DEPARTMENT OF COMPUTER SCIENCE

MACHINE LEARNING

HEART DISEASE PREDICTION **GROUP_2**

TEAM MEMBERS		
NAME	ID	CONTRIBUTION %
KAMILE SEIDU	10822436	25%
DANIEL ANSAH DELADEM	10712655	20%
FREDA OWUSUAA BAAH	10808851	10%
HELENA AMISSAH	10821309	20%
AKUA OSEI-NKWANTABISA		10%
BERTHA ABENA JONES	10844150	15%

GITHUB REPOSITORY LINK:

https://github.com/KamileSeidu/Group_2_Heart_Disease_Prediction

INTRODUCTON

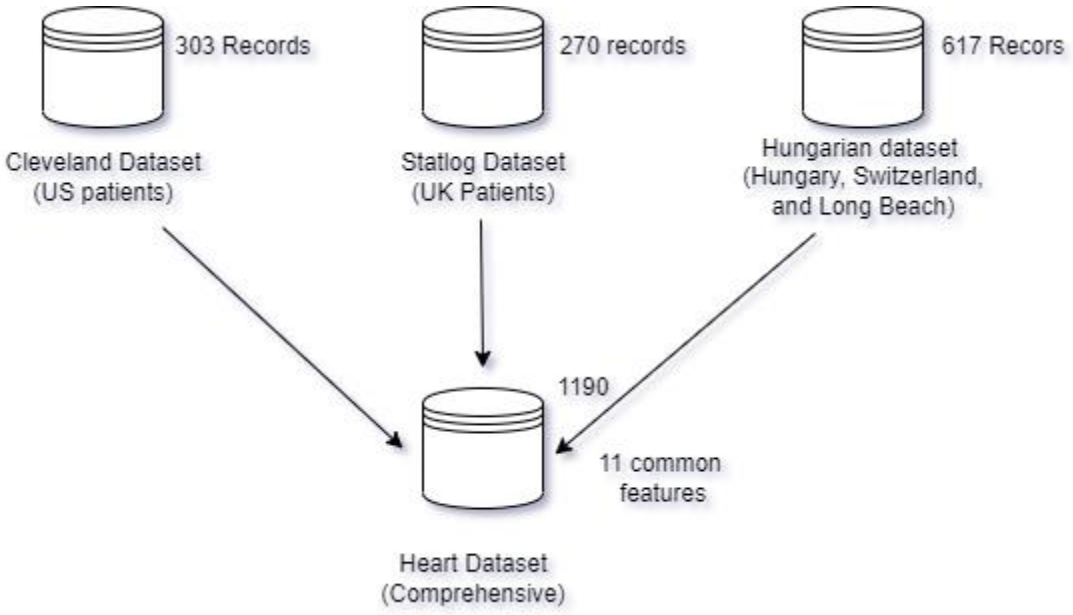
The term "heart disease" refers to a variety of diseases and ailments that affect the heart and circulatory system. Cardiovascular illnesses are another name for them. It is one of the leading causes of disability all around the world. Because the heart is one of the most important organs in the body, illnesses of the heart can impact other organs and parts of the body. Heart disorders come in a variety of shapes and sizes. The most frequent ones involve coronary artery narrowing or blockage, heart valve malfunction, heart enlargement, and a variety of other problems that contribute to heart failure and heart attacks.

According to the Centers for Disease Control and Prevention (CDC), heart disease is one of the major causes of mortality in the World. High blood pressure, high cholesterol, and smoking are three important risk factors for heart disease that affect over half of all Americans (47 percent). Diabetic status, obesity (high BMI), lack of physical exercise, and excessive alcohol use are all important indicators. In healthcare, detecting and avoiding the variables that have the greatest influence on heart disease is critical. As a result of computational advancements, machine learning methods may be used to find "patterns" in data that can be used to forecast a patient's status.

OBJECTIVES OF THE NOTEBOOK

This notebook aims at predicting heart diseases base on the data set given. We will be comparing models to see which one performs better than the other. We will be using some common Python libraries, such as pandas, numpy, matplotlib and Seaborn.

DATASET CURATION

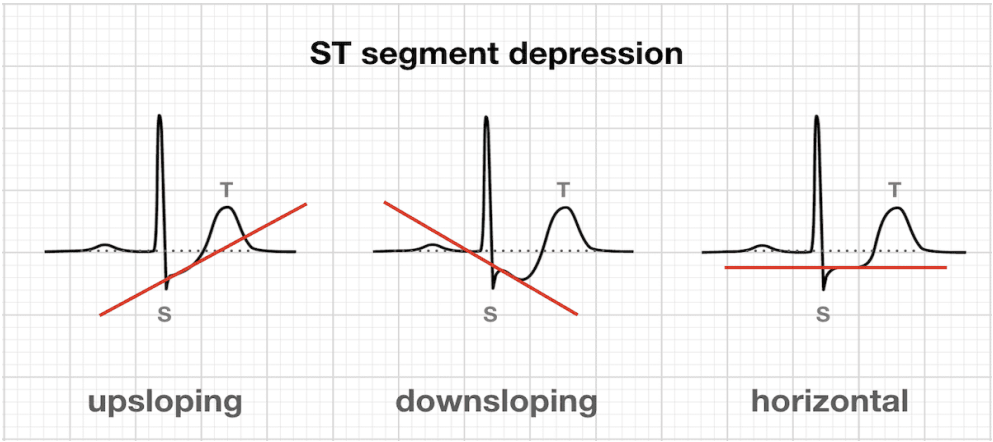


EXPLANATION OF THE VARIABLES OF THE DATASET

S. NO	Feature	Description	Data type
1.	Age	Patients Age in years	Numeric
2.	Sex	Gender of Patient (Male - 1, Female - 0)	Normal
3.	Chest pain type	Type	Nominal
4.	Resting BP	Level of blood pressure at resting mode in mm/HG	Numerical
5.	Cholesterol	Serum cholesterol in mg/dl	Numeric
6.	Fasting blood sugar	Blood sugar levels on	Nominal

		fasting > 120 mg/dl represents as 1 in case of true and 0 as false	
7.	Resting ECG	Result of electrocardiogram while at rest are represented in 3 distinct values	Normal, Wave, Nominal
8.	Max Heart rate	Maximum heart rate achieved	Numeric
9.	Exercise angina	Angina induced by exercise 0 depicting NO 1 depicting Yes	Nominal
10.	Old peak	Exercise induced ST-depression in comparison with the state of rest	Numeric
11.	ST slope	segment measured in terms of slope during peak exercise	Normal, Flat, Nominal
TARGET VARIABLE			
12.	Target	It is the target variable which we have to predict 1 means patient is suffering from heart risk and 0 means patient is normal.	Heart Risk, Normal

ST SEGMENT DEPRESSION



GITHUB LINK:

https://github.com/KamileSeidu/Group_2_Heart_Disease_Prediction