

## PART A: Innovation Development Plan (Product/Service/Process)

### 1. Team Details:

#### Team Lead:

**Name** : SAI HARITHA S

**Email** : saiharitha69@gmail.com

**Contact No.** : 9500455005

**Aadhar No.** : 3093 3545 9043

#### Team Member details excluding Team Lead;

Sr.No.	Name	Dept	Gender	Email	Contact No.
1	SAI HARITHA S	CSE	FEMALE	Saiharitha69@gmail.com	9500455005

#### Team Mentor details:

Sr.No.	Name	Dept	Gender	Email	Contact No.
1	VANI RAJASEKAR	CSE	FEMALE	Vanikecit.cse@kongu.edu	9659902769

#### Institute Name and Address:

KONGU ENGINEERING COLLEGE  
THOPPUPALAYAM  
PERUNDURAI, ERODE- 638 060

### 2. Name of the Innovation/Prototype:

Medicare for prevention of Corona Virus Using Machine Learning

### 3. Sub-Theme:

Working Conditions, ensuring occupational health and safety issues

### 4. Define the problem & relevance to today's market/society/industry need:

As of now covid-19 or corona virus is one of the global pandemic which has been spreading day to day, even the people from younger to elder know what is mean by covid-19, this led me to do this project, where it takes 2 to 3 days to find the occurrence of covid-19 in the particular person, but this software helps to detect the virus in the human body within couple of minutes using user input or X-Rays

### 5. Explain the solution developed (Max 100 words)\*:

The main aim of rising this project is to find the occurrence of covid-19 in human being within couple of minutes using two methods

1. User Input

2. X-rays.

In first method the user should give the input in the sidebar which consists of symptoms of corona virus, or using the drag bar else the user can fill in numerical values, after checking the exact body condition, Second method can be done using X-rays, in this method the X-rays of the particular person is given as user input which will help to detect the particular person is suffering from corona virus or not.

## **6. Describe the product/process/ service and write how it is innovative / unique:**

Machine Learning plays a major role in this project and they were almost 4 algorithms which are used in the deployment of this software and the best among them is used for the detection of corona virus. Machine learning in medicine has recently made headlines. It's safe to say there are too many manual processes in medicine. Health Catalyst believes machine learning is the life-saving technology that will transform healthcare. It's been said before that the best machine learning tool in healthcare is the doctor's brain.

**The algorithm which is used in the Prediction of Covid-19 is RANDOM FOREST CLASSIFIER**

### **RANDOM FOREST ALGORITHM**

Random forests is a supervised learning algorithm. It can be used both for classification and regression.

It is also the most flexible and easy to use algorithm. A forest is comprised of trees. It is said that the more trees it has, the more robust a forest is. Random forests creates decision trees on randomly selected data samples, gets prediction from each tree and selects the best solution by means of voting.

Random forests has a variety of applications, such as recommendation engines, image classification and feature selection.

It can be used to classify loyal loan applicants, identify fraudulent activity and predict diseases. It lies at the base of the Boruta algorithm, which selects important features in a dataset.

### **STEPS INVOLVED IN THIS MODEL**

The steps involved in deployment of this project is

#### **COLLECTION OF DATA:**

The datasets are collected from the kaggle website and fed into the machine and we use pandas and numpy to read the datasets

#### **ANALYSE THE DATA**

The data which has been collected from the website should be analyzed using various library packages

#### **CLEAN THE DATA**

If any of the instances consist of NA values it should be replaced by zero or average of other values which is called as data preprocessing

## **SPLIT THE DATA**

The datasets are separated into testing data and training data using numpy and pandas in order to test the accuracy level and to know other further information

## **ANALYSE THE ALGORITHMS**

In the development of this model we have used 4 algorithms to check the accuracy score and errors by the way Random forest algorithm gives the highest accuracy score as 94% and less errors, So this algorithm is chosen for the detection

## **PREDICT THE OUTPUT**

Finally the output will be displayed after undergoing various process and calculations

## **WORKING MODEL:**

The symptoms of corona virus such as fever, cold, shortness\_of\_breath, headache, throat pain, DPF, Age is given as input whereas the output will be 0 or 1 (yes or no) category, which helps to predict the person is suffering from Covid or not. The datasets are fed into the machine first it contains the description of the data and we should analyse various data based on the library packages in python (i.e.) Seaborn, count plot, heat map and so on and in which each data of the particular input undergoes data pre-processing which helps to clean the data after this process where the visualization of each inputs is represented using pyplot. Finally by using Random Forest the accuracy score and the classification will be displayed c data the user has given in the sidebar

## **DETECTION OF COVID-19 CASES IN THE WORLD**

Once after displaying the accuracy score and classification, which will be followed by the video using bar chart race which consists of the ups and downs of the country affected by corona virus in which every datasets are collected from the kaggle website and based on that the particular video will be displayed after checking the box on the sidebar

## **DETECTION OF COVID-19 ACROSS INDIA IN EACH AND EVERY STATE**

In this method it consists of almost 4 unchecked boxes named Active Cases, Confirmed Cases, Death Cases and Recovered Cases in sidebar after checking each and every box the particular chart will get displayed based on the unique identification number given by the authority of India which helps to find which state has been affected more and helpful in data visualization

## **DETECTION OF COVID-19 USING X-RAYS**

In this method where the X-Ray of the user is given as the input which undergoes image processing using openCV [Computer Vision] which gives the output as the cropped image and also gives the information about the detection of corona virus in the human.

**7. How is your proposed product/ process/service being different/ better from a similar product/ process/ service, if any, in the market\*:**

Compare to other papers this model will give you greater accuracy score and less errors and it also gives you the total information regarding corona virus cases and it will be more user friendly and informative to the environment in which every people can use it for free of cost and also it is used to consume time and lives of many peoples

**8. If your Innovation is technology based, then specify the TRL Level (Technology Readiness Level) and features of Prototype.\*:**

**It comes under the category of TRL 8**

This project is already executed in the current environment and checked for peoples suffering from corona virus for almost 20 to 25 peoples which gives the correct result and accuracy but the result will be more accurate if we deployed using deep learning

Project Executed in

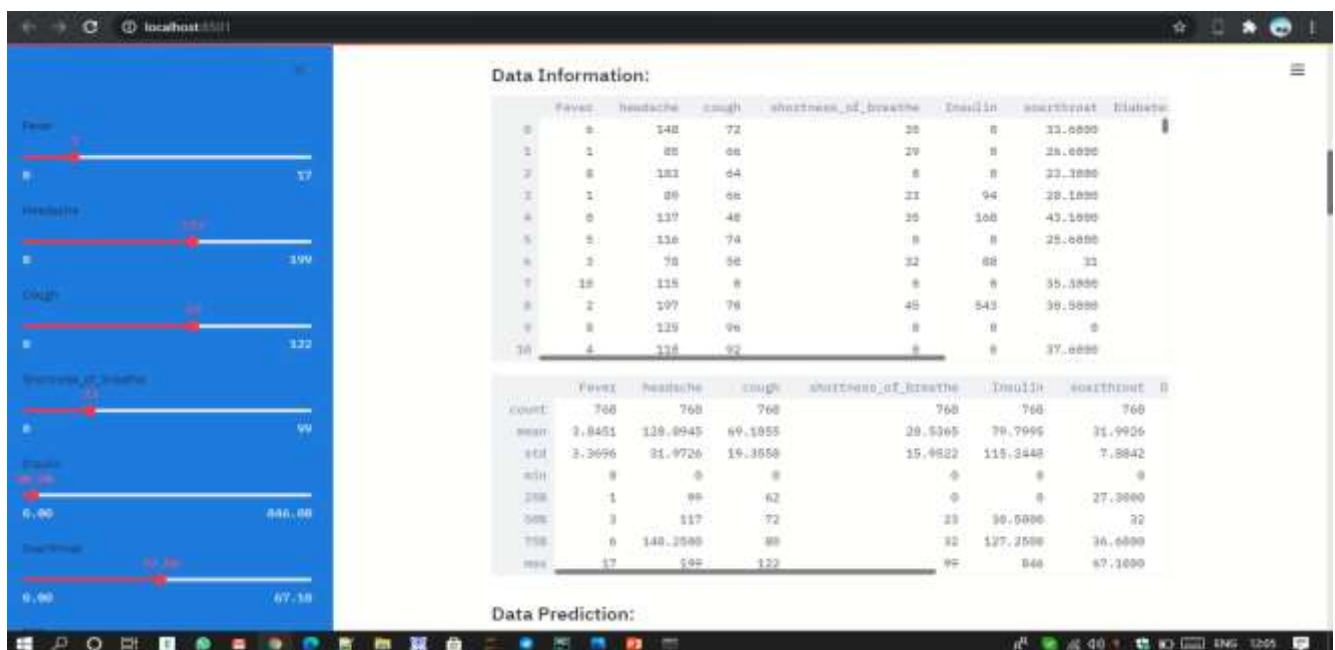
**IIC – Idea presentation [Won 3<sup>rd</sup> prize]**

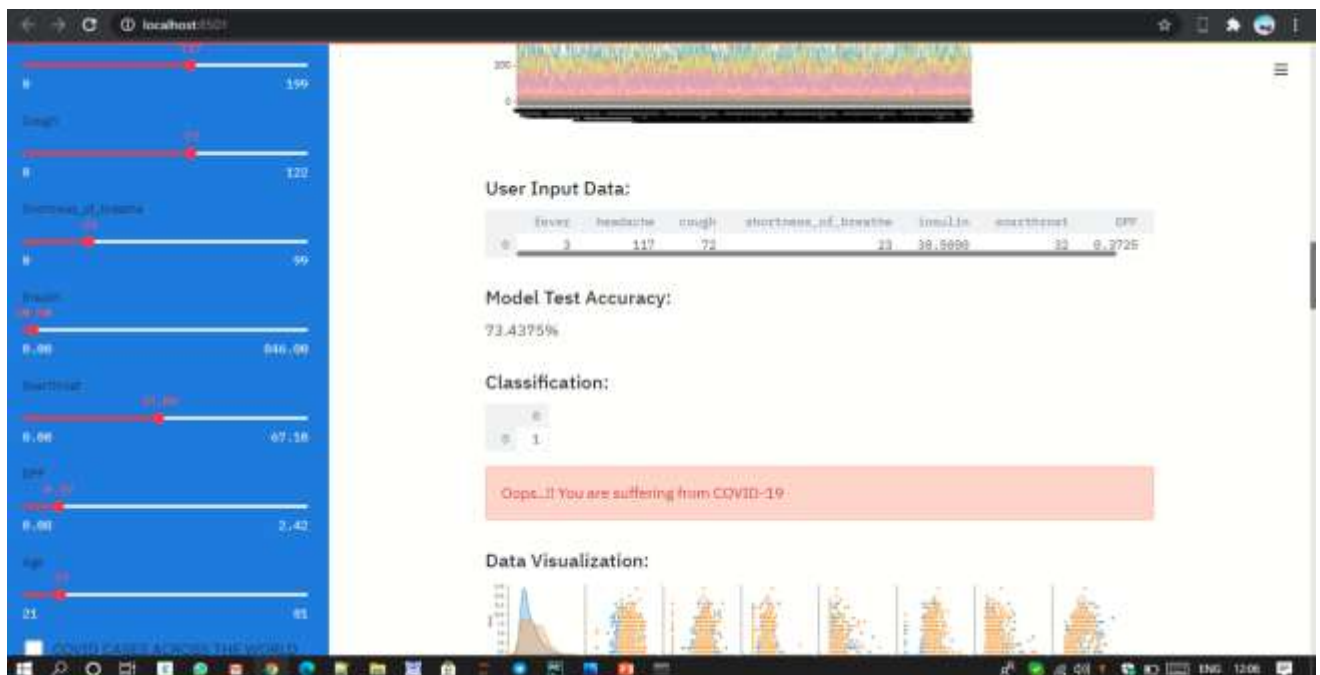
**IIT MADRAS – Paper Presentation [Won 3<sup>rd</sup> Prize]**

**RENAISSANCE – Project Presentation [Won 2<sup>nd</sup> prize]**

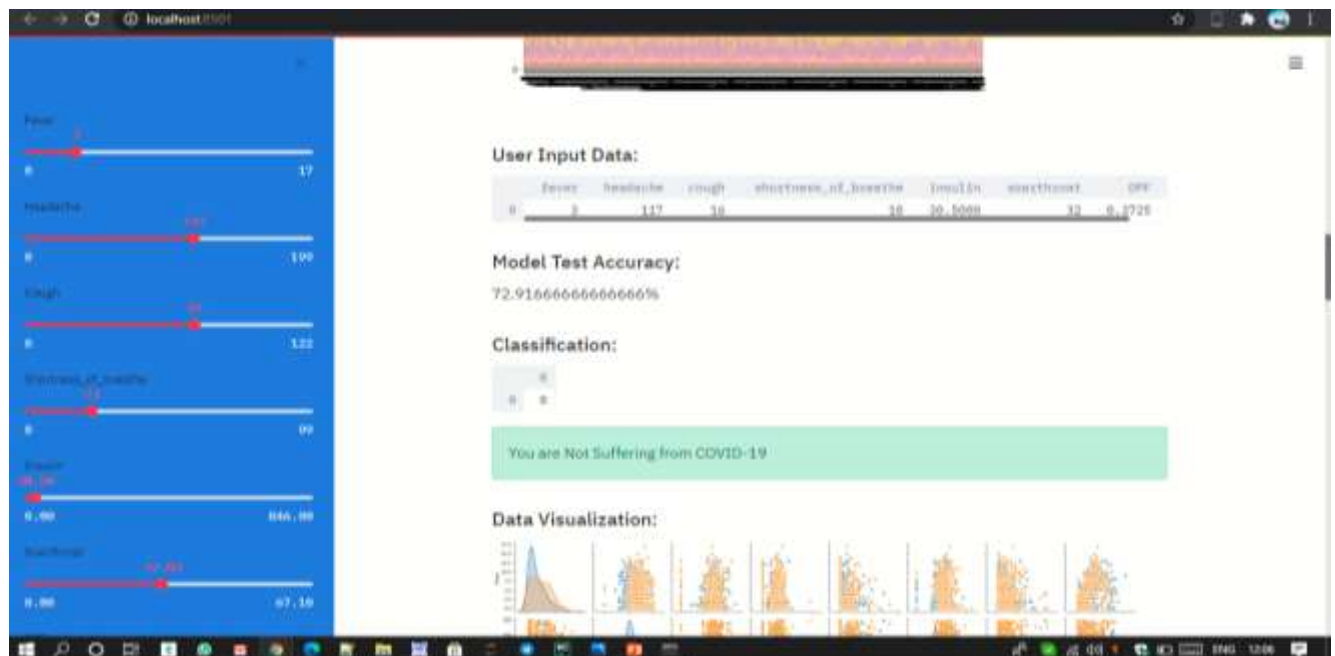
**9. Photos**

**FRONT PAGE DATA DESCRIPTION**

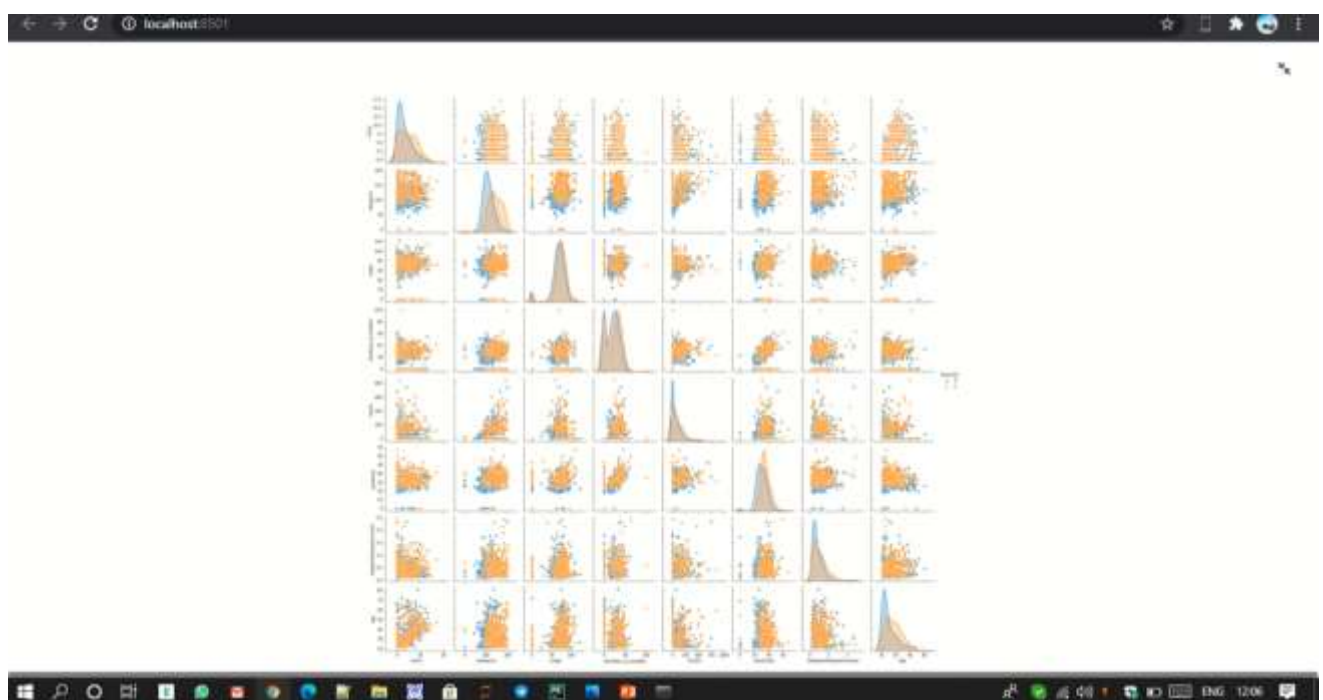




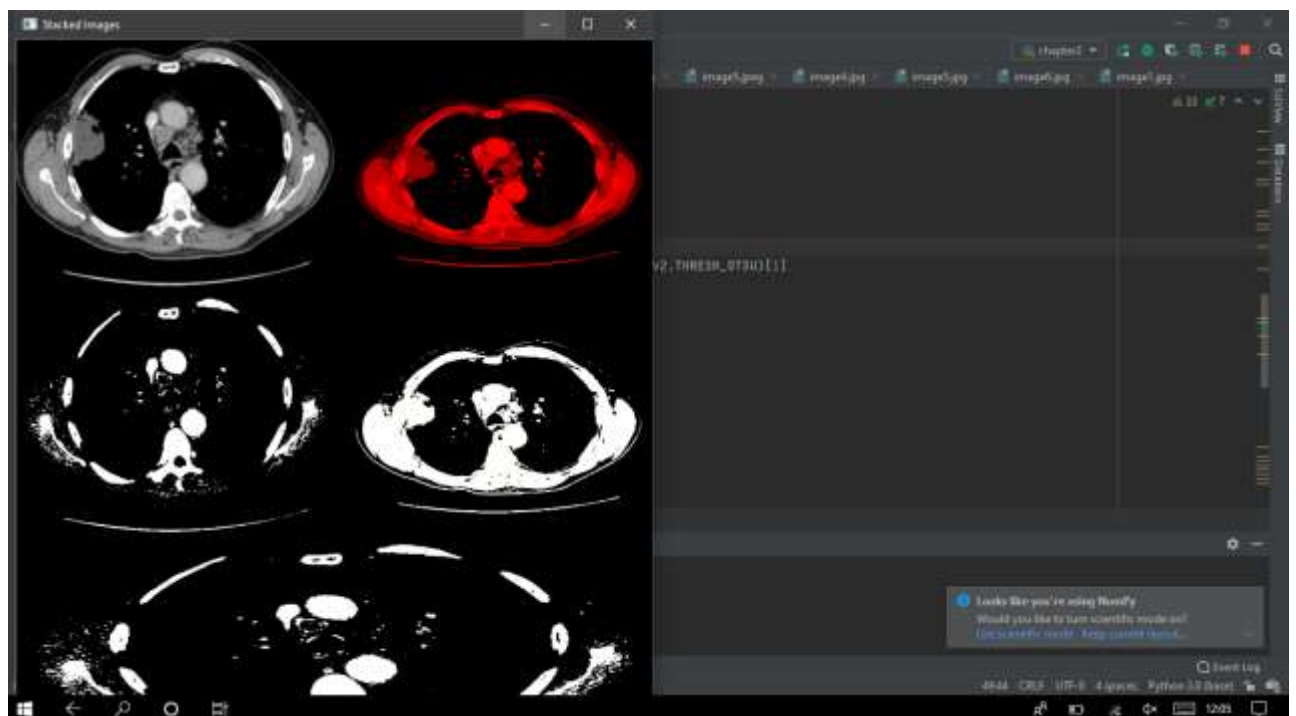
## ACCURACY SCORE AND CLASSIFICATION



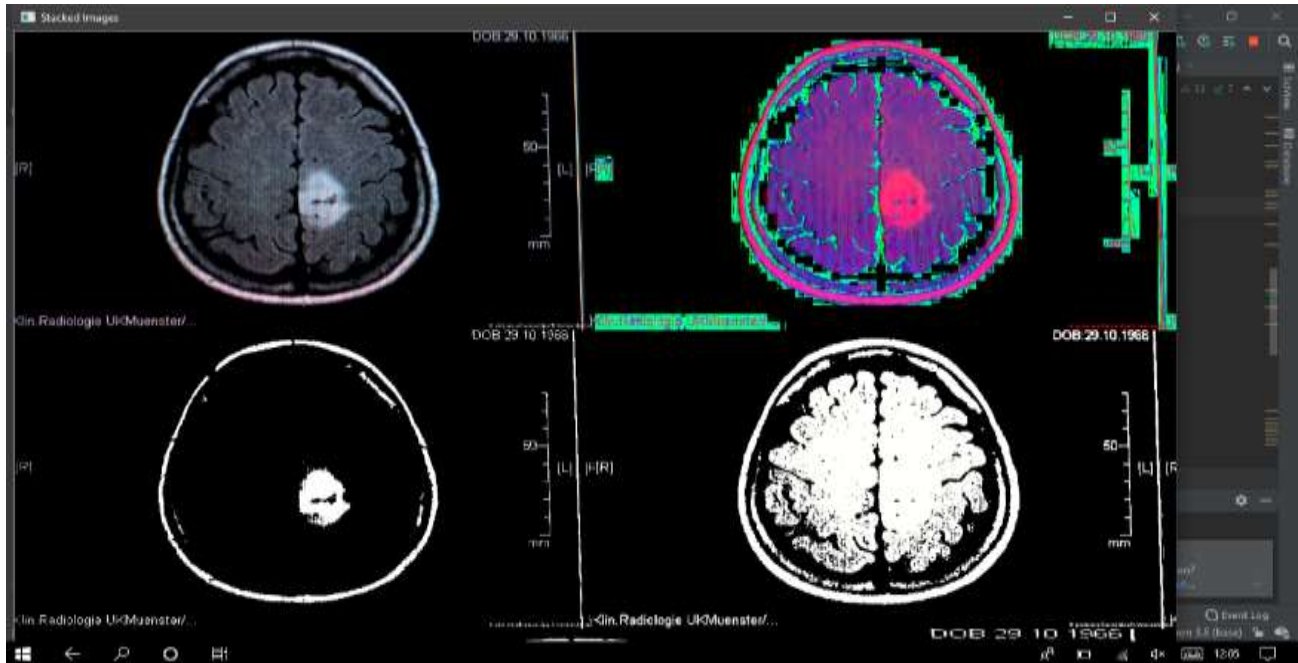
## DATA VISUALIZATION USING PYPLOT FOR USER INPUT







## BRAIN IMAGE PROCESSING USING OPENCV



### 10. Video of Prototype\*:

<https://drive.google.com/file/d/15uz0X1JeJa1exKilotEA6yWGVg-afKu-/view?usp=sharing>

### 11. Feasibility of solution (SMART)

#### 11.1. Specific- Specify the features of Innovation/ Prototype:

**VERSION** WINDOWS 10

**SERVER** STREAMLIT SERVER

**EDITOR** PYTHON 3.9

#### **REQUIREMENTS:**

1. Anaconda should be installed in the particular system
2. Server will be running on the localhost
3. Should run in the anaconda PowerShell prompt in order to run the server
4. The file should save with the extension of .py (i.e) Python
5. Permission for antivirus should be given as ok or off the antivirus



## COMMANDS TO RUN THE SERVER

1. Navigate to the directory in which the file has been saved
2. Stream lit run "filename.py"

It is user friendly and easily accessible to everyone in which every people can use it for free of cost with the above requirements.

### 11.2. Measurable- Mention the approach to convert prototype in to market ready product with milestones:

It is already existing tool in the market in which it has been detected for almost 25 patients which Shows the accuracy score high and the right classification. And also it gives the approximate output for the patients not suffering from corona virus and the less accuracy score

### 11.3. Attainable- Explain how you are going to achieve the commercialization objective with the available resources at your disposal:

The main aim of rising this project is to

1. Consume Time
2. Easily Accessible
3. User Friendly to the environment
4. Available at free of cost

### 11.4. Realistic- what kind of skillset of team and resources required to achieve the goal in specific time period? :

Knowledge in Machine Learning algorithms such as Random Forest Classifier, Decision Tree Algorithms, Logistic Regression, Linear Regression, Support Vector Machine and some of the algorithms used and by image processing and video processing made easier to detect the virus at the faster rate and in faster duration of time and also using Open CV [Open Computer Vision] and some of the library packages in python like numpy, pandas, data man plication are used for data visualization, data prediction and data classification

- |                         |                        |
|-------------------------|------------------------|
| 1. Reading of Data      | : Pandas               |
| 2. Splitting of Data    | : Numpy                |
| 3. Testing and Training | : Data Manu plication  |
| 4. Video Processing     | : Bar Chart Race       |
| 5. Stacked Images       | : Open CV              |
| 6. Image Processing     | : Stream lit in Python |

**11.5. Timeline- Develop a timeline against the milestones for taking prototype to the commercial level/startup stage:**

DATE	TIMELINE
June 2020	Started to do the project
September 2020	Completed Phase 1
October 2020	Completed Phase 2
November 2020	Presented the Full project
January 2021	Implemented in the real time environment for diagnosis and testing of covid-19 patients

**12. Applicability of Solution**

**12.1. Usability: what is the usability of your innovation/ prototype:**

In the month of November the complete part of the project has been deployed and in the month of January 2021 it has been tested for the real time corona patients and for the peoples not suffering from corona virus it gives us the approximate level of accuracy and the classification

**12.2. Scalability: how your innovation/ prototype will be scalable at market level:**

It is not for sales in the market, it is for free of cost to the users or in the hospital

**12.3. Economic sustainability: Explain the potential of innovation/ prototype to become profitable or financially viable:**

This project is deployed using machine learning algorithms but, it can be done more accurately and more informative with the help of deep learning which can give more accurate results and the easy diagnosis of corona virus in X-rays

**12.4. Environment Sustainability: How your innovation/ prototype is environment friendly or Address environmental problems:**

As it is running on the server local host, every peoples around the world can use this simultaneously there is no need of running the server again and again. After diagnosing, the corresponding details will be displayed along with the user input followed by the data visualization and if another person wants to check it can be done by refreshing the page where the previous data will get deleted

**12.5. Is there any Intellectual Property (IP) Component associated with innovation/ prototype? if yes, explain:**

No, there is no any Intellectual property component associated with this innovation

**12.6. Define the potential market size (in terms of INR) and target customers:**

It is free of cost which is easily accessible and user friendly to the environment and there is no Potential market size available to the target customers

### **9 stages of TRL:**

TRL 0 : Idea. Unproven concept, no testing has been performed.

TRL 1 : Basic research. Principles postulated observed but no experimental proof available.

TRL 2 : Technology formulation. Concept and application have been formulated.

TRL 3 : Applied research. First laboratory tests completed; proof of concept.

TRL 4 : Small scale prototype built in a laboratory environment ("ugly" prototype).

TRL 5 : Large scale prototype tested in intended environment.

TRL 6 : Prototype system tested in intended environment close to expected performance. TRL

TRL 7 : Demonstration system operating in operational environment at pre-commercial scale.

TRL 8 : First of a kind commercial system. Manufacturing issues solved.

TRL 9 : Full commercial application, technology available for consumers.