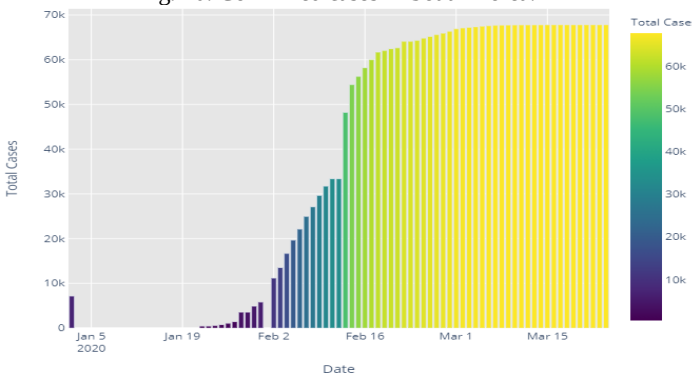


Fig. 11. Confirmed cases in Wuhan.

Fig. 13 shows the plotting of Italy, India and South Korea scenario. South Korea already started achieving the sigmoid pattern because of theory regressive testing they placed testing booths with proper safety to get people tested. Whereas India and Italy were still on a rise with patients. Wuhan as already close to achieved the sigmoid pattern was left out in this comparison. With the increase in number of cases the graph becomes bluer. The legend on right side of the graph helped in depicting the total cases in countries.

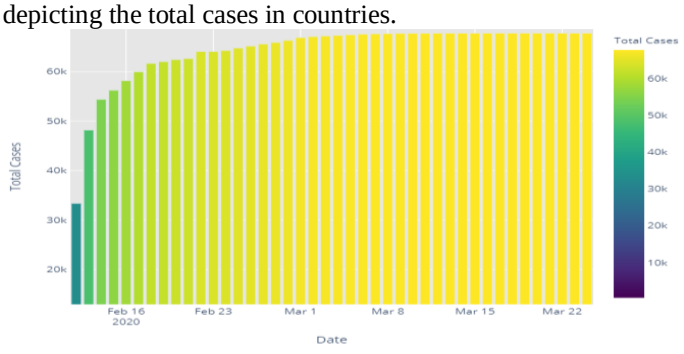


Fig. 12. Confirmed cases in Wuhan close to achieve sigmoid pattern.

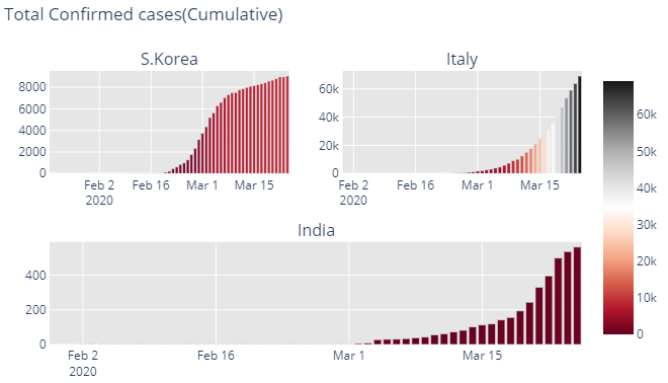


Fig. 13. Total confirmed cases (cumulative) for India, Italy and South Korea.

Fig. 14 represents the tread after crossing 100 days. It was analyzed that in Italy after crossing of the 30 days there were more than 70,000 people infected where on other hand South Korea did pretty well keeping the patients count below 10,000. The visualization inferences tell us the following:

- The number of days after the COVID-19 cases crosses 100 vs total number of cases in each country.
- Both Italy and South Korea crossed the mark of 5600 in the next 13 days.
- Number of cases detected (trend) in India was less as compared to Italy and South Korea.

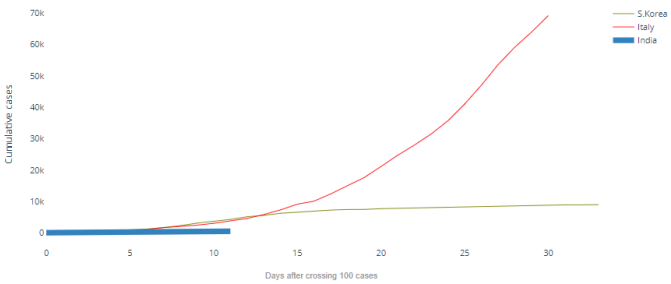


Fig. 14 Trend for India, Italy and South Korea after 100 cases.

The reason for India testing so little according to CNN Report [16] was due to the less spread of disease among people. As early "evidence" health authorities took 826 samples collected from patients suffering from acute respiratory disease from 50 government hospitals across India between 1st and 15th March which tested negative for coronavirus. Also, hospitals did not report any spike in admissions of respiratory distress cases. India was tested below scale due to the fear of its under-resourced and uneven public health system could be swamped by the patients. But according to the PM announcements it was clear that India was buying time to stock up on testing kits and add isolation and hospital beds [17][18].

Even India had 8 doctors per 10,0000 people as compared to 41 in Italy and 71 in Korea in Italy and 71 in Korea. There was one state-run hospital for more than 55,000 people. (Private hospitals were out of reach for most individuals). Another problem was that India had a poor culture of testing and most

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 97-101
Published Online December 2020 in IJEAST (http://www.ijeast.com)

of the people with flu symptoms avoid to consult doctors and instead try home remedies or go to pharmacies. There was a scarcity of isolation beds, trained nursing staff, medics, ventilators and intensive care beds.

C. Results of Worldwide Data Exploration

Fig. 15 showed the exploring of the worldwide data. There were 336.004k people which were confirmed of the deadly corona virus from which 98.334k recovered and 14.643k died. Whereas from Fig. 16 it was concluded that between February 12th 2020 and February 13th 2020 there was an increase of confirmed cases in just one day which was abnormal. From our study we found that on 12th February 2020 the main body responsible for researching about coronavirus came up with new method of counting patients which were infected with the virus but on 13th February they realized that the method was wrong and changed it. But till then the cumulative graph for the world scenario for confirmed patients was already altered because of that mistake.



Fig. 15. Worldwide NCOVID-19 cases for confirmed cases.

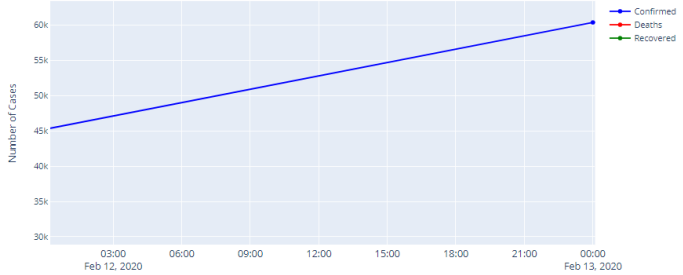


Fig. 16. Increase in confirmed cases between 12th and 13th February.

D. Results of Predicting Future Worldwide Cases using Prophet

The predicted value for confirmed cases on 26th March 2020 came out to be 3,55,136 and for 30th March it came out to be 4,18,529. All predicted values for confirmed cases are mentioned in Table II. The forecast for confirmed cases is presented in Fig. 17. In Fig. 18, the prediction components for confirmed cases illustrated the number of confirmed patients will go down near Wednesday and then again raise near Saturday of the week. The huge dip in number of confirmed cases near Wednesday was because in china at that particular day very a smaller number of people were found COVID-19 positive. China being the huge contributor in populating the affected crowd made the forecast like this. The forecast for death cases is presented in Fig. 19. The predicted value for death cases on 26th March 2020 came out to be 15,411 and for 30th March it came out to be 18,490. All predicted values for death cases are mentioned in Table III. The death cases also came down near Wednesday because the number of confirmed cases were predicted to be less at that day which showed the consistency of the graphs as shown in Fig. 20. The forecast for recovered cases was presented in Fig. 21 which showed the most coinciding graph even in exponential increase of predicted and actual cases there was an amazing cohesiveness. The predicted value for recovered cases on 26th March 2020 came out to be 1,03,001 and for 30th March it came out to be 1,13,244. All predicted values for death cases are mentioned in Table 4. The recovered components are plotted in Fig. 22. From the subplots the recovered cases were seen more in Sundays and Mondays because the most number of identified cases according to model were also on Sundays and Mondays. It takes a hit till Friday because of less case and rises on Saturday again. The predicted values for confirmed, death and recovered patients was mentioned in Table 2– 4 respectively.

Table-2 Predicted Values for Confirmed Cases

Date	Predicted value	Lower limit	Upper limit
2020-03-26	355136.872975	334384.109371	375726.515251
2020-03-27	372235.326939	349712.320279	391835.450565
2020-03-28	388674.964143	367308.789183	411177.125876
2020-03-29	405307.954676	381419.927783	428149.690375
2020-03-30	418529.648467	395683.667642	443064.760034

Table-3 Predicted Values for Death Cases

Date	Predicted value	Lower limit	Upper limit
2020-03-26	15411.877522	14483.880532	16355.741169
2020-03-27	16199.617603	15223.621567	17106.213775
2020-03-28	17006.904589	16017.036648	17958.214719
2020-03-29	17831.505383	16831.699633	18852.858632
2020-03-30	18490.986176	17455.813093	19549.969302

Table-4 Predicted Values for Recovered Cases

Date	Predicted value	Lower limit	Upper limit
2020-03-26	103001.351979	101349.595952	104789.401191
2020-03-27	105338.025206	103466.650449	107416.338645
2020-03-28	108256.310913	106204.677291	110231.644632
2020-03-29	111119.296134	108721.510751	113799.330776
2020-03-30	113244.961701	110536.366609	116213.952425

If we look at the world's data and its forecast, it is clear to say that India might face one of its worst days if strict measures are not taken against COVID-19. If we look at the data, the number of cases in India is rising but not like in Italy, Wuhan and South Korea. From proposed research work we can say that the number was less because of low awareness and less tests being conducted in India. Currently India is a deadly and risky zone as there are very few COVID-19 test centers in India as compared to its population.

PM Modi announced Janta Curfew in India on 22nd March. From 24th March there is a complete lockdown the entire country. Definitely it was for our own good.

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 102-110
Published Online December 2020 in IJEAST (http://www.ijeast.com)

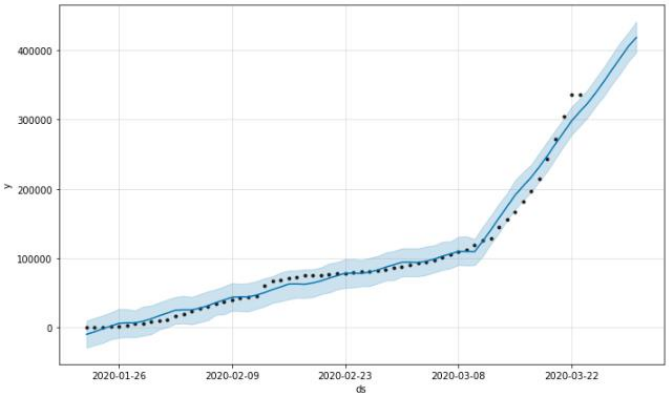


Fig.17. Plotting forecast for confirmed cases.

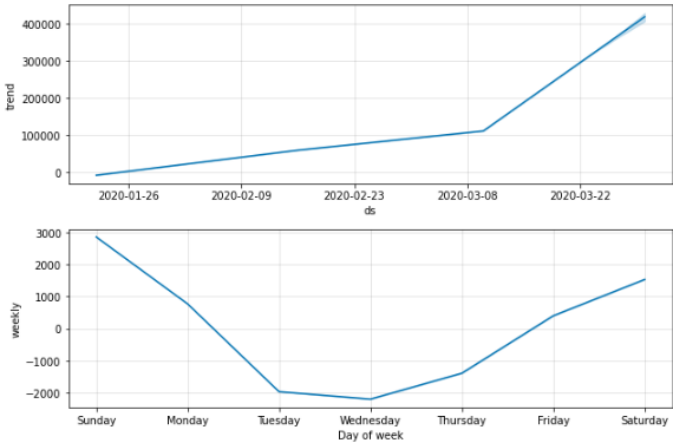


Fig. 18. Displaying forecast components for confirmed cases.

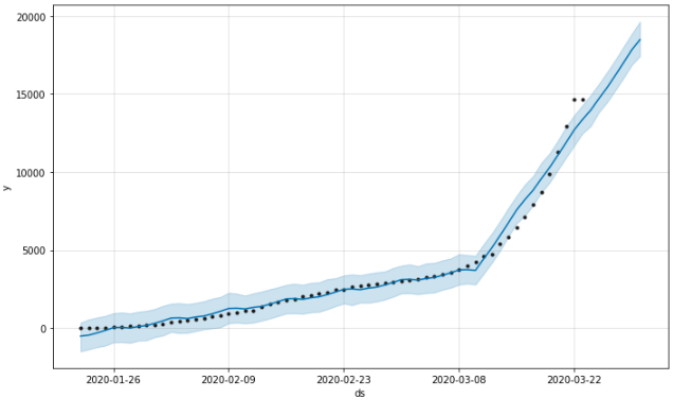


Fig. 19. Plotting forecast for death cases.

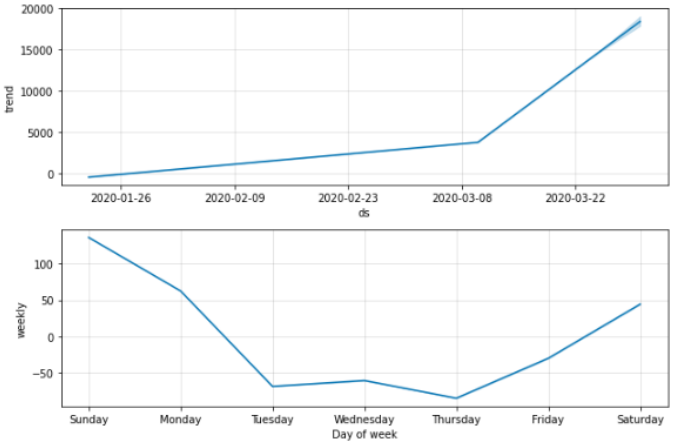


Fig. 20. Displaying forecast components for death cases.

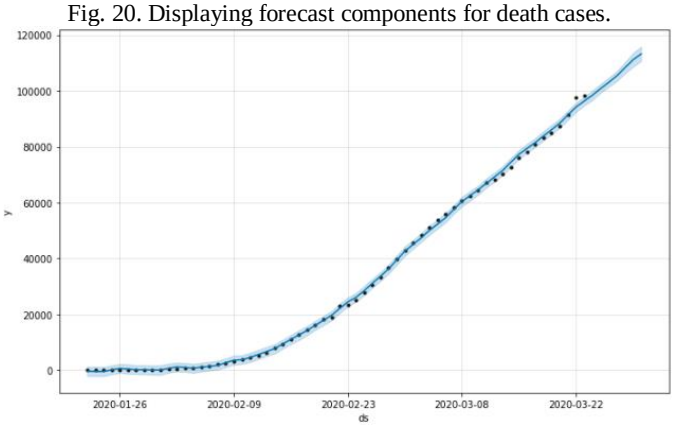


Fig. 21 Plotting forecast for recovered cases.

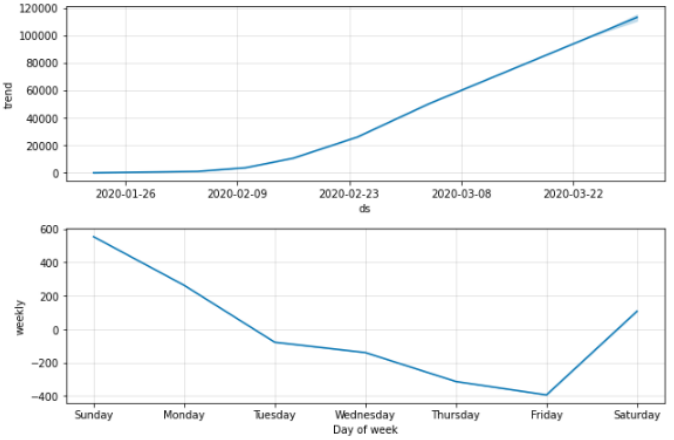


Fig. 22. Displaying forecast components for recovered cases.

IV. CONCLUSIONS

For proper visualization and to draw accurate inferences it is proved that the proposed model fits the best. The analysis of present condition in India can be performed for any country based on datasets provided of that country. From finding total cases considering the foreign and country national people to finding active cases. Visualizations on maps is the best way to recognize the disease spreads in a country in the results. Covering every state comparison for cases to determine the performance of that state to finding the pattern in raising cases for a particular country is achieved in the results. A global comparison is necessary when we talk about pandemics which



is achieved though the results finding similar trends to other countries followed by exploring the data worldwide. In case of stopping the disease from spreading a proper technique can only be implemented and thought of when we know the coming forecast due to that particular disease which is also achieved in the results. At last only suggestions to India for coming days is to implement regressive testing, use thermal imaging cameras on public places and impose heavy fines on those who break laws which was done by China and other countries.

This model can be used by any country or government organization to non-government organizations who want to know the impact of a virus in worldwide scenario and prepare before to combat it. Results obtained by the model can be used as a reference in finding out the mistakes on country as well as global level.

Future work will be directed towards the analyzation on large worldwide datasets and then performing comparison of other countries with India using machine learning models. Further comparing with time series models.

V. REFERENCES

[1] Organization W. H. (2020). Naming the Coronavirus Disease (COVID-19) and the Virus that Causes it, World Health Organization.

[2] World Health Organization (WHO) (2020), Novel Coronavirus—China.

[3] Huang C., Wang Y., Li X., Ren I., Zhao J., and Hu Y. et al (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *The Lancet*, vol. 395, no. 10223, (pp. 497 - 506).

[4] Organization W. H. (2020). Statement on the second meeting of the International Health Regulations (2005), Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV).

[5] Organization W.H. (2020). WHO Director-General's opening remarks at the media briefing on COVID-19 .

[6] Apolloni A., Poletto C., and Colizza V. (2013), Age-specific contacts and travel patterns in the spatial spread of 2009 H1N1 influenza pandemic, BMC *Infectious Diseases, vol.* 13, no. 1, (pp.176).

[7] Eubank S., Guclu H., Kumar V. A., Madhav A. S., Marathe, Z., Toroczkaiand Z., and Wang N. (2004), Modelling disease outbreaks in realistic urban social networks, *Nature*, vol. 429, no. 6988, (pp. 180 –184).

[8] Griffin R. H. and Nunn C. L. (2012), Community structure and the spread of infectious disease in primate social networks, *Evolutionary Ecology*, vol. 26, no. 4, (pp.779–800).

[9] Repaka A. N., Ravikanti S. D., and Franklin R. G. (2019), Design and implementing heart disease prediction using naives bayesian, in *Proc. of 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, (pp.292-297), Tirunelveli, India.

[10] Stojanovic O., Leugering J., Pipa G., Ghozzi S., and Ullrich A. (2019), A bayesian monte carlo approach for predicting the spread of infectious diseases, PloS One, vol. 14, no. 12.

[11] Walters C. E., Mesle M. M. and Hall I. M. (2018), Modelling the global spread of diseases: A review of current practice and capability, *Epidemics*, vol. 25, pp.1-8, Dec.

[12] Arun S. S. and Iyer G. N. (2020), On the Analysis of COVID19-Novel Corona Viral Disease Pandemic Spread Data Using Machine Learning Techniques, in *Proc. of 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, (pp.1222-1227), Madurai, India.

[13] Song X., Xiao J., Deng J., Kang Q., Zhang Y, and Xu J. (2016), Time series analysis of influenza incidence in Chinese provinces from 2004 to 2011, *Medicine*, vol. 95, no. 26.

[14] Lee K., Agrawal A., and Choudhary A. (2017), Forecasting influenza levels using real-time social media streams, in *Proc. of IEEE International Conference on Healthcare Informatics (ICHI)*, Park City, UT, USA , (pp.409-414).

[15] Johns Hopkins University Data Repository. Cssegisanddata. Accessed: Mar. 25, 2020. [Online]. Available: https://github.com/ CSSEGISandData.

[16] Ministry of Health and Family Welfare, Statewise Status. Accessed: Mar. 25, 2020. [Online]. Available: https://www.mohfw.gov.in/

[17] Biswas S. (2020), Coronavirus: Why is india testing so little, *BBC News*, Mar.

[18] New A. J. (2020), India's poor testing rate may have masked coronavirus cases.

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 8, ISSN No. 2455-2143, Pages 102-110
Published Online December 2020 in IJEAST (http://www.ijeast.com)

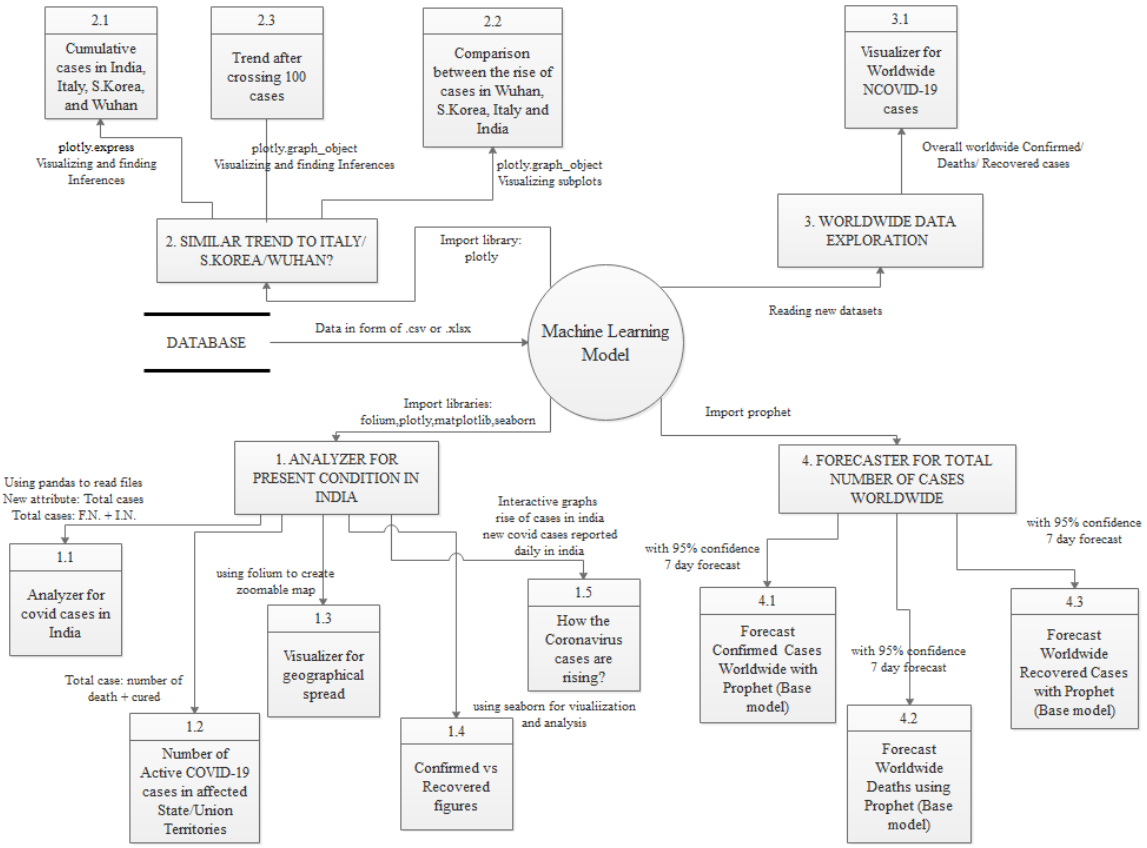


Fig. 1. Data flow diagram of machine learning model.

IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE AND TECHNOLOGY

AUTHORS

Kanika Bhalla

Ashish Kumar

ABOUT IJEAST

International journal of Engineering Applied Science and Technology (IJEAST) is a peer-reviewed, open access journal that publishes high - quality research paper in the field of Engineering, Applied Science and Technology.

IJEAST aims to provide a platform for researchers, academicians, and professionals to share their innovative ideas, research findings, and practical experiences with the global scientific community.

JOURNAL METRICS

H-INDEX	i10-INDEX	IJEAST CITATIONS	PEER-REVIEWED & OPEN ACCESS
33	154	9,647 (IJEAST CITATIONS)	PUBLISHED MONTHLY

PEER REVIEWED

All submission are rigorously peer reviewed to ensure quality.

AUTHORS ARE OUR PRIORITY

We value our authors and recognize their contribution to the advancement of research and knowledge.

OPEN ACCESS

Free and unrestricted access to research for all.

GLOBAL REACH

Connecting researchers and professionals worldwide.

TIMELY PUBLICATION

We ensure a swift and efficient publication process.

For more information, visit our website
www.ijeast.com

editor@ijeast.com

www.ijeast.com

India

2455-2143