**REFLECTION REPORT**

**MACHINE-LEARNING APPLICATION ON COVID-19 DATA SET**

**Group 6**

Naragam yamini chandu

Tadikonda chittaranjan

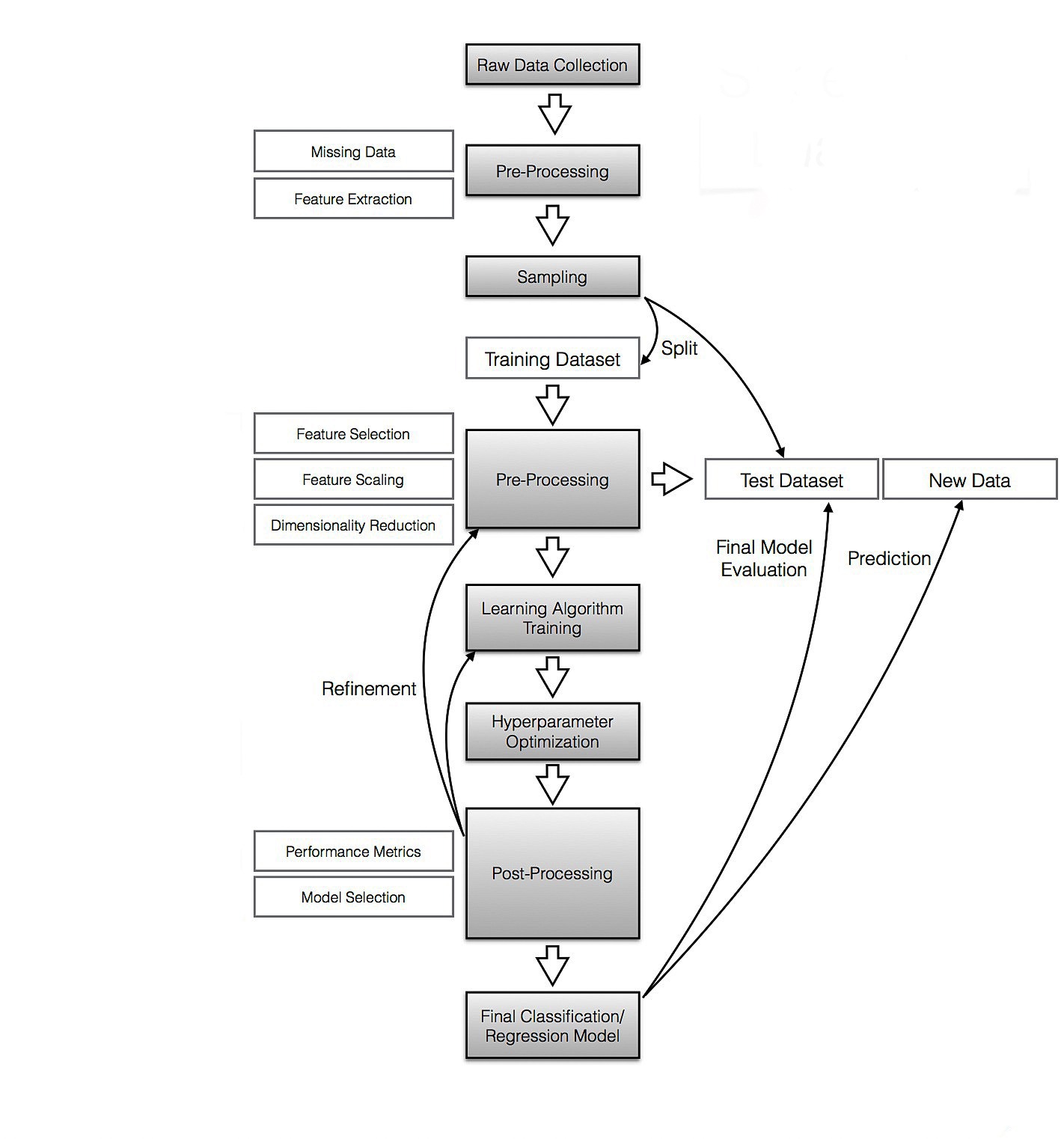
G S V M Pattabhi ram

Medisetti Venkata sudhindra

**INTRODUCTION**

On January 30, 2020, and March 11, 2020, respectively, the World Health Organization (WHO) formally declared the SARS-CoV-2 outbreak a Public Health Emergency of International Concern. WHO recommended nations use stringent social segregation and quarantine measures to protect the public from the virus spread. Covid-19 has severely affected the lives of people around the world it is like the world has been stopped for like a short period due to the lockdown and rise of covid cases around the various places on this planet. Here in this project, we can able to see how many humans were affected with covid 19 virus and perform analysis to check how males are affected, and females. We plot the analysis charts for the cases recorded and try to get a sense of the outcome like which section of people are affected worse and their patterns for the spread of the covid-19. Along with we are trying to find what is the number of people reported to have been admitted and among which are trying to segregate(divide) into different categories and get an understanding of how hospitalization affects people’s health conditions. By finding these results, we can predict and analyze pandemic affects the world's health condition and population.

**FLOWCHART**

****

**SUMMARY**

By using Pandas, we analyzed, and visualized the open covid-19 raw data. The Pandas package used is indeed a very powerful Python package in terms of extracting, grouping, sorting, analyzing, and plotting the data. From bring the raw unstructured data, merging them together and brought into one dataset.

Later brought with research question where the motivation of the project is build. In building the prediction model it is important to visualize the data.

The following are the outcomes from open covid-19 data. In this project we have taken 3 types of datasets based on years and in each year we found how many cases were detected for humans, separated based on male and female. We did visualization based on gender. We segregated by using newly hospitalized patients and in that, we did for tested, conformed, and recovered patients and we used analysis, and visualization to optimize the data. we did analysis charts for the cases.

**CONCLUSION**

In this work, we have done analysis based on machine learning in covid-19 open raw data. By measuring the accuracy of different algorithms, we found that the most suitable algorithm for predicting drugs based on various conditions is Random Forest. We believe that employing more sophisticated features and applying more powerfulmachine learning models, deep learning approaches can help to enhance the performance of the system.

Thank you,

Group 6