**ANALYSIS OF COVID-19 BASED UPON SYMPTOMS**

***A synopsis submitted in partial fulfillment of the requirements***

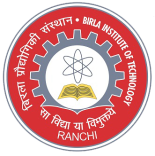
***for the award of the degree of***

**MASTER OF COMPUTER APPLICATIONS**

**(CA 656)**

BY

**Meghaj Kumar Mallick (MCA/25017/18)**



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA**

**JAIPUR CAMPUS, JAIPUR**

**SP-2021**

**INDEX PAGE**

|  |  |  |
| --- | --- | --- |
| SR NO. | TOPIC | PAGE NO. |
| 1 | ABSTRACT | 3 |
| 2 | INTRODUCTION | 3 |
| 3 | MOTIVATION | 4 |
| 4 | OBJECTIVE | 4 |
| 5 | SOFTWARE & HARDWARE REQUIREMENT | 4 |
| 6 | REFERNCES | 5 |

1. **ABSTRACT**

The novel corona virus disease 2019 (COVID-19) pandemic caused by the SARS-CoV-2 continues to pose a critical and urgent threat to global health. This project is based upon the analysis & prediction of the COVID-19 (Corona Virus) disease by their symptoms.

This project describes the application of machine learning. We will use various machine leaning algorithm such as logistic regression, random forest, naive bayes algorithm & neural network.

In this process, we propose a machine-learning model that predicts a positive COVID-19 infection by asking basic questions that are based upon the symptoms of the disease i.e. fever, difficulty in breathing, dry cough etc. We are also going to use the ensemble technique of bagging approach to improve the final result. Finally we can predict the result on a GUI.

1. **INTRODUCTION**

The outbreak of the novel corona virus in early December 2019 in the Hubei province of the People’s Republic of China has spread worldwide. This pandemic continues to challenge medical systems worldwide in many aspects, including sharp increases in demands for hospital beds and critical shortages in medical equipment, while many healthcare workers have themselves been infected.

In this project we will use machine learning approach to identify the symptoms provided by the users. 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.

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.

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. **MOTIVATION**

We all have been affected by the current COVID-19 pandemic. However, the impact of the pandemic and its consequences are felt differently depending on our status as individuals and as members of society.

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 Corona virus is increasing rapidly day by day, it is impossible to test due to the time and cost factors.

Thus, we need to create an application that could analysis this disease & it will help to save life of many persons. The project will focus upon the analysis & prediction of COVID-19 which based upon the symptoms.

1. **OBJECTIVE**

The objective of this project is to easily identify the COVID-19 disease by its symptoms. It will help to reduce the risk of getting affected. By using the machine leaning model we will try to indentify the symptoms of this disease. This will help to analysis the COVID-19 & reduce the risk of getting affected from this disease.

1. **SOFTWARE & HARDWARE REQUIREMENT**

* Package Requirement : Tinker, NumPy, Scikit Learn, Matplot Library
* Platform: Jupyter-Notebook, Anaconda, Python-IDE.
* Operating System : Microsoft Windows 7 or Above
* Processor : Intel Core i3 or above
* RAM : 4 GB or above
* Hard Disk : 250 GB or above

1. **REFERENCES**

[1] Punn, N. S, Sonbhadra, S. K. & Agarwal, S. COVID-19 Epidemic Analysis using Machine Learning and Deep Learning Algorithms. medRxiv, https://doi.org/10.1101/2020.04.08.20057679 (2020).

[2] Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. <https://doi.org/10.1016/S1473-3099> (20)30120-1

[3] Mei, X. et al. Artificial intelligence–enabled rapid diagnosis of patients with

COVID-19. Nat. Med. 26, 1224–1228 (2020)

[4] Feng, C. et al. A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. medRxiv, https://doi.org/10.1101/2020.03.19.20039099 (2020).

[5] Hastie, T., Tibshirani, R. & Friedman, J. In The Elements of Statistical Learning: Data

Mining, Inference, and Prediction (eds. Hastie, T.,Tibshirani, R. & Friedman, J.)337–387 (Springer,2009).