

“ Coronavirus Probability Detector ”

SUBMITTE

Name: Hemant Dharmendra Khambait

Roll No: N20111042

Guided By
Dr. Jyoti yadav

**M.Sc.(Computer Science) – SEM
III**



INTRODUCTION :

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS- CoV-2). It was first identified in December 2019 in Wuhan, China, and has resulted in an ongoing pandemic. The first case may be traced back to 17 November 2019.

As of 1 February 2022, more than 395 million cases have been reported across 188 countries and territories, resulting in more than 5,705,754 deaths. More than 315 million people have recovered. The virus is primarily spread between people during close contact, most often via small droplets produced by coughing, sneezing, and talking. The droplets usually fall to the ground or onto surfaces rather than travelling through air over long distances.

Less commonly, people may become infected by touching a contaminated surface and then touching their face. It is most contagious during the first three days after the onset of symptoms, although spread is possible before symptoms appear, and from people who do not show symptoms.

PROBLEM STATEMENT :

India has a high population country and hence it is not possible to check everyone for the SARS-CoV-2 virus which is the cause of the corona virus outbreak.

Everyone who is living in this country is affected due to the high population as there not enough facilities for each person. The spread of false news is causing the people to worry. By the use of self evaluation for themselves and their family members they can take the decision to inform the authorities and ask for help.

If someone enters false details then that would result in a small error in the assessment of the outbreak.

PROJECT AIM :

As we continue to navigate the COVID-19 pandemic, people around the world often ask themselves the same question as they head into various buildings and scenarios:

What changes can we make to reduce the risk of COVID-19 transmission?

The Coronavirus has rapidly spread to all parts of the world. Research is continuing to find a cure for this disease while there is no exact reason for this outbreak. As the number of cases to test for Coronavirus is increasing rapidly day by day, it is impossible to test everytime due to the time and cost factors. Over recent years, machine learning concepts has turned very reliable in the medical field.

To help provide some insight into this question, This project “Coronavirus Probability Detector” using python and machine learning concepts is very different yet helpful for all of us in such a pandemic situation.

We will be using Jupyter Notebook for the initial development and then create UI which tells whether the person has an infection or not based on input features using Virtual Studio Code IDE.

IDEA : In such situation, it is necessary to prioritise the patients and patients with more chances of having infection should be tested first. For ex - imagine a situation where 20,000 patients have come for check-up but only 8,000 devices are available. In such scenario, top 8,000 patients with more chance of infection should be given priority because they have more chance of spreading the infection. Others can be tested later since they might not be infected by the virus.

SOLUTION :

We can first collect data regarding the age and symptoms of people affected by COVID-19. The actual type of data can be decided by the specialists from medical field. I hereby present just a sample data which is randomly generated by me for an example.

S.No.	Age (in years)	Body temp. (in °F)	Fatigue	Cough	Body pain	Sore throat	Breathing difficulty	Infected
1.	38	99.42	1	0	1	0	0	0
2.	72	104.65	1	1	1	2	2	1
3.	56	102.32	0	1	1	0	1	1
4.	52	102.50	1	0	0	1	1	0
5.	45	98.65	0	1	1	1	0	0
6.	84	101.59	1	1	1	0	2	1
7.	62	103.07	1	0	1	0	2	1
8.	44	100.96	0	1	0	1	1	1
9.	68	100.35	1	1	0	2	0	1
10.	65	102.69	1	1	1	1	1	1

Note:

In the above data, in fields of 'Fatigue', 'Cough', 'Body pain' and 'Infected':

0 – No

1 – Yes

In fields of 'Sore throat' and 'Breathing difficulty':

0 – Not at all

1 – Little bit

2 – Too much

The above data is just a randomly generated sample.
Government can avail data for 50,000-60,000 people across globe.

Now a machine learning model can be trained and tested on the data collected, where the 'features' will be 'age' and 'symptoms' and 'label' will be 'probability of patient getting infected'.

So now when a new patient will come, we can simply ask the patient about age and various symptoms and the model will tell the probability of patient being affected.

For example, if patient gives the information: Age - 42 years

Body temp. - 102.65

°F Fatigue - 1

Cough - 0

Body pain - 1

Sore throat -

1

Breathing difficulty - 0

The machine will give output – 46% (just an example, not by model)

So, this way the probability of all the patients can be recorded and then sorted in decreasing order simultaneously so that patients with more chance of infection get priority and be tested first.

Whatever discussed till now was about the method of working of model in scientific terms. In common man's terms, it will be like a software with user friendly interface where patient can fill a digital form asking various details about symptoms. On clicking the submit button, the patient will be able to know the probability of him getting infected by COVID-19.

Although doctors can use this by prioritising the patients, even the people at homes can check their probability of having infection by COVID-19 and can better stay at home instead of going outside and visiting doctor if probability is less than 40% or so. This will also prevent crowd

Machine learning model parameters

- A team of doctors can sit down to find out the best model parameters.
- A sample set of such parameters is as follows:
 - ✓ Features
 - Average Fever - Continuous
 - Body Pain - 0/1 - Binary
 - Age - Discrete
 - Runny Nose
 - Difficulty breathing - Categorical : -1/0/1
 - ✓ Labels:
 - Probability of Covid-19 Infection



Activate Windows
Go to Settings to activate Windows.



The Idea

- The idea is to stop the transmission by prioritizing tests and hence detecting the cases quickly
- 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 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 potential candidates for conducting random tests



CORONA VIRUS
Activate Windows
Go to Settings to activate Windows.



Browser tabs: jupyter notebook - Google Search, Project Jupyter | Home, CodeWithHarry - YouTube, Download Python | Python.org, Home Page - Select or create a..., My Solution - Jupyter Notebook, sklearn.linear_model.LogisticRegression

Address bar: localhost:8889/notebooks/My%20Solution.ipynb

Navigation: Apps, Gmail, YouTube, Maps, Python Django Tut..., JavaScript Tutorials..., Site administration..., C++ Tutorials in H..., Web Development..., Python tutorials for...

Jupyter My Solution Last Checkpoint: 6 minutes ago (unsaved changes) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

```
Out[64]: array([0, 0, 0, ..., 1, 1, 1], dtype=int64)
```

```
In [64]: from sklearn.linear_model import LogisticRegression
```

```
In [65]: clf = LogisticRegression()
         clf.fit(X_train, Y_train)
```

```
Out[65]: LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
                           intercept_scaling=1, l1_ratio=None, max_iter=100,
                           multi_class='auto', n_jobs=None, penalty='l2',
                           random_state=None, solver='lbfgs', tol=0.0001, verbose=0,
                           warm_start=False)
```

```
In [67]: clf.predict([[100, 1, 22, 1, 1]])
```

```
Out[67]: array([0], dtype=int64)
```

```
In [ ]:
```

Typesetting math: 0%

Windows taskbar: Type here to search, Taskbar icons, System tray: ENG

Brave notification: Solana Riptide Hackathon, \$5 million in prizes+seed funding | Starts Feb. 2 | Join now, Activate Windows, Go to Settings to activate Windows.

Downloads/Coronavirus-Probabi... x index.html

localhost:8889/view/Downloads/Coronavirus-Probability-Detector-main/index.html

Apps Gmail YouTube Maps Swayam Central PUCSD Placement... GitHub Dashboard Telegram Solve Me First | Hac... Object Oriented Pr... Reading list

COVID-19 Detector Home About Us Contact Us Search Search

Coronavirus Probability Detector

Enter fever Value

Enter Age Value

Body Pain

Runny Nose

Breathing Difficulty

Submit

Activate Windows
Go to Settings to activate Windows.
Show all

Coronavirus-Probabi...zip

Type here to search

16°C 07:53 AM 18-02-2022

Downloads/Coronavirus-Probabi... x show.html x index.html

localhost:8889/view/Downloads/Coronavirus-Probability-Detector-main/show.html

Apps Gmail YouTube Maps Swayam Central PUCSD Placement... GitHub Dashboard Telegram Solve Me First | Hac... Object Oriented Pr... Reading list

COVID-19 Detector Home About Us Contact Us Search Search

Coronavirus Probability Detector

Thanks for using Covid probability detector.

Patient's probability of Infection is {{inf}}%. Thank You

Go Back

Activate Windows
Go to Settings to activate Windows.
Show all

Coronavirus-Probabi...zip

Type here to search

16°C 07:53 AM 18-02-2022

OBJECTIVES :

The Objectives of the project are:

- To Stop the transmission by prioritizing tests and hence detecting the cases quickly.
- For Collecting Data of symptoms of COVID-19.
- We use a machine learning model is then trained on the data to find out the probability of a person having the infection.
- The model is then used to find out whom to test for the infection first under a limited testing capacity.
- This model can be used to find potential candidate for conducting random tests.
- Provide knowledge, change attitudes, commitments to COVID-19 prevention and control.
- Provide support for COVID-19 prevention and control.

TECHNOLOGY CONCEPT :

The developed hack is End user input based Web application. It takes the input from the user for their details and shows a list of symptoms which they may be experiencing and can confuse with Covid-19.

On the basis of their entered choices the recieved input will be processed and their probability will be shown to them. If the probably is severe or high than the person is prompted to consult a doctor. His input will be stored in the system for Machine Learning.

On the basis of this the priority of the tests can be done so as to reduce the transmission of the disease and also to make the most of limited testing capacity in such highly populated country such as India. A Machine Learning Model is then trained on the data to find out the probability of a person having the infection.

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 potential candidates for conducting random tests.

TECHNOLOGIES USED :

Python for data scrapping , preprocessing , visualization etc.

Libraries Used :

- 1) Pandas
- 2) sklearn
- 3) Numpy
- 4) Plotly

Django and Virtual Studio Code IDE.

BENEFITS FROM THIS :

There is a situation of fear among the people and spread of fake news is further making the situation worse. This Hack can be used to control this fear by self assessment of the user and their family. Also, if the prediction comes severe the authorities will be informed and this will be used to control the COVID-19 Pandemic upto a very large extent.

REFERENCE BOOKS :

Mastering Python For Data Science : Samir

Madhavan Learning With Python

WEBSITE :

<https://api.covid19india.org>

www.wikipedia.com

<https://www.w3schools.com/>

<https://stackoverflow.com/>

<https://plotly.com/>

<https://www.pythonanywhere.com>

THANK YOU