**Descriptive Analytics for heart failure**

**Abstract:**

The goal of this project is to understand which factors are most Causing heart failure and how to get prevention from it.

Heart disease is one of the major causes of mortality in the world today. Prediction of cardiovascular disease is a critical challenge in the field of clinical data analysis. With the advanced development in machine learning (ML), artificial intelligence (AI) and data science has been shown to be effective in assisting in decision making and predictions from the large quantity of data produced by the healthcare industry. Initially ML was used to find degree of heart failure, but also used to identify significant features that affects the heart disease by using correlation techniques. There are many features/factors that lead to heart disease like age, blood pressure, sodium creatinine, ejection fraction etc. In this paper we propose a method to finding important features by applying machine learning techniques. The work is to design and develop prediction of heart disease by feature ranking machine learning.

**Design:**

Heart failure is a common event caused by cardiovascular diseases and this dataset contains features that can be used to predict a possible heart disease.

People with cardiovascular disease or who are at high cardiovascular risk (due to the presence of one or more risk factors such as hypertension, diabetes, hyperlipidemia or already established disease) need early detection and management wherein a machine learning model can be of great help.

**Data:**

This dataset was created by combining different datasets already available independently but not combined before. In this dataset, 5 heart datasets are combined over 12 common features.

Rows: 999

Columns: 12 columns which are Age – Sex – Chest Pain Type – Resting BP – Cholesterol - Fasting BS – Resting ECG – Max HR – Exercise Angina – Old peak – ST Slope - heart disease.

**Algorithms:**

Removing the outlier in the cholesterol column by replace it with the mean because it was normal distribution.

Ranking the cholesterol by condition and create new column which is cholesterol type.

Calculate the patient how have the height cholesterol and have heart disease by mask function.

Calculate the patient how older than 40 and have heart disease by mask function.

Calculate the patient how have heights blood pressure and have heart disease by mask function.

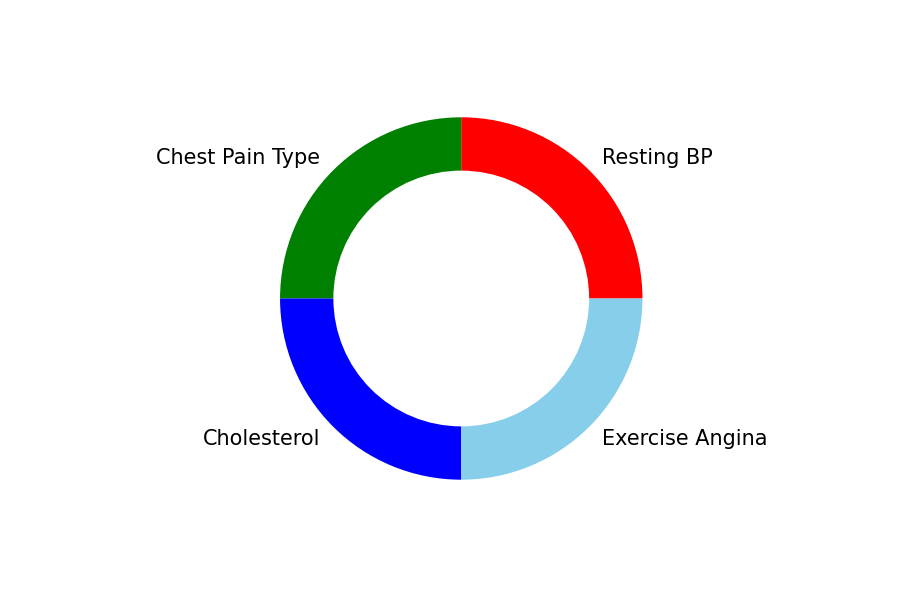
**Tools:**

The technologist that I used are Python, Jupyter Notebook.

The libraries are pandas, matplotlib.

**Communication:**

These are the columns that I will focus on to predict if there are a relation to cause a heart failure.

****