

Covid-19 Cases prediction

Meet Pedhadiya

Information and Communication Technology
School of Engineering and Applied Science, Ahmedabad University
Ahmedabad, India
email: meet.p2@ahduni.edu.in

Krunal Pagdar

Information and Communication Technology
School of Engineering and Applied Science, Ahmedabad University
Ahmedabad, India
email: krunal.p@ahduni.edu.in

Ayush Kaneria

Information and Communication Technology
School of Engineering and Applied Science, Ahmedabad University
Ahmedabad, India
email: ayush.k1@ahduni.edu.in

Mihir Chauhan

Information and Communication Technology
School of Engineering and Applied Science, Ahmedabad University
Ahmedabad, India
email: mihir.c@ahduni.edu.in

Abstract— Covid-19 diseases spread all around world for a more than 1 year and WHO declares COVID-19 as global pandemic. Today number of covid-19 cases exceeded 100 Million+ across the globe and deaths due to Covid-19 exceeded 2.5M + across the globe. In this pandemic situation COVID-19 case prediction helps a lot to the medical management to distribute medical recourses and take care of all the precautions. In this paper we have described our approach and work done for the implementing machine learning (ML) algorithm to predicts the COVID-19 cases in India.

Keywords—Covid-19, Pandemic, Prediction, Daily confirmed cases, Active Cases, Daily Tests, Algorithms

I. INTRODUCTION

The SARS-CoV-2 is the cause of Covid-19 diseases which originated in the city Wuhan, China at the end of the year 2019. In the beginning, it affects Wuhan very badly. Gradually it spreads all around the world. It has been a big threat to global health as WHO (World Health Organization) declared COVID-19 diseases as a global pandemic in March 2020. At the end of June 2020 number of COVID-19 cases are exceeded 10 Million+ across the globe. India reported the first case of COVID-19 on 30th January 2020. During that time international traveller, Indian students who studied abroad came back to India, etc. results in an increase in cases of COVID-19 in India. Today India has crossed 10 Million + cases of COVID-19. In a rapidly evolving pandemic, it is important to have proper analysis and prediction of cases in the future. Inefficient distribution of medical resources results in improper medical facility and it affects the recovery rate. If we have a good model that predicts the new cases well, it will result in the efficient distribution of medical resources in arising cases in covid-19.

We found an efficient dataset that contains daily positive cases and deaths due to the COVID-19 disease of India. The dataset contains daily cases, deaths, tested reports, etc. since the first registered case in India. Dataset also contains State-wise data [3].

We intend to apply a machine learning (ML) algorithm to the dataset and predict the cases of COVID-19 cases.

II. LITERATURE SURVEY

This pandemic has had a major impact on the health and well-being of the global population. Various research works have been done to forecast the cases of COVID-19.

A.Khakharia et al.[1] have developed an outbreak prediction system for COVID-19. For that, they analyze data of the top 10 highly dense and populated countries (India, Bangladesh, the Democratic Republic of Congo, Pakistan, China, Philippines, Germany, Indonesia, Ethiopia, and Nigeria). In the paper, the authors proposed a prediction model to forecast the count of new cases likely to arise in 5 successive days. For the prediction, the author has used 9 different machine learning (ML) algorithms. The 9 different algorithms were Auto-Regressive Moving Average (ARMA), Auto-Regressive Integrated Moving Average (ARIMA), Support Vector Regressor (SVR), Linear Regressor polynomial (LRP), Bayesian Ridge Regression (BRR), Linear Regression (LR), Random Forest Regressor (RFR), Holt-Winter Exponential Smoothing (HW), and Extreme Gradient Boost Regressor (XGB).

The author found that different country has different behaviour of increasing and decreasing in COVID-19 cases. So that not every algorithm could give high accuracy of prediction of cases for each country. They implemented all 9 algorithms to different 10 countries' data. They found that different countries have different accuracy with different models implemented. Some counties have good accuracy with one algorithm while other countries may not have good accuracy with that particular algorithm.

Y. Zoabi [2] has implemented a model that predicts COVID-19 diagnosis based on symptoms. They have established a machine learning approach that trained on records from 51,831 tested individuals. The model is using 8 binary features i.e., sex, age>60 years, known contact with an infected individual, and the appearance of five initial clinical symptoms.

III. IMPLEMENTATION

Before you begin to format your paper, first write and save

There are many sources available on the internet that provides a COVID-19 datasheet. Some of the datasets have a large number of errors or very less amount of data. We found a data set from the website “covid19india.org” which has the most appropriate dataset. These datasets contain different kinds of information like all India, state-wise, district-wise, and with their daily information and time-series and in long formation about COVID-19 cases. We have done analytics on the datasets and compare them.

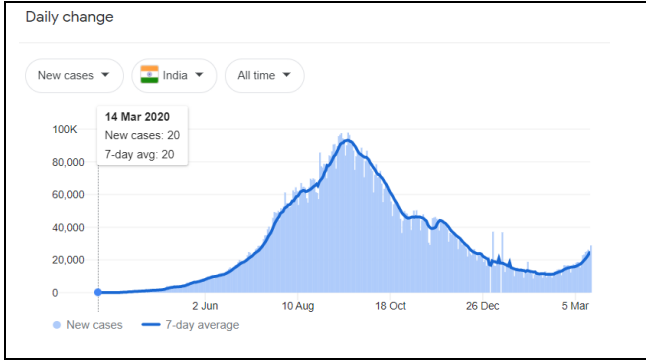


Figure (a)

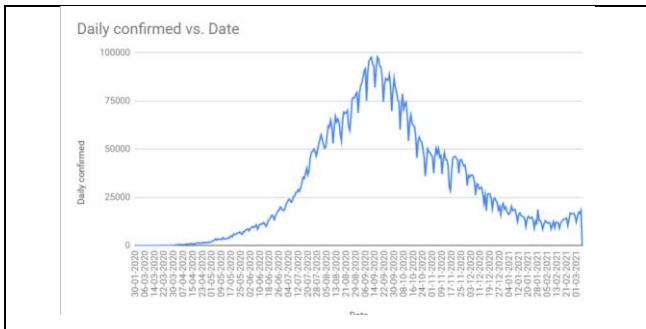


Figure (b)

We have done some data analysis on the datasets and one of the comparisons is shown in figure(a) and figure(b).

Figure (a) represent data of COVID-19 daily cases in India from our selected datasets and figure (b) represents graphs provided by John Hopkins University.

We had compared data analysis for all states wise data as well. Data is almost the same in all cases.

IV. RESULT AND CONCLUSION

We have got a datasheet and now we will implement a machine learning (ML) algorithm to predicts the COVID-19 cases based on given data.

REFERENCES

Article's reference:

- [1] A. Khakharia et al., “Outbreak Prediction of COVID-19 for Dense and Populated Countries Using Machine Learning,” *Ann. Data Sci.*, vol. 8, no. 1, pp. 1–19, 2021, doi: 10.1007/s40745-020-00314-9.
- [2] Y. Zoabi, S. Deri-Rozov, and N. Shomron, “Machine learning-based prediction of COVID-19 diagnosis based on symptoms,” *npj Digit. Med.*, vol. 4, no. 1, pp. 1–5, 2021, doi: 10.1038/s41746-020-00372-6.

Dataset reference:

- [3] <https://www.covid19india.org/>
- [4] <https://github.com/nshomron/covidpred>