**COVID 19 CASES ANALYSIS**

PHASE2-DATA ANALYTICS WITH COGNOS: GROUP2

**INTRODUCTION:**

COVID-19 Cases Analysis is a critical examination of the global pandemic's data, trends, and impact. It involves studying infection rates, mortality statistics, and vaccination progress to gain insights into the virus's spread and public health responses. This analysis helps inform policymakers, healthcare professionals, and the public on effective measures to combat the pandemic.

By analyzing COVID-19 cases, we can better understand and respond to the ongoing crisis.In this phase we are going to explain about design and ideology that are going to present to solve this problem.

**Dataset link:** [**https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases**](https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases)

To this problem this dataset is given to us so by using this dataset we are going to solve our problem.

In phase1 we have defined certain steps to solve the problem step by step. Now we are going to explain which methodology we are going to use to solve this problem in each step.

**Clearly define the problem:**

The problem statement for Tamil Nadu COVID-19 data analysis is to comprehensively examine and interpret the COVID-19 data specific to the state of Tamil Nadu, including infection rates, vaccination progress, healthcare infrastructure utilization, and the impact on various demographics. This analysis aims to identify trends, hotspots, and potential areas of concern to inform public health policies, resource allocation, and containment strategies, ultimately contributing to the effective management and mitigation of the COVID-19 pandemic within the state.

**Data collection:**

The dataset is already given for us:

**Dataset link:**

[**https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases**](https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases)

**Preparing of the data:**

First we have to understand what was the data we are going to analyze for this we have to clean and process the data by using suitable techniques like **dropping the null values, data types, remove the duplicate values, visualize the missing values drop the duplicates, by using the suitable functions like drop, is null etc.…**

**Exploratory data analysis:**

This was the most important step in this project so we have to represent our data in the **understandable visualization tools like pie chart, bar graph, histogram** to represent the relation between the two attributes in the given column.

**Feature selection:**

In this step we are going to explain about the features or attributes that are going to be selected in the dataset and we have to **represent the relationship between the data visualization.**

**Model selection:**

We are going to solve the problem by using various algorithms support vector machine, linear regression, logistic regression, naive Bayes, decision trees (random forest and ETC), K-nearest neighbor, and neural networks (multilayer perceptron)**,** we can solve by using any of this model and choose the accurate model.

**Model training and validation:**

We are using various model validation methods like **label encoder, ascending values, dis plot, standardization and scaler.**

**Model evaluation:**

We divide our data set by using the **training and testing, predicting the values by using the 2 dimensional planes as a result the accuracy model** will tell us which algorithm we use.

**Result representation:**

This project can be represented in many ways like using the ROC curve, confusion matrix etc… We are going to use the **confusion matrix** to solve this.

**Reporting and visualization:**

To inform stakeholders about the **churn prediction and insights provide the periodical reports and dashboards** we are going to do this in a confusion matrix and accuracy of the data.

**Business action:**

Businesses have adapted by enabling remote work, improving sanitation, using **digital tools, diversifying supply chains, boosting online presence, and contributing to pandemic relief efforts**, showcasing resilience amid COVID-19 challenges.