A

THESIS REPORT

ON

Covid-19 Prediction & Notification System

Submitted in partial fulfillment of the

Requirements for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

Computer Science & Engineering

by

SONU KUMAR

1709032

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

B.I.T. SINDRI

SINDRI – 808123 (INDIA)

July, 2021

CANDIDATE'S DECLARATION

I hereby declare that the work carried out in this report titled **Covid-19 Prediction & Notification System** is presented on behalf of partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** with specialization in Computer Science & Engineering **submitted** to the department of Computer Science & Engineering**, B.I.T. Sindri, India**, under the supervision and guidance of **Prof. S.C. Dutta**, Associate Professor, Information Technology, B.I.T. Sindri, India.

I have not submitted the matter embodied in this report for the award of any other degree or diploma.

Date: - 31 July 2021   **Student Name: Sonu Kumar**

Place: Sindri   **Roll No.: 1709032**

B.I.T. SINDRI

(DEPARTMENT OF SCIENCE & TECHNOLOGY, GOVT. OF JHARKHAND, RANCHI)
P.O. SINDRI INSTITUTE, DHANBAD-828123

<u>DECLARATION CERTIFICATE</u>

This is to certify that the work presented in this thesis entitled **Covid-19 Prediction & Notification System** submitted by Sonu Kumar, under supervision and guidance of **Prof. S.C. Dutta**, Assistant Professor, Information Technology, B.I.T. Sindri, India.

To the best of my knowledge, the content of this dissertation does not form a basis of the award of any previous degree to anyone else.

Prof. S.C. Dutta,   Dr. F Ansari,

Associate Professor,   Head of Department,

Information Technology   Computer Science & Engineering

B.I.T. SINDRI

<u>CERTIFICATE OF APPROVAL</u>

The forgoing thesis entitled **Covid-19 Prediction & Notification System** is here by approved as a creditable study of research topic carried out and has been presented in satisfactory manner to warrant its acceptance as prerequisite to the degree for which it has been submitted.

It is understood that by this approval, the undersigned do not necessarily endorse any conclusion drawn or opinion expressed there in, but approve the thesis for the purpose for which it is submitted.

Prof. F Ansari

Head of Department

Dept. of Computer Science & Engineering   (External Examiner)

B.I.T. Sindri, Dhanbad

CONTENTS:

INTRODUCTION

Covid-19 pandemic has overwhelmed all the testing facilities in the developed nations, let alone the developing and the underdeveloped ones. Different people have different symptoms with different severity. This situation called for a mechanism which will prioritize people with severe symptoms and old age. The number of testing labs, machines and medical staffs are limited. The world is in a state of medical emergency because of the sudden upsurge of medical demand. We need to do millions of tests per day which sounds unrealistic. Everybody cannot be tested at the same day at the same time. We need a mechanism which will help the medical staffs to form a priority queue based on the condition of the person who wants to get tested. Since the symptoms of Covid is same as common cough and cold, this creates a lot of confusion. This project will be able to segregate people with common cold from covid suspects. This project uses machine learning to determine the degree of urgency of getting tested.

We will also see a Covid-19 notification system which will give real time notifications of Covid-19 figures across the country which will make them alert of the current scenario and instil a sense of need to have Covid appropriate behaviour. This notification system will show state wise data which will help the government, the people and the scientists to understand the severity of the pandemic in each state. This will also help people choose the states where they can travel and where they cannot. This will also help the state governments to decide the strictness of the lockdown in their respective states. The states where the cases are low can have free movement of people and open commercial and recreation places whereas the states with high covid numbers must stay under strict lockdown. Also, the data of number of active cases will help the governments arrange beds, oxygen and medical supplies. Since all the data will be available to the public via this notification system there will be more transparency in the system because the governments will not be able to hide any data.

OBJECTIVE

The prime objectives of this project are:

➢ To predict the chances of having COVID-19 infection for a patient.

➢ To classify a person as less prone or more prone to COVID-19 infection so that the person having higher chances should be tested first and the person having minor risks can just quarantine themselves.

➢ To filter out people who genuinely need medical attention from the masses

➢ To segregate patients with common cough and cold from people who genuinely have covid, since both have similar symptoms.

➢ To reduce the burden of medical service provider up to some extent because all the medical systems across all countries are overwhelmed.

➢ And to display the status of covid-19 active cases, recovery, and deaths in different states of the country.

➢ To promote transparency in covid data by letting people know the accurate figures so that the covid management authorities don't hide the real covid numbers.

➢ To show the real time covid numbers to the people so that people understand the threat posed by the pandemic and they are convinced to follow covid appropriate behavior.

➢ To help the government know the real time covid numbers so that they can make arrangement for treatment of covid patients.

BACKGROUND

Problem Definition and Algorithm:

Task Definition:

Data can be collected on the symptoms of Covid-19. A **machine learning model** is then trained on the data to find out the probability of a person having the infection. The more the probability the more the patient needs medical attention. The model is then used to find out whom to test for the infection first under a limited testing capacity. The same model can be used to find out potential candidates for conducting random tests.

We cannot avoid the curiosity about knowing what's happening in the world. People nowadays spend most of their time in searching for the number of cases about new patients or deaths in their state. This program will not only help them by providing real-time updates but also while saving their time. The **covid notification system** obtains real time data at particular intervals of time from the websites preferably government websites which update their data on the real time basis and displays it on the screen. It frequently accesses those websites to refresh the data.

Algorithm:

Features:

- Body temperature – Continuous
- Body pain – Categorical
- Age – Discrete
- Runny nose – Categorical
- Cough - Categorical
- Difficulty in breathing – Categorical
- Patient travelled in affected area - Categorical

The above are the input features. **Logistic regression** is used in this project for training the model and for the prediction of patient's chances having infection

of COVID-19. The output is a value between 0 and 1. The more the number is closer to 1, the more is the urgency of getting tested.

Technology used:

- **Python 3.7.9**
- **Flask**
- **Scikit-learn**
- **HTML**
- **CSS**
- **Bootstrap**

Experimental Evaluation:

Methodology:

Logistic Regression: Logistic regression is the appropriate regression analysis to conduct when the dependent variable is dichotomous (binary). Like all regression analyses, the logistic regression is a predictive analysis. Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables. Sometimes logistic regressions are difficult to interpret; the Intellects Statistics tool easily allows you to conduct the analysis, then in plain English interprets the output.

Information about the data:

	fever	bodyPain	age	runnyNose	diffBreath	visitedAffectedArea	infectionProb
0	102	1	44	1	-1	1	1
1	102	1	16	0	0	0	0
2	103	0	25	1	-1	0	0
3	99	0	20	1	-1	0	0
4	103	1	94	1	1	0	0

Variable types :

```
In [5]:  df.info()

         <class 'pandas.core.frame.DataFrame'>
         RangeIndex: 10115 entries, 0 to 10114
         Data columns (total 7 columns):
          #   Column              Non-Null Count  Dtype
         ---  ------              --------------  -----
          0   fever               10115 non-null  int64
          1   bodyPain            10115 non-null  int64
          2   age                 10115 non-null  int64
          3   runnyNose           10115 non-null  int64
          4   diffBreath          10115 non-null  int64
          5   visitedAffectedArea 10115 non-null  int64
          6   infectionProb       10115 non-null  int64
         dtypes: int64(7)
         memory usage: 553.2 KB
```

	fever	bodyPain	age	runnyNose	diffBreath	visitedAffectedArea	infectionProb
count	10115.000000	10115.000000	10115.000000	10115.000000	10115.000000	10115.000000	10115.000000
mean	100.022145	0.493821	50.642313	0.501236	0.012259	0.507860	0.503707
std	1.423726	0.499987	28.913035	0.500023	0.814303	0.499963	0.500011
min	98.000000	0.000000	1.000000	0.000000	-1.000000	0.000000	0.000000
25%	99.000000	0.000000	25.000000	0.000000	-1.000000	0.000000	0.000000
50%	100.000000	0.000000	51.000000	1.000000	0.000000	1.000000	1.000000
75%	101.000000	1.000000	76.000000	1.000000	1.000000	1.000000	1.000000
max	104.000000	1.000000	100.000000	1.000000	1.000000	1.000000	1.000000

Notification System:

As this is the desktop notifier application, we will be using the **plyer module** in our program. The Plyer module does not come built-in with Python. To install it externally, write the following command on your terminal:

pip install plyer

This program will give you the real-time update about the number of new cases, deaths, and the recovered cases of coronavirus according to time (after 1 hour or 2 hours) within a state. It will also provide information about how many of them are Indian national and foreign nationals.

FUTURE SCOPE

All the new Covid-19 test data should be added to the database which will enlarge the database. The larger the database, the better will be the future prediction. We may more parameters to the data set which will increase the accuracy of this probability detector by many folds. We may add a predictor of oxygen and beds which will make it easy for the government and the oxygen generation plants to understand the oxygen demand objectively. As this virus belongs to the SARS virus family which has already infected many people in the past, there is a possibility that a virus from the same family will infect human beings in the future. We may see similar symptoms of that virus as well. This machine learning model and the data can be used for prediction then.

The notification system can be used in various fields for the real time display of data. It just accesses the websites and fetches data from there to display them to the people. That system can be used to display any data of public interest. A template can be built where the people should be given the option to choose which type of data they want to see. The data can be related to any disease, or progress report of a any games or real time satellite data or anything which the viewer wants to see. We can also add the provision of fetching data from multiple websites to cross check the data to know about it's authenticity. We can also rate those websites according to the authenticity of data they provide. This will help people to know which website provides authentic data and which website provides outdated or forged data.
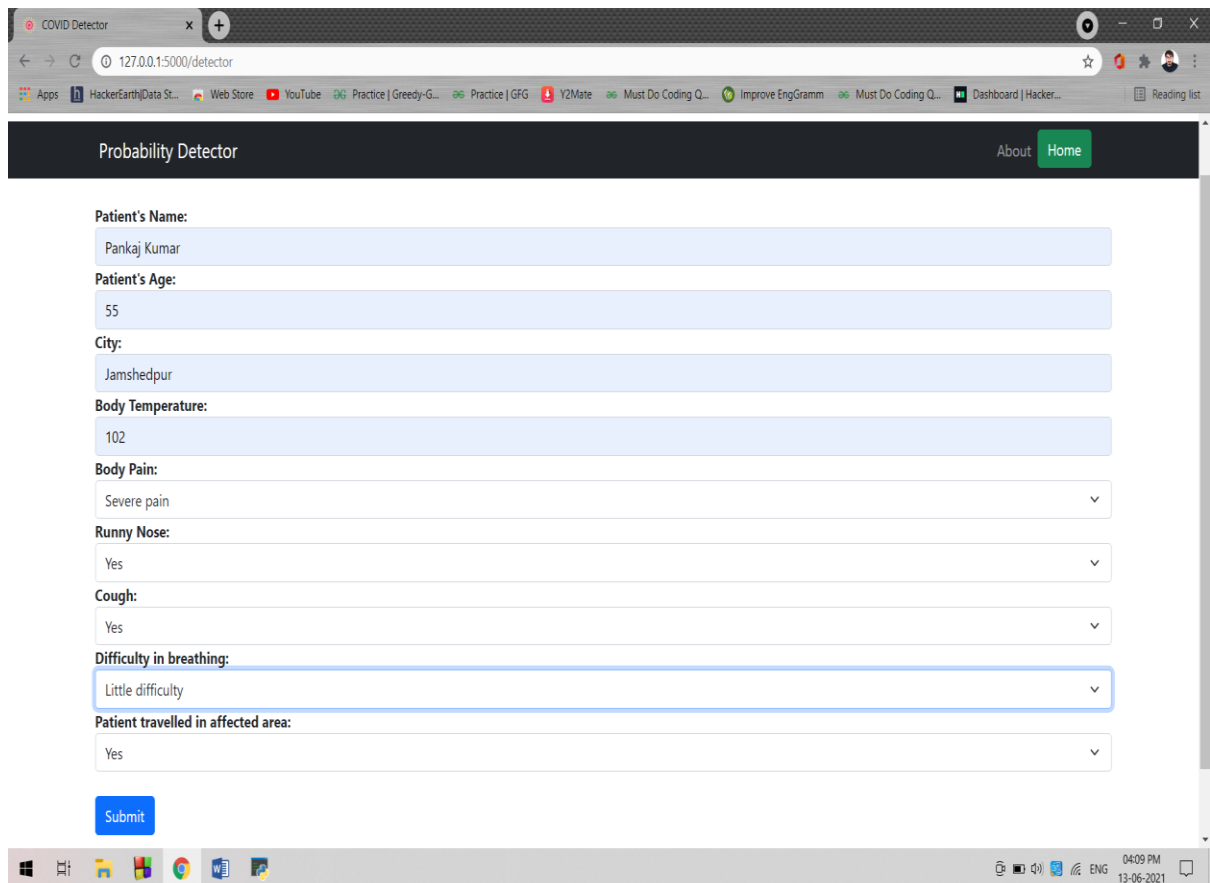
WEB APP VIEW

(Github link: https://github.com/panks123/Covid19-Detector-and-Status-TrackerFinal-Year-Project)

Landing page: This is the first web page or the home page of the web application. The home page shows that the project is about Covid Predictor and Status tracking. It contains two buttons which will take us to the Probability Detector and State wise covid tracker.
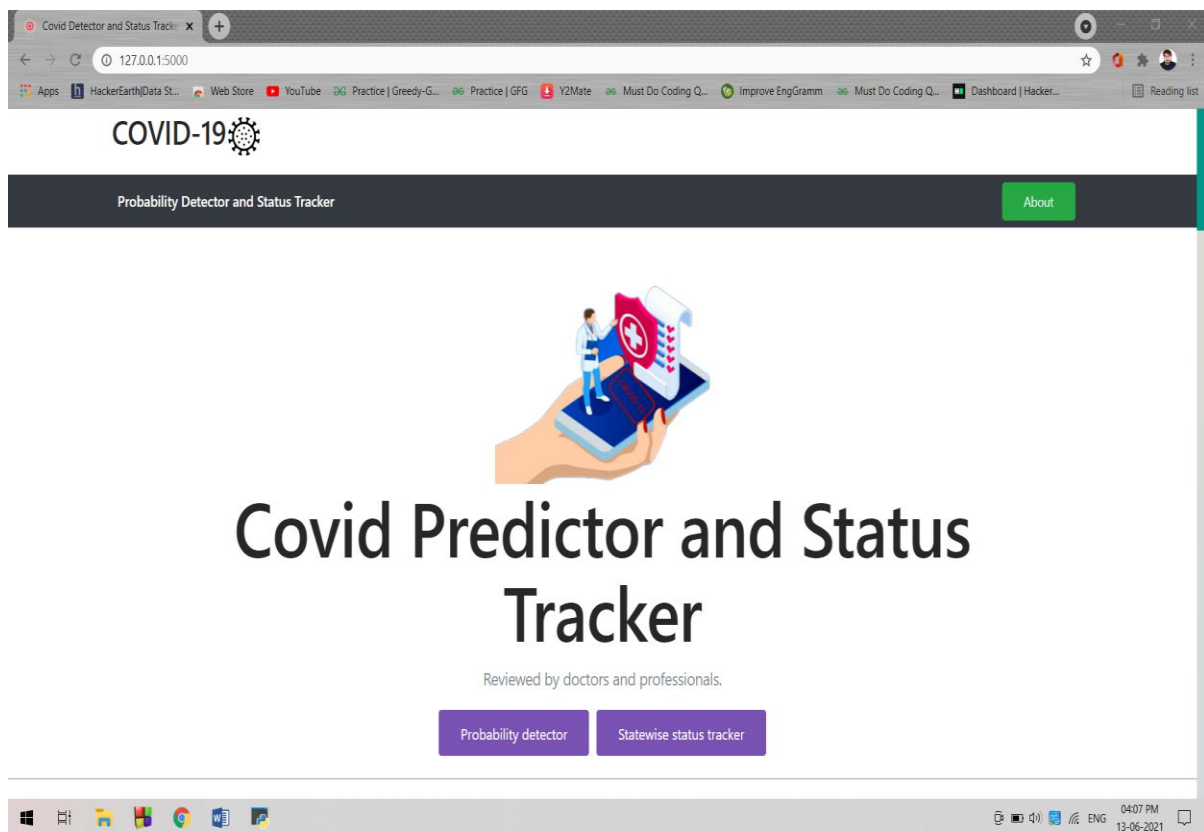
Predictor app:

We can reach this page through the buttons given in the previous page which is called the home page. It is a form which takes input as various symptoms of the patient.
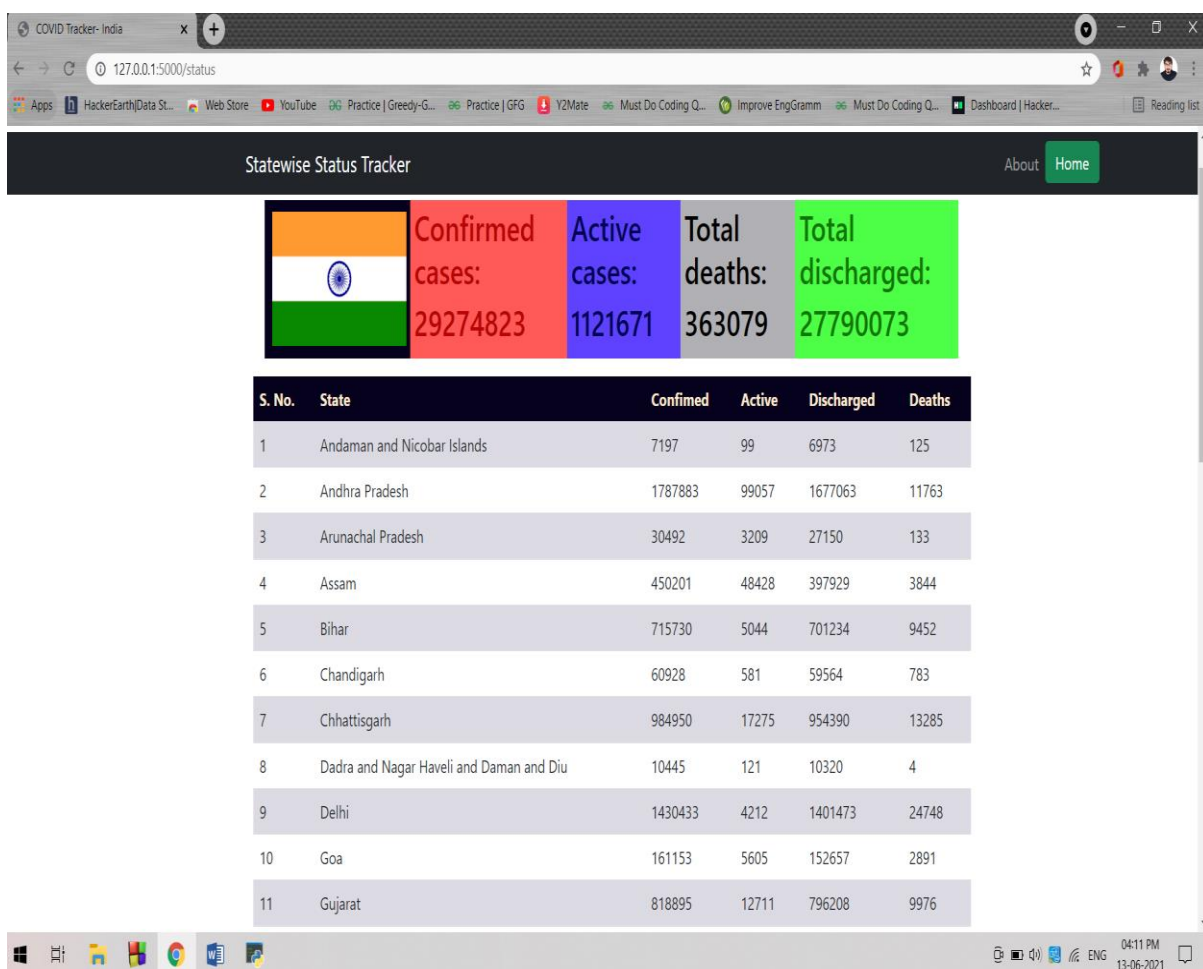
Result page:

On submitting the form this page will display the result of the person in terms of probability of having chances of COVID-19 infection. The more the probability the more the patient needs urgent medical assistance.

The page also highlights the preventive measures of Covid. Users can learn them which will help them to learn Covid appropriate behaviour.
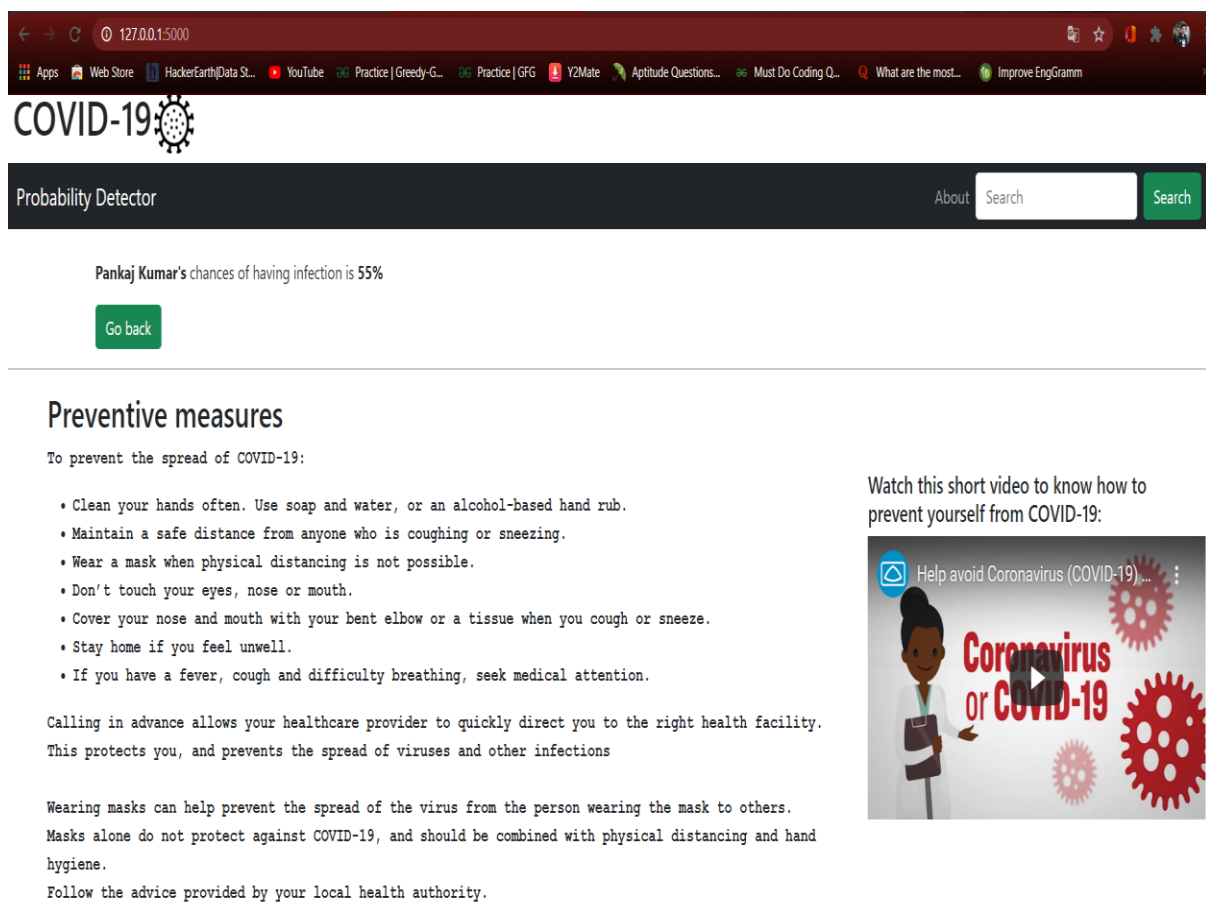
Status Page:

It displays real time covid-19 cases occurring across several states of India in the form of a table. It obtains real time information from different websites preferably the government websites because they are of highest credibility and displays them on the screen to give an idea of the ongoing covid scenario in different states of the country to the people and the state and the central government to that they can be better prepared.

<u>RESULT</u>

We have successfully applied Logistic Regression model to predict the chances of having COVID-19 infection in a person. Our model is providing the probability number which is telling us the chances a person is having it. We used the real covid data set to predict accurate and reliable data.

Also we are displaying the real time status of Covid-19 cases across several states of India by fetching data from reliable websites like the government website and helping the government to spread covid awareness and also benefitting the people by providing them with transparent data.

<u>CONCLUSION</u>

We have successfully created a machine learning model and presented it as a web app which can be supplied to various hospitals, doctors and other medical service organizations so that it can aid in examining the symptoms of the patient. This gives the advantage of classifying people based on the likeliness of having Covid-19 and test the people with higher probability at the earliest. The people with lower risk can quarantine themselves till their chance arrives. This will filter out people who genuinely need medical attention from the masses and reduce the burden on the medical service providers of India.

We have also created a notification system successfully which fetches data from reliable websites and displays it. It promotes transparency in data display. It is able to refresh the data frequently so that it is able to provide real time data to the users.

REFERENCES AND BIBLIOGRAPHY

- Scikit-learn : https://scikit-learn.org/stable

- **Data visualization**:
  Matplotlib
  (https://matplotlib.org/tutorials/index.html)

- **Exploratory data analysis:**
  Numpy
  (https://numpy.org/doc/stable/user/tutorials_index.html)
  Pandas, Matplotlib
  (https://matplotlib.org/tutorials/index.html)

- **Exploratory data analysis:**
  Numpy
  (https://numpy.org/doc/stable/user/tutorials_index.html)
  Pandas, Matplotlib
  (https://matplotlib.org/tutorials/index.html)

- Mastering Python for Data Science, Sameer Madhavan 2nd Edition.

- Hands-On Machine Learning with Scikit-Learn & TensorFlow (Concepts, Tools and Techniques to Build Intelligent Systems), *Aurélien Geron*

***** E.O.F. *****