**Navie-Bayes Classifier for Predicting the Novel Corona virus**

**Alternate Title:** Prediction of COVID-19 Using Machine Learning Classification Algorithms

**Aim**:

To detect the COVID-19 using Machine Learning Classification Algorithms

**Abstract:**

The Corona-virus 2019(COVID-19), which first occurs in Wuhan city of China in December 2019, spread quickly around the world and became a plague. Due to the regular increase in cases, the number of COVID-19 test kits available in the hospital is minimal. Since an automated detection system is needed as an alternative diagnosis to prevent COVID-19 from spreading among people. It is necessary to detect the positive cases as early as possible so as to prevent the further spread of this disease. Application of Machine learning techniques coupled with the COVID-19 dataset. Our goal is to enhance the performance of the model by removing unnecessary and insignificant attributes from the dataset and only collecting those that are most informative and useful for the classification task. Thus the main focus of the system is to make use data analytics to predict the presence of the Virus or not.

**Existing System:**

In Existing system, they use several types of test like swab test, nasal aspirate and sputum test. The people have symptoms that match to the Covid symptoms. But they afraid to take the Covid test. So we proposed a system, they can check through online.

**Proposed System:**

The Covid-19 pandemic has consumed the world, with 192 countries affected and global cases nearing 150 million. To overcome the fallback in the existing system we propose a Machine learning based system to overcome the patient from the afraid and make it as user friendly. We have proposed an automatic prediction of Covid-19 using Covid dataset. The final result shows in Web app.

**Modules:**

* Data Collection
* Convolution Neural network algorithm
* Prediction

**Data Collection:**

Data were drawn from a dataset provided via Kaggle. It consists of number of data Collected a dataset from Kaggle, which has Symptoms like Breathing Problem, Fever, Dry Cough, Score throat, Running Nose, Asthma, etc…. The aim of the dataset was to investigate various ways of effectively detecting Covid virus infections.

**Algorithm Implementation:**

The Classification Algorithms to produce the best result in prediction. We are using Navie-Bayes,Random Forest and Logistic Regression to predict the Covid-19 using ML. The clustering of datasets is done on the basis of the variables and criteria of Decision Tree (DT) features. On an analysis conducted within various algorithms, the Random Forest was found to provide highest efficiency. Then, the classifiers are applied to each clustered dataset in order to estimate its performance. The best performing models are identified from the above results based on their low rate of error.

* Logistic Regression
* Random Forest Classifier
* Navie-Bayes

**Prediction:**

Preprocessed data are trained by ML algorithm and input given by the user, goes to the trained dataset. After prediction the predict value Shown as a whether the person affected by covid-19 or not. And user gets output via Web app.

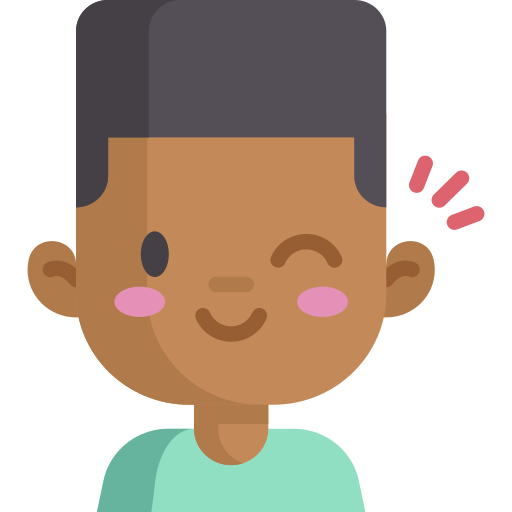
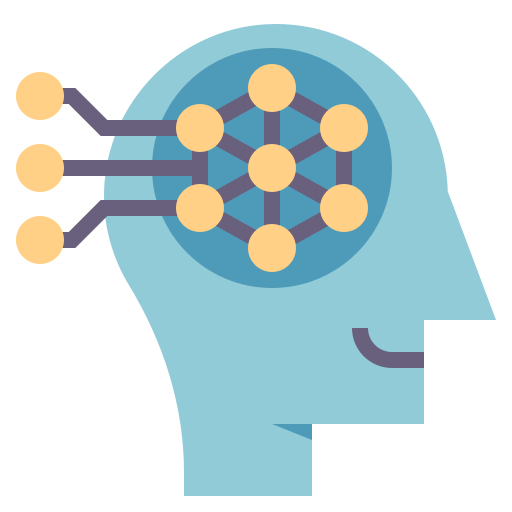
**Software Requirements:**

* Operating System : Windows 10 (64 bit)
* Software : Python 3.7
* Tools : Anaconda (Jupyter Note Book IDE)

**Hardware Requirements:**

* Hard Disk : 500GB and Above
* RAM : 4GB and Above
* Processor : I3 and Above

**Architecture Diagram:**



Dataset Collection

Preprocessing

Covid-19 Detection

Machine Learning

User

Web app

Symptoms Matched

Symptoms Matched

Symptoms Unmatched