Mini Project -I Report
on
Mini Project Title

by

Ayush Bhandarkar
UID : 2020400001

Anmol Chokshi
UDI : 2020400003

Adish Padalia
UID : 2020400031

Guide Name:
Prof. Swapnali Kurhade

Information Technology Department

Bharatiya Vidya Bhavan's

Sardar Patel Institute of Technology

Munshi Nagar, Andheri(W), Mumbai-400058

University of Mumbai

April 2022

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Ayush Bhandarkar (2020400001)

Anmol Chokshi (2020400003)

Adish Padalia (2020400031)

Date: 29th April

Acknowledgements

We feel great pleasure in presenting the stage one report of our mini project titled ‘Covid Prediction for next 30 days’. We have channelized our best efforts towards a systematic approach to the project, keeping in mind the aim we need to achieve.

We are highly grateful to our project guide Prof. Swapnali kurhade, Department of Information Technology, Sardar Patel Institute of Technology (SPIT) for constant encouragement, effort and guidance. She has always been involved in discussing our topic at each phase to make sure that our approach was designed and carried out in an appropriate manner and that our conclusions were appropriate, given our results.

Ayush Bhandarkar

Anmol Chokshi

Adish Padalia

Abstract

In most developing countries, most data is still captured, organized and maintained using paper forms. This is also true for businesses. Forms are an important part of most transactions. However, it can be challenging to aggregate, share, and analyze the data collected using paper forms. Better management and processing of forms and applications is indispensable to improving customer experiences. But, typing those forms into a spreadsheet is time-consuming, mundane and dull. Also, there can be data discrepancies and missing values that need to be dealt with. This paper presents a system that uses computer vision to capture data from paper forms. It analyzes business needs and generates templates that mirror all the requirements .The extracted data is then used to reduce labour by auto-completing the relevant and required fields that the template asks for. Also, the extracted data is stored in a structured format, exportable as a CSV file or into Google documents. Thus, the system considers not just conversion of information from paper to electronic format, but also identifies opportunities to eliminate paper as the vehicle for providing information at the source of origination.

Table of Contents

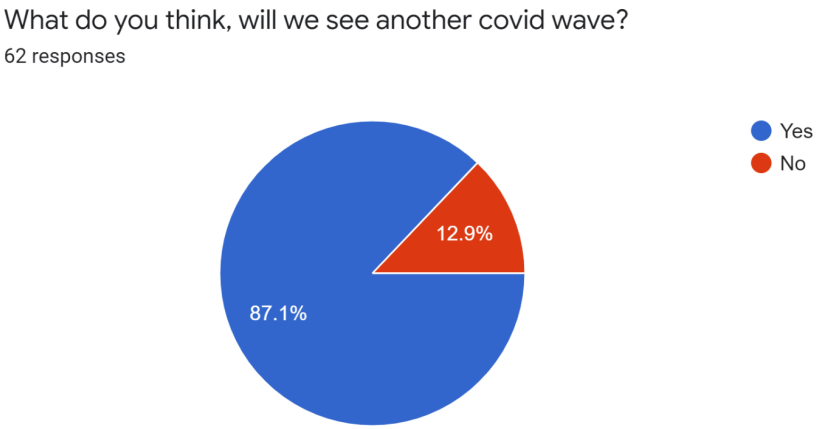
1. Introduction

1.1. Problem Statement :

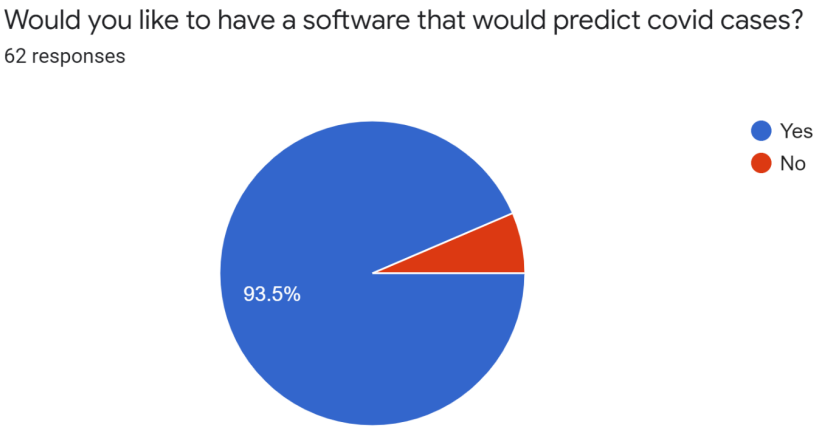
- Coronavirus disease (COVID-19) is an inflammation disease from a new virus. The disease causes respiratory ailment with manifestations, and in progressively serious cases, the problem in breathing and death. Coronavirus has been perceived as a worldwide pandemic
- Initially we used to be unsure about how the cases will rise or fall in the upcoming days. Thus we could not prepare for the same and lost a lot of time and resources.
- Using this software, we will be able to predict the covid cases for the next 30 days. This will ensure we can make adequate preparations for the same.

1.2. Literature Survey/Market Survey :

- We had floated a survey to ask people about their opinion on another Coronavirus wave and their view on whether a software that could predict covid cases would be useful to them.
- The response we got was as follows:



(Figure-1)



(Figure-2)

- It is clear from this Market survey that the need for a prediction website for covid cases is important and useful for all users.

1.3. Scope and Objectives :

- Predicting covid cases will ensure that we adequately prepare for them and take the required provisions before time.
- To help to better understand the current situation.
- By analyzing the area wide spread we can ensure that the resources like beds and oxygen reach places where they are required
- We will use the Facebook prophet model for the task of Covid-19 cases prediction for the next 30 days. The Facebook prophet model uses a time series method for forecasting. We will show a graph of expected cases.
- We will show the user the distribution of covid cases, deaths, recoveries and active cases around the world.
- At the end we will predict the covid cases for the next 30 days using previous data and visualise it using a graph.This would be done as per the state or country entered by the user in India's page or world's page respectively.

1.4. Assumptions :

- Boundaries : Covid has been around only for 2 years so the data is available only for that time.
- Assumptions : We are assuming that the data we are taking is accurate and there is no change in government policies regarding covid.

1.5. Constraints :

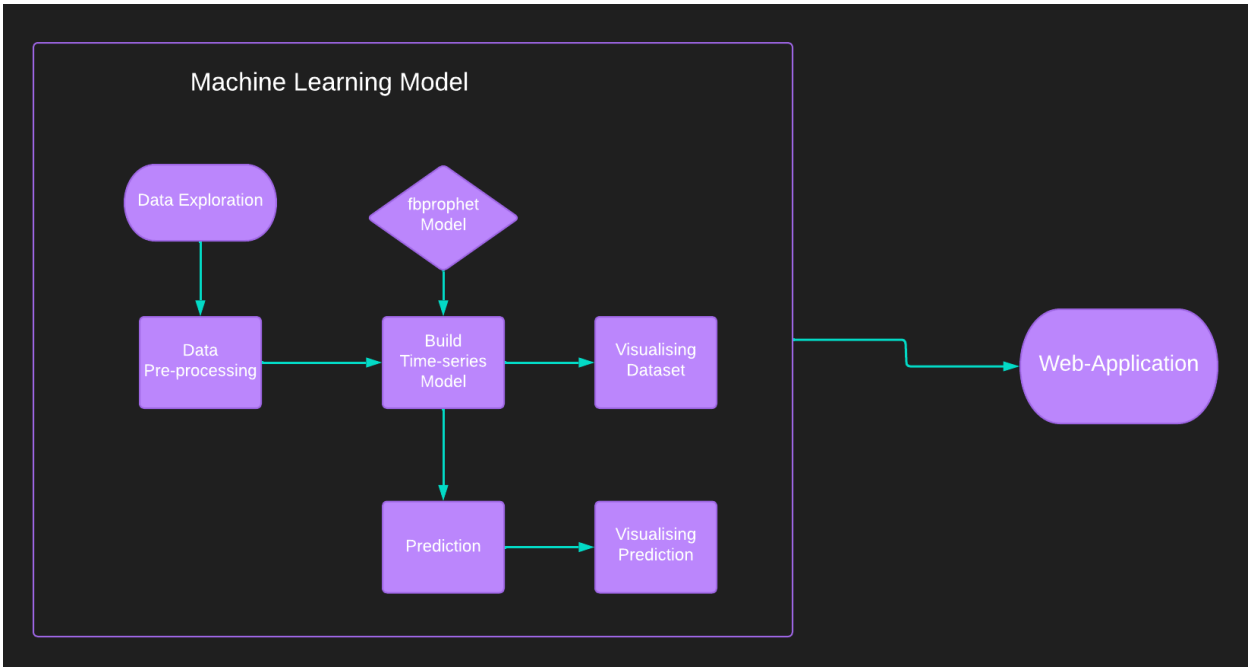
- Work constraints : Some of the data might be missing so we have assumed the missing data to be the average of all the entries present in the dataset. The data for all the states in India is available till August 2021.

2. Proposed System

2.1. Architecture Diagram

The system consists of two major components namely:

1. Machine learning model
2. Web application



(Figure-3)

2.1.1. **Data exploration:** Initially we found 2 separate datasets, one for the world which would be country-wise and would include covid-19 cases, deaths and recoveries day-wise and another for India which would be state-wise which would also include covid-19 cases day-wise.

2.1.2. **Data Pre-processing:** Modifies the dataset in a way such that it can be suitable for the machine learning model

2.1.3. **Build time series model:** We build the fbprophet model to predict the covid-19 cases for the next 30 days. We choose this model as it gives us 92.47% accuracy

2.1.4. **Data-visualisation:** Initially we have visualised the data set in various representations such as bar graph, simple line graph etc for the covid-19 cases, deaths and recoveries for the world country-wise and India state-wise.Then we have visualised a graph in which we could observe the prediction and the actual dataset we have collected.

2.1.5. **Web-application:** We have shown the Machine learning model results by building a user friendly website.

2.2. Algorithms used :

(Figure-4)

- Facebook Prophet model is based on a time series model.
- It helps us better analyse and visualise the data.
- It makes accurate future predictions by taking into account the previous data of date and parameter for which predictions are to be made as an input.
- It has easy implementation and also gives us weekly,monthly and yearly trends.

Facebook Prophet operates similarly to scikit-learn, so first we instantiate the model, then call .fit(ts) passing the time series through it. When calling .predict(ts), the Prophet outputs a lot of information. Luckily, the developers added a method called .make_future_dataframe(periods = 10) that will easily collect all of the output in an organized way. This method outputs an empty pandas dataframe that we will fill with the forecast using the .predict(ts)method. The forecast will contain a prediction for every historical value present in the dataset plus additional forecasts for the number of periods passed through the method (in the case above 10). There are many columns of useful information in this future dataframe but the most important ones are:

- ds contains the timestamp entry of the forecast
- yhat contains the forecasted value of the time series
- yhat_lower contains the bottom of the confidence interval for the forecast
- yhat_upper contains the bottom of the confidence interval for the forecast

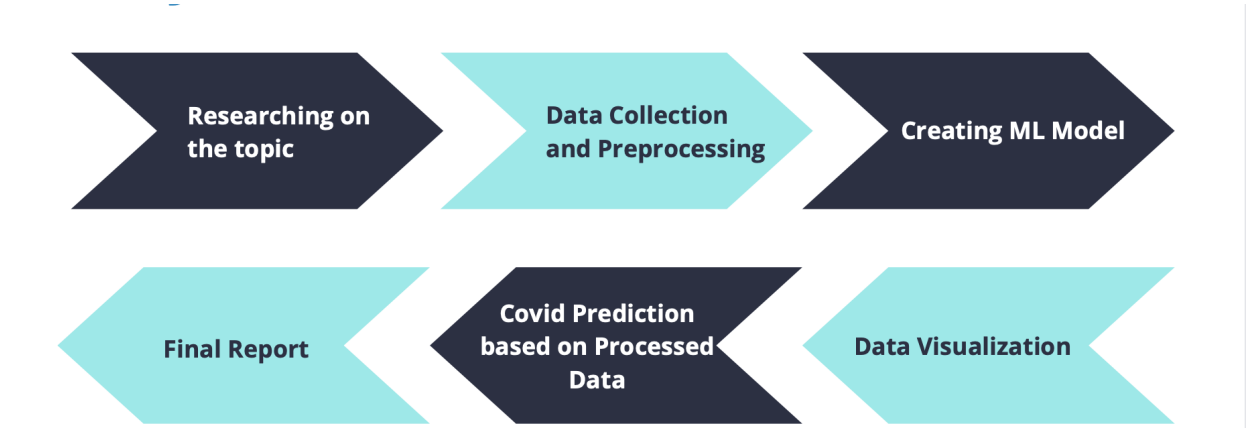
A .plot() function is also provided for easy plotting of the original data, the forecast and the confidence interval of the model.

We have divided the machine learning model into various steps and we are calling the functions as and when required.

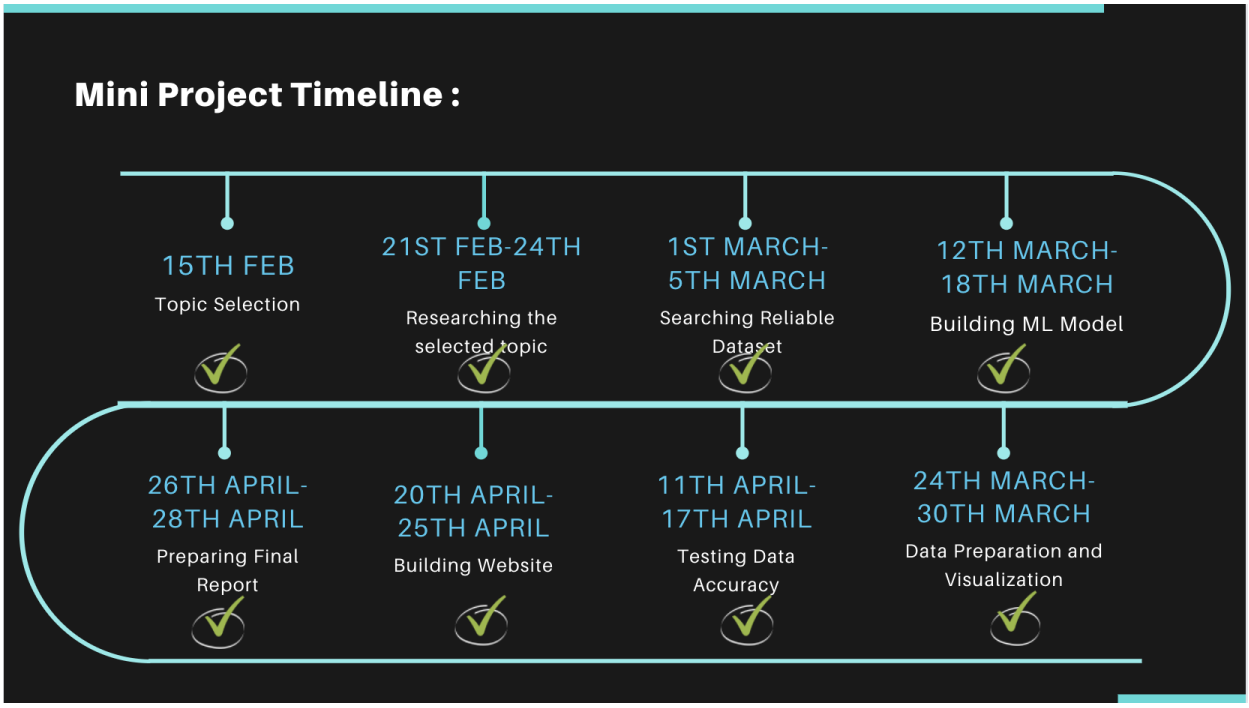
- We have created several plotting functions using plotly and matplotlib such as bar graph,line graph,geographical representation.
- We have created a function which would take in the state or the country and show the predicted graph accordingly.

3. Project Plan

Modules of the Project :



(Figure-5)



(Figure-6)

4. Implementation :

We have created the following major web pages in our website :

Home Page:

- This is the page our website will open to. It consists of a navbar common to all pages with links to various pages in our website.
- It also displays the latest figures regarding covid cases in India.
- Further it contains 2 cards to take users to the world and india page along with a short description for each.

World Page:

- In this page, we have first depicted the live count of covid cases,deaths, recoveries and active cases and following that we have represented the country-wise distribution of covid-cases using a world-map.
- Next, we have shown the top 10 worst affected countries due to the covid-19 pandemic using different kinds of graphs.
- We have also shown the graph for recoveries.
- At the end, we have created a form which would ask users about the country and predict the covid cases for the next 30 days using graphical representation.

Country Page:

- This page takes the input from the user in the form mentioned in the above page.
- This page shows the users a graphical representation of cases along with a prediction of the number of cases for the next 30 days, for the country they entered.
- If the user enters an invalid country they are redirected to the same page.

India Page:

- In this page, we have represented the state-wise distribution of covid-cases using India's map.
- We have also shown the graph for cured cases and active cases.

- At the end, we have created a form which would ask users about the state and predict the covid cases for the next 30 days using graphical representation.

State Page:

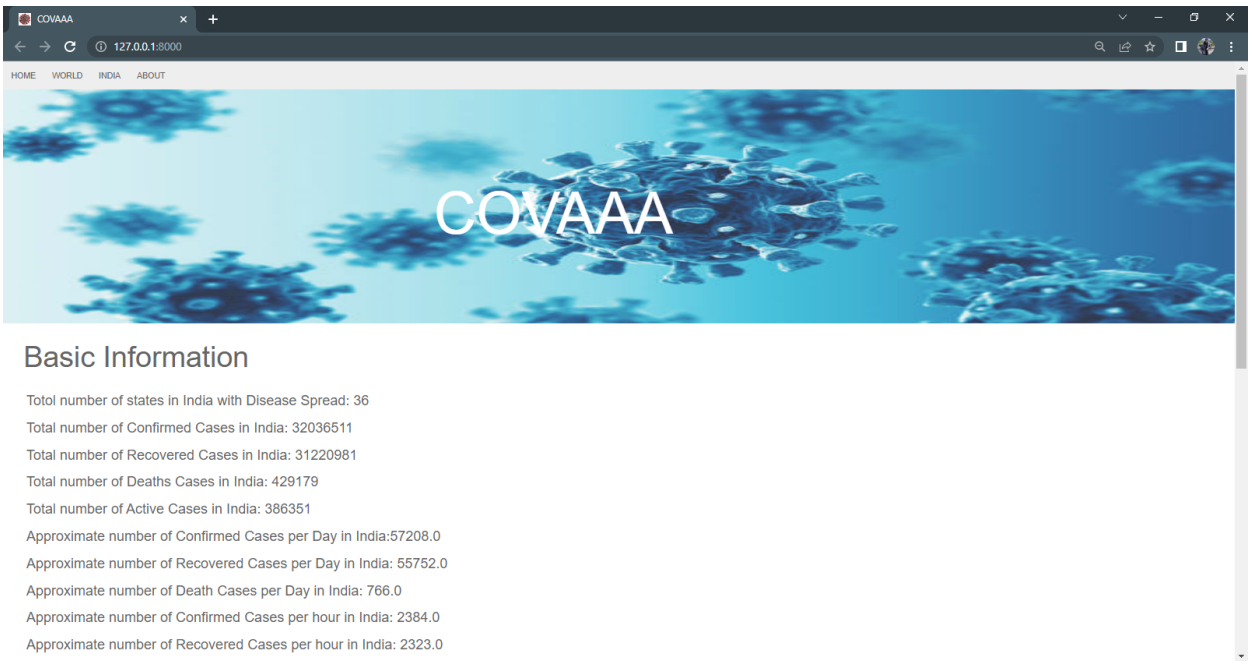
- This page takes the input from the user in the form mentioned in the above page.
- This page shows the users a graphical representation of cases along with a prediction of the number of cases for the next 30 days, for the state of India they entered.
- If the user enters an invalid state they are redirected back to the same page.

Tech Stack used :

1. HTML
2. CSS
3. Pandas
4. Numpy
5. FBprophet
6. Django

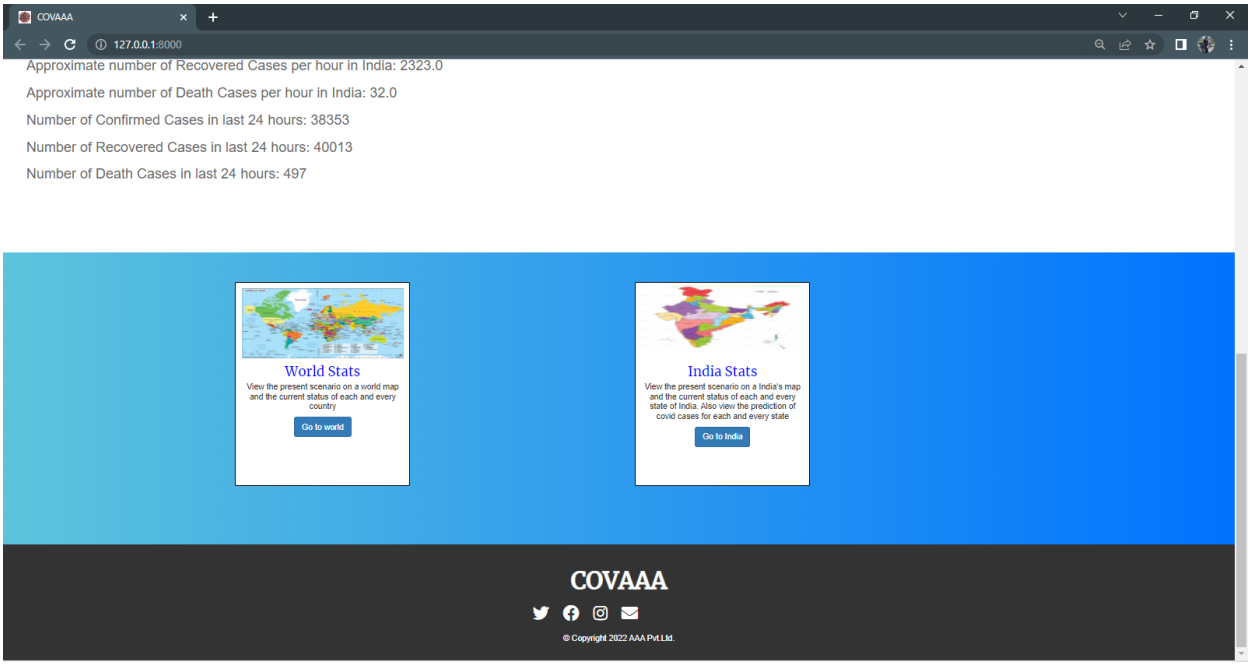
Our Website

1. Home Page



(Figure-7)

On the home page, we display the daily updates and statistics for India's Covid cases.



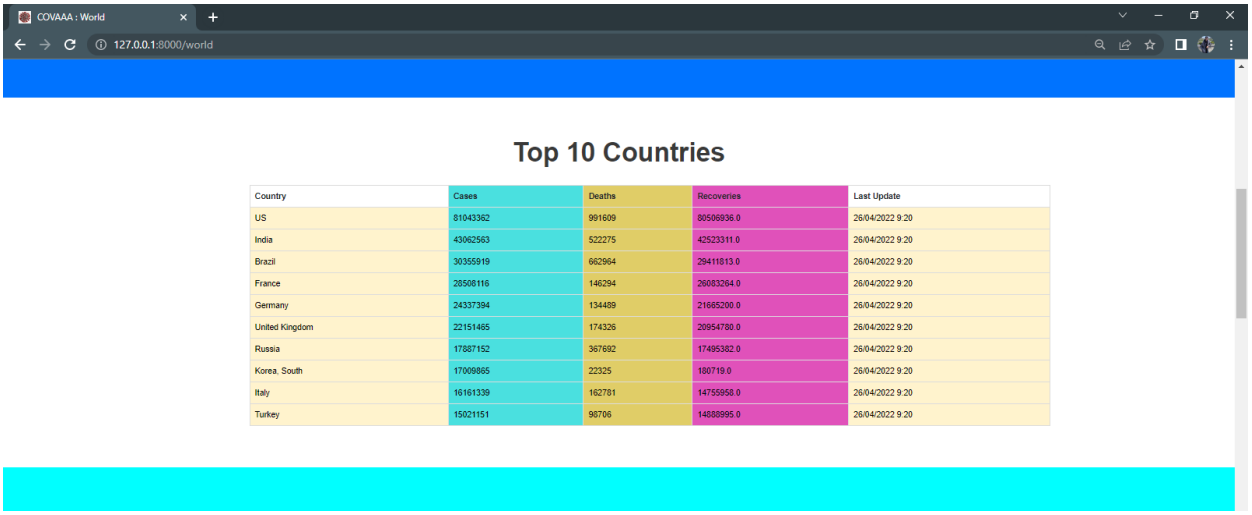
(Figure-8)

Here, you can go to 2 pages → India and World pages

2. World Page

(Figure-9)

This is the world page, here we display the total number of different cases → Confirmed, Death, Recovered and Active cases.

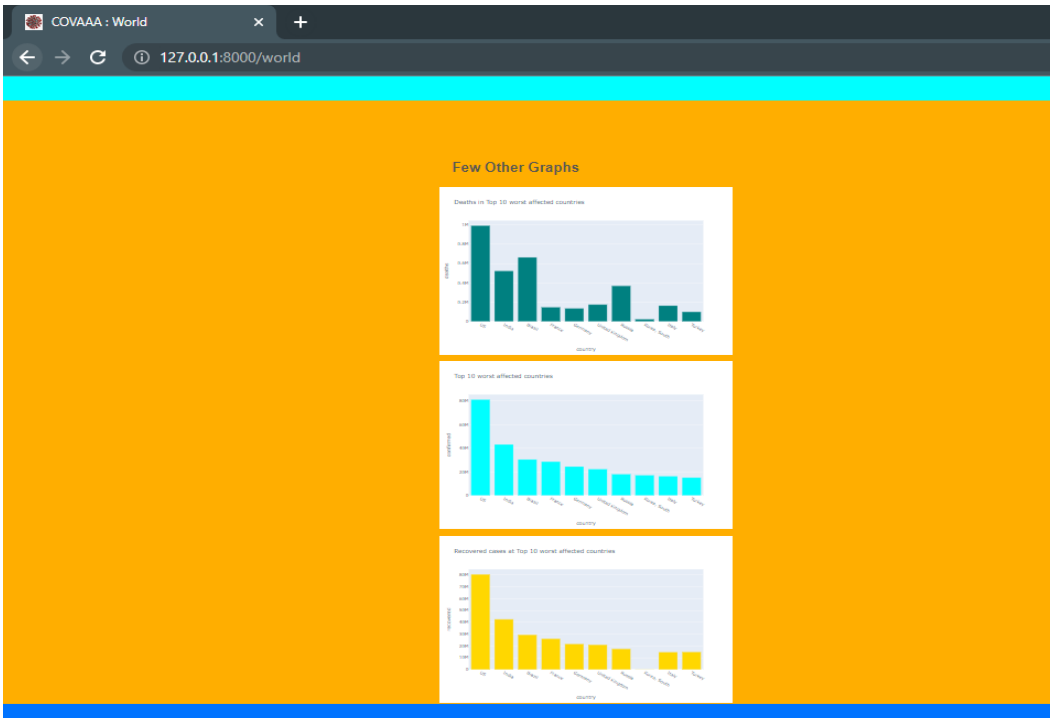


(Figure-10)

This is in the world page, here we display the top 10 affected countries in table format.

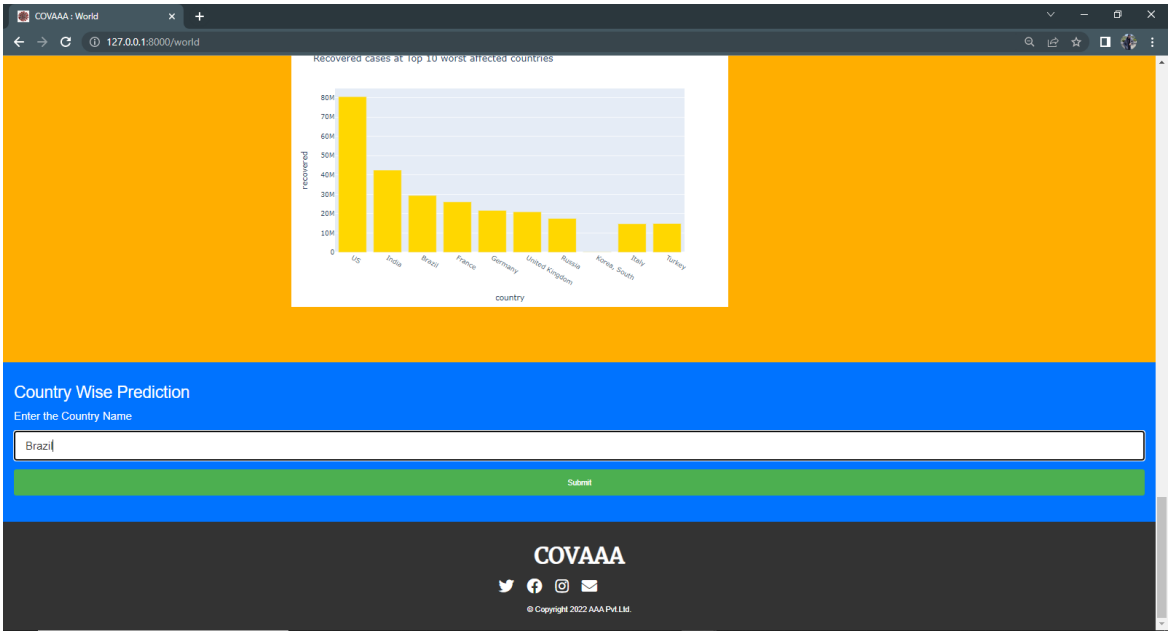
(Figure-11)

This is in the world page, here we display the top 10 affected countries in Bar graph format.



(Figure-12)

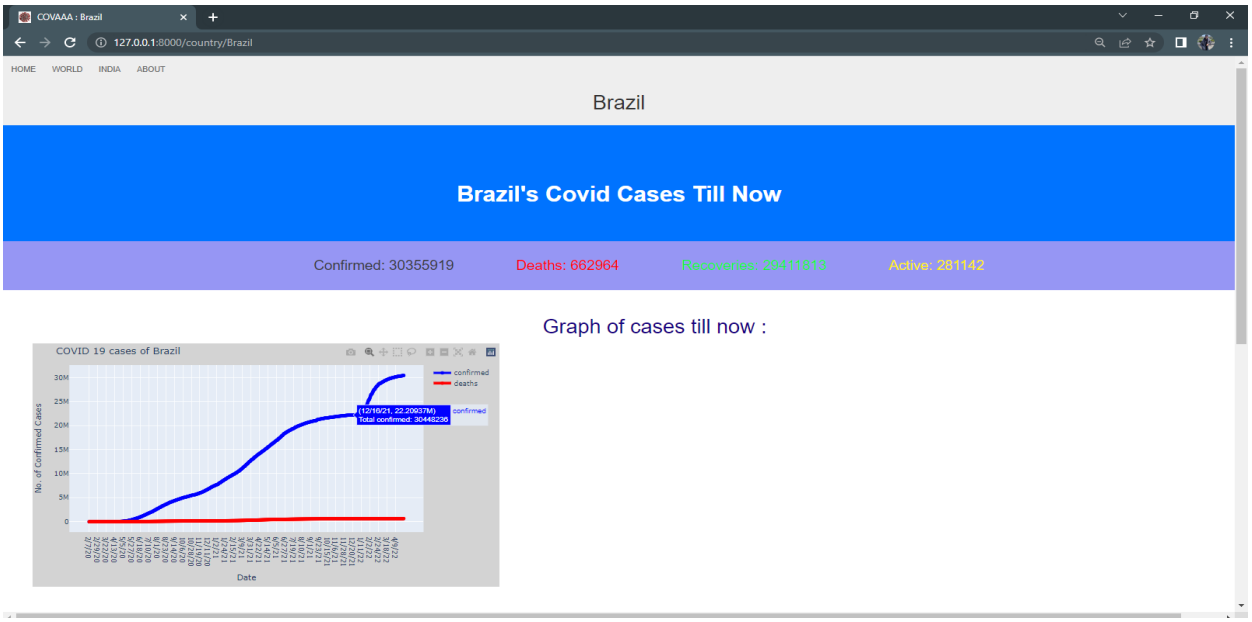
This is the World page, here we display the confirmed, recovered and death cases graphs for the top 10 countries in a Bar graph format.



(Figure-13)

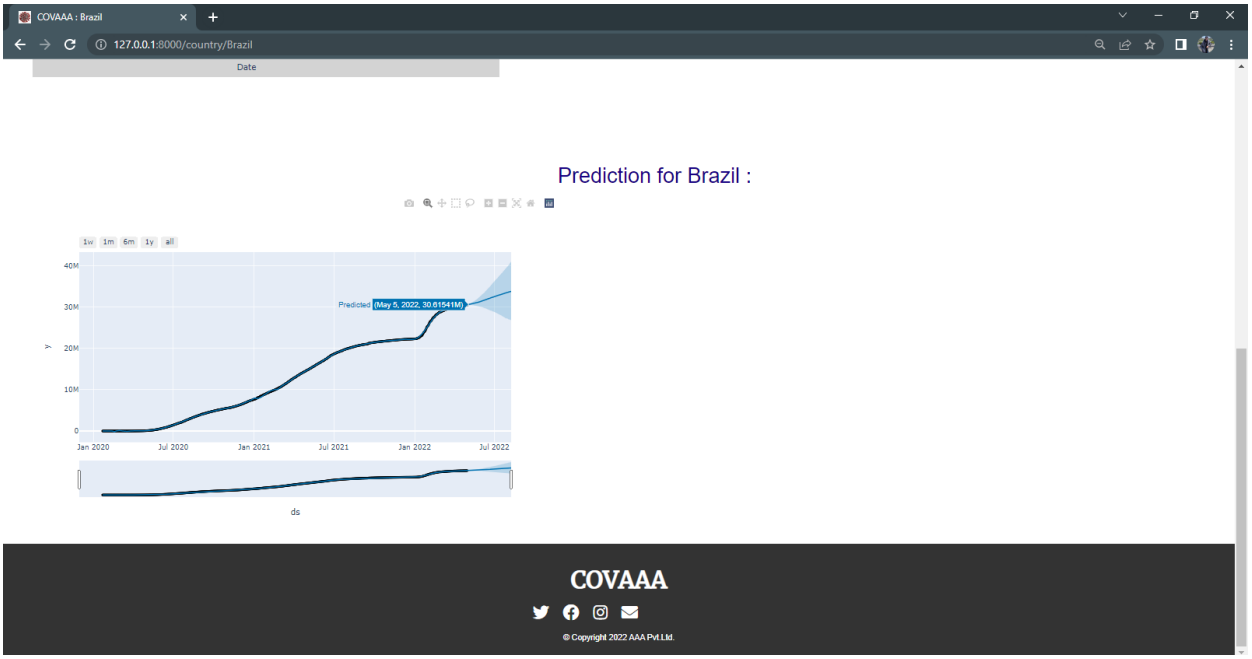
At the end of the World page, we take user input of a country and display the prediction of the cases for the respective country.

3. Country Page :



(Figure-14)

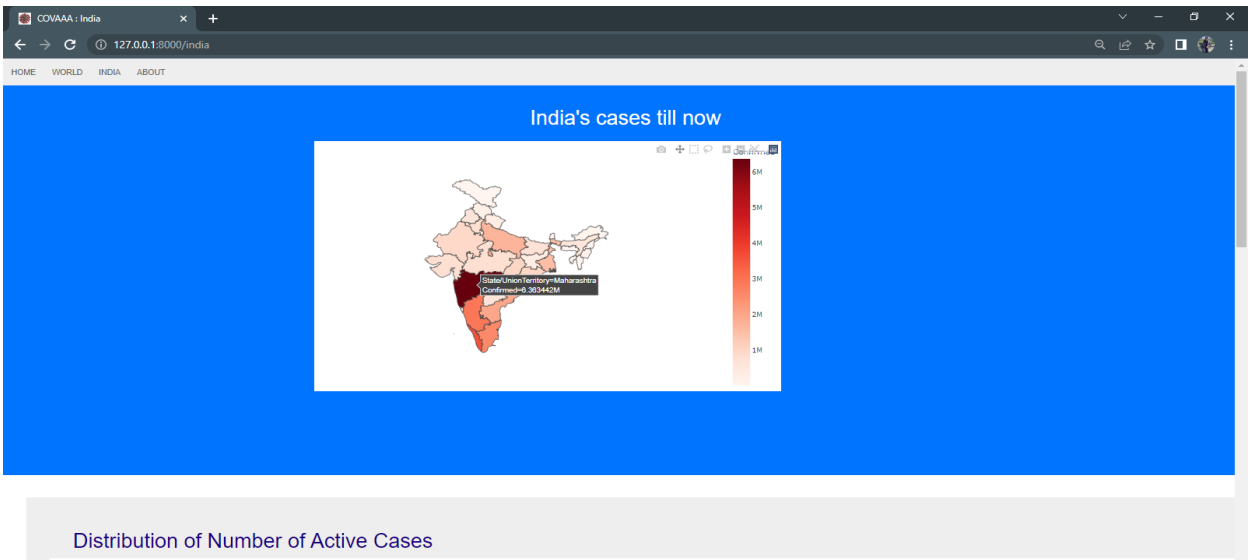
This page shows the users a graphical representation of cases for the country they entered.



(Figure-15)

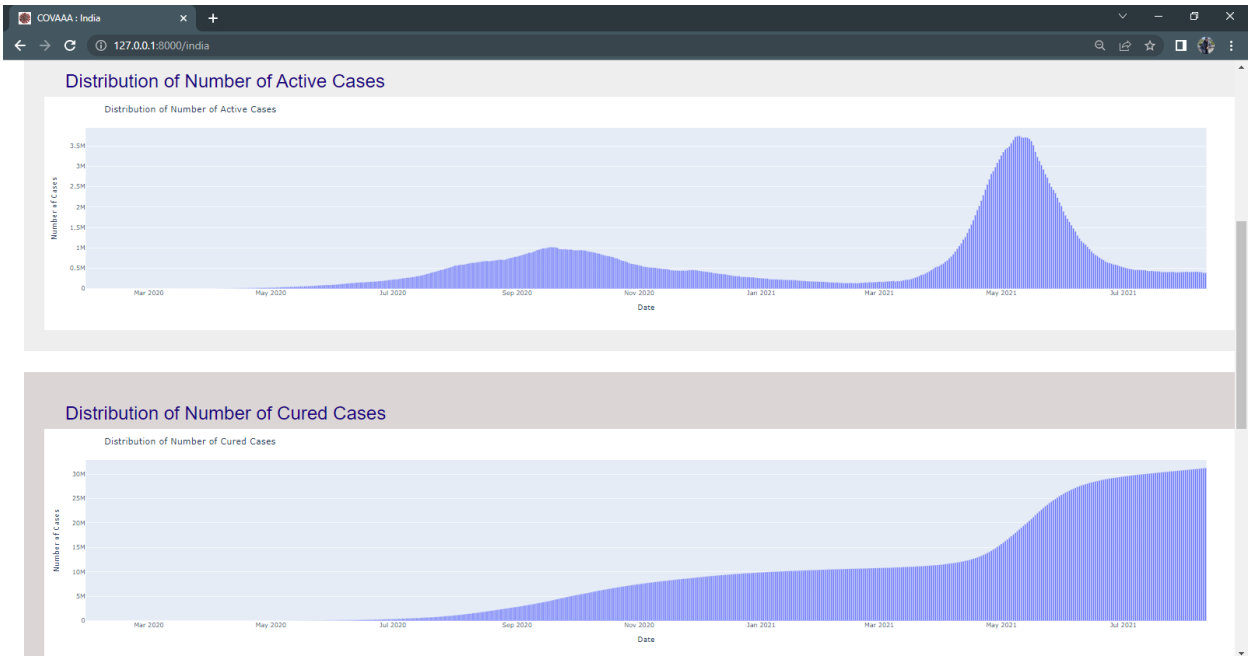
It also shows a prediction of the number of cases for the next 30 days in that country.

4. India Page



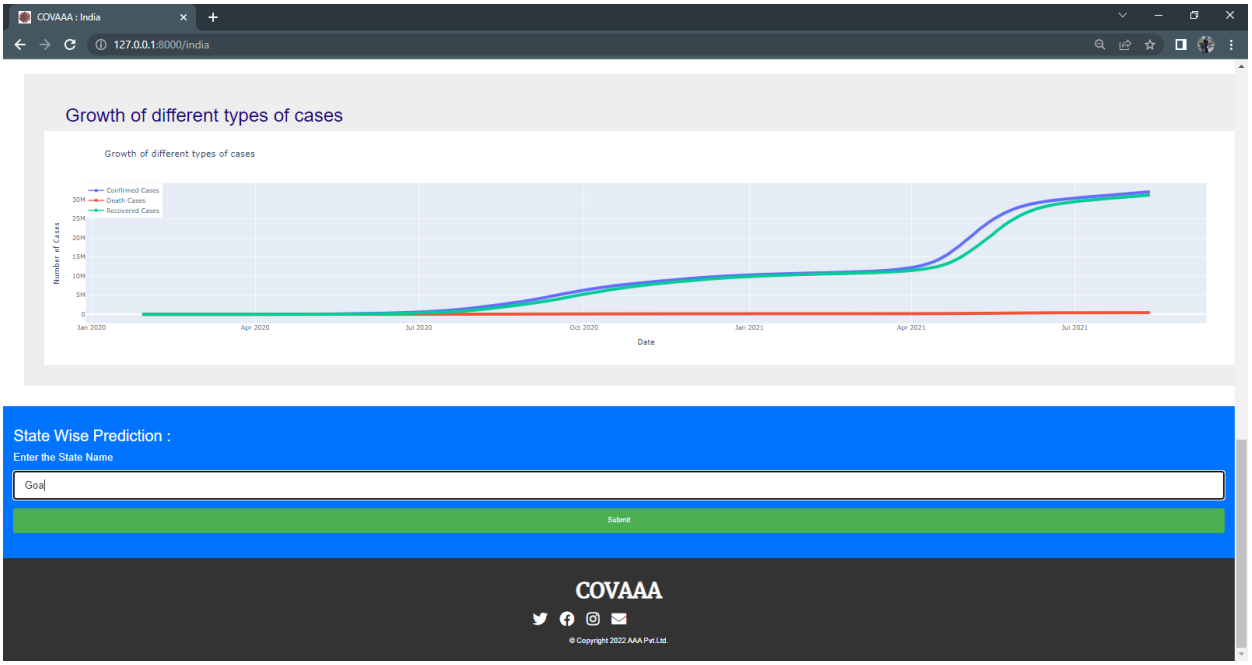
(Figure-16)

This is the 2nd page → India. Here at the top we display the map of India showing the total confirmed cases as you hover over the state.



(Figure-17)

This is the India page, here we display the distribution of Number of Active cases and Distribution of number of cured cases.



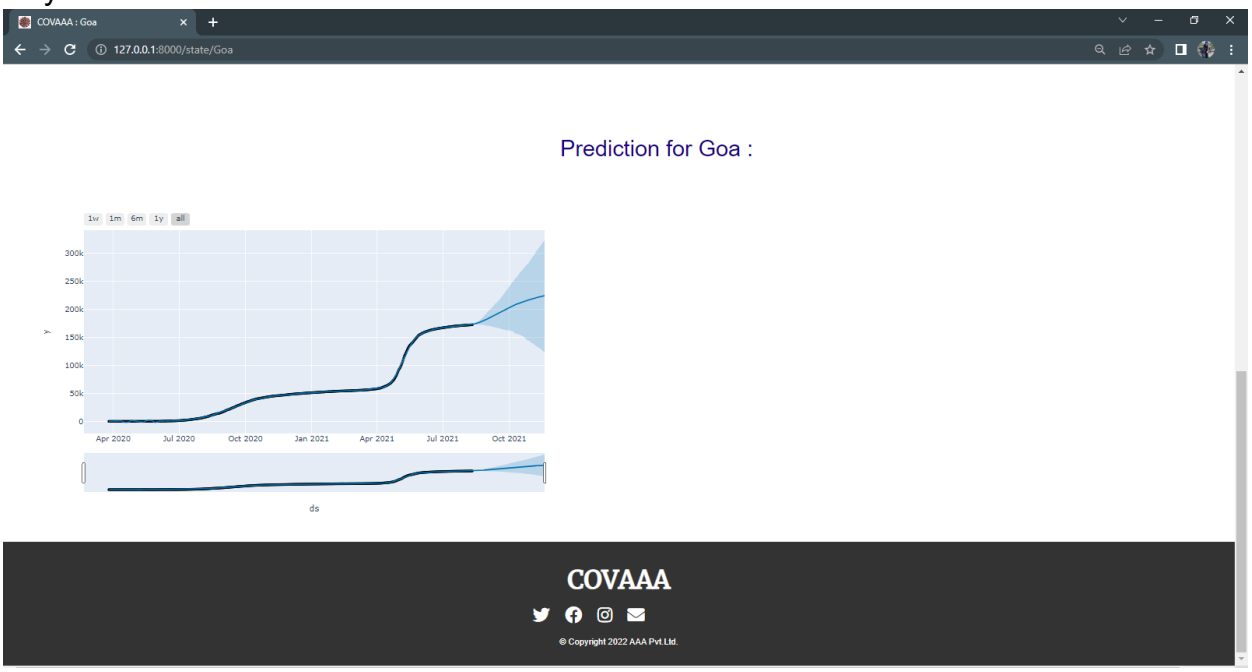
(Figure-18)

Here at the end of India page, we take user input of the state and then after clicking on the submit, the user can see the prediction of the cases for the next 30 days of the respective state.

5. State Page

(Figure-19)

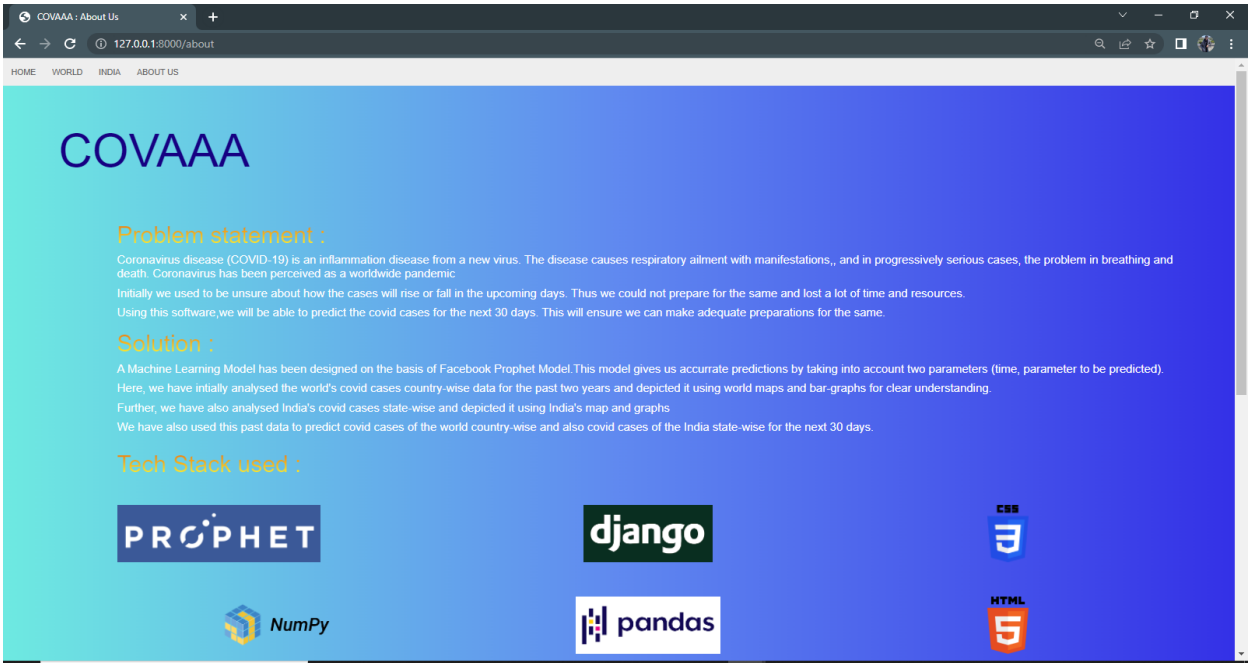
This page shows the users a graphical representation of cases for the state of India they entered.



(Figure-20)

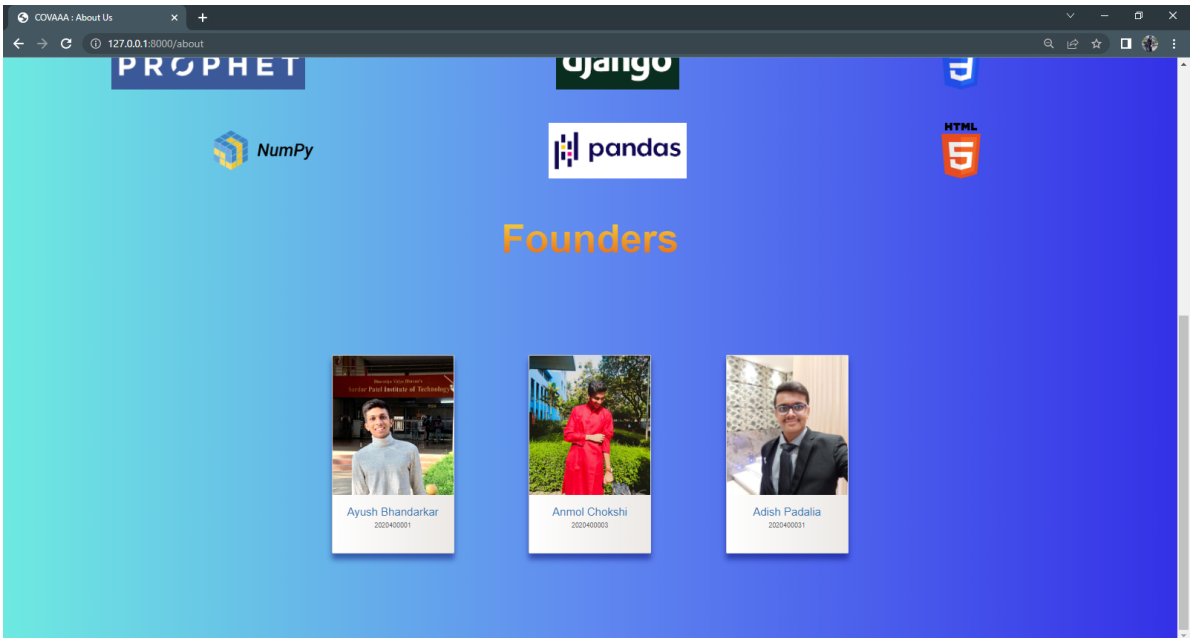
It also shows a prediction of the number of cases for the next 30 days for the state they have entered.

6. About Us



(Figure-21)

Here users can see information about the website, what it does and the problem that it was created to solve. The Tech stack used is also displayed.



(Figure-22)

Along with the Tech Stack, the creators of the website are also shown,

4. Conclusion and Further Work :

Conclusion :

- We have built a website to show users different statistical, graphical and visual representations of the spread of Coronavirus in the world, India and in their choice of country and state.
- We have used the facebook Prophet model to predict cases based on the data of cases available till now. On comparison with real world results, we have concluded that the Prophet model can predict cases with 92.47% accuracy.
- We have used various python libraries to calculate and plot graphs, ML models to predict data, HTML and CSS for an interactive frontend and Django to build our website by connecting all these elements.
- Thus we have made an informative, user-friendly and visually appealing website for all users.

Further Work :

1. Different variants : Visualisation of cases based on different variants.
2. Pincode : Building an API to detect the pincode of the area the user lives and showing the covid data of the area and the neighbouring areas.
3. Age wise prediction : Prediction of the covid case growth for different age groups.

References :

1. M. Rohini, K. R. Naveena, G. Jothipriya, S. Kameshwaran and M. Jagadeeswari, "A Comparative Approach To Predict Corona Virus Using Machine Learning," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 331-337, doi: 10.1109/ICAIS50930.2021.9395827.

Available : https://ieeexplore.ieee.org/document/9395827

2. Locquet, M., Diep, A.N., Beaudart, C. et al. A systematic review of prediction models to diagnose COVID-19 in adults admitted to healthcare centers. Arch Public Health 79, 105 (2021).

Available at: https://archpublichealth.biomedcentral.com/articles/10.1186/s13690-021-00630-3

3. Sujath , R., Chatterjee, J.M. & Hassanien, A.E. A machine learning forecasting model for COVID-19 pandemic in India. Stoch Environ Res Risk Assess 34, 959–972 (2020).

Available at: https://link.springer.com/article/10.1007/s00477-020-01827-8#Sec6

4. Article Source: A machine learning based exploration of COVID-19 mortality risk Mahdavi M, Choubdar H, Zabeh E, Rieder M, Safavi-Naeini S, et al. (2021) A machine learning based exploration of COVID-19 mortality risk. PLOS ONE 16(7): e0252384.

Available at: https://journals.plos.org/plosone/article/citation?id=10.1371/journal.pone.0252384

5. S. Shaikh, J. Gala, A. Jain, S. Advani, S. Jaidhara and M. Roja Edinburgh, "Analysis and Prediction of COVID-19 using Regression Models and Time Series Forecasting," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021, pp. 989-995,

Available at: https://ieeexplore.ieee.org/document/9377137

6. L. W. Mary and S. A. A. Raj, "Machine Learning Algorithms for Predicting SARS-CoV-2 (COVID-19) – A Comparative Analysis," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021,

Available at: https://ieeexplore.ieee.org/document/9591801