# Name: Saniya Mushtaq

(Student ID: STU-ML-251-486)

Name of model:

Heart\_Disease\_using\_SVM.ipynb

# Heart Disease Prediction using Machine Learning

## 1. Introduction

Heart disease is one of the leading causes of death worldwide. Early prediction can help in timely intervention and better treatment. In this project, we apply machine learning techniques to predict the likelihood of heart disease based on patient attributes.

## 2. Dataset Description

We used a simplified heart disease dataset based on the UCI repository. The dataset includes features such as age, sex, chest pain type, blood pressure, cholesterol level, blood sugar, ECG results, maximum heart rate, exercise-induced angina, ST depression, slope, number of major vessels, thalassemia, and a target variable indicating disease presence.

## 3. Data Preprocessing

Initial data cleaning involved removing duplicate records and handling missing values. The dataset was also scaled using Standard Scaler to normalize feature ranges. No categorical encoding was needed as all variables were already numeric.

## 4. Exploratory Data Analysis (EDA)

EDA was performed using visualizations such as heat maps and count plots. The heat map revealed correlation among features, helping identify those with significant impact on the target. Count plots showed class distribution in the target variable.

## 5. Feature Selection

Features most correlated with the target were selected using the correlation matrix. The top 5 features were chosen to train the models, reducing noise and improving performance.

## 6. Model Building and Evaluation

Two models were trained and evaluated: Logistic Regression and Random Forest. Evaluation metrics included Accuracy, Precision, Recall, F1-Score, and the Confusion Matrix. Both models showed good performance, with Random Forest slightly outperforming Logistic Regression.

## 7. Bonus: Hyper parameter Tuning

Although optional, Grid Search CV was prepared for hyper parameter tuning on Random Forest. This technique allows systematic testing of parameters to find the optimal model settings.

## 8. Conclusion

This project demonstrates how machine learning can be used for predicting heart disease. With proper data preprocessing, EDA, feature selection, and model tuning, ML models can aid in early detection and diagnosis, potentially saving lives.