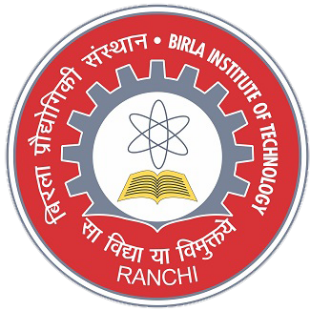
**BIRLA INSTITUTE OF TECHNOLOGY MESRA, JAIPUR CAMPUS**

****

**SYNOPSIS OF PREDICTION OF COVID-19 DISEASE BASED UPON SYMPTOMS (CA656)**

**Submitted to:**

Dr. Piyush Gupta

(Project Guide)

Dr. Madhavi Sinha

(H.O.D Computer Science Department)

**By:**

MEGHAJ KUMAR MALLICK

(MCA/25017/18)

3RD YEAR, VI SEMESTER

1. **INRODUCTION**

This project is based upon the analysis & prediction of the COVID-19(corona virus) disease by their symptoms. This entire process is done by collecting the data from the user. These data will help to indentify whether any person is suffering from COVID-19 or not, which is based upon some predefined standard symptoms. These symptoms are based on the guidelines given by the World Health Organization (WHO) & the Ministry of Health and Welfare, India

For this project we will use various machine learning algorithm to predict the symptoms. In this project we will get to know about

* The dataset contains seven major variables that will bring an impact on whether any person is suffering from corona or not.
* **Country**: List of the countries a person has visited.
* **Age**: Classification of the age group for each person, based on WHO age standard group.
* **Symptoms**: According to WHO there are five major symptoms such as Fever, Tiredness, Difficulty in Breathing, Dry Cough & Sore throat.
* **Other Symptoms:** Other symptoms include Pain, Nasal Conjection etc.
* **Severity**: The level of severity, Mild, Moderate & Severe.
* **Contact**: Whether person came to contact with a COVID-19 patient.
* With all these categorical variables, a combined for each label in the variable will be generated & therefore in total **3,16,800** combinations are created.
* We need two kinds of data in **csv** (comma separated values) format such as **raw** **data & cleaned data**.
* In raw data it contains all possible labels of variables, which is used to generate cleaned data.
* The cleaned data contains all possible from raw data, which can be used for analysis.
* The cleaned data might contain some dummy variables.

1. **PROBLEM STATEMENT**