

# Cardiovascular Diseases Risk Prediction

## Problem Statement :

- Heart disease continues to be a significant global health concern, impacting millions of lives and posing a substantial burden on healthcare systems. Early detection and accurate prediction of heart disease play a crucial role in managing and preventing its onset, as well as improving patient outcomes. Over the years, advancements in medical research, technology, and data analysis have significantly enhanced our understanding of heart disease prediction. This brief introduction aims to provide an overview of the progress made in this field, highlighting key methodologies and approaches utilized in predicting heart disease. By exploring the various risk factors, diagnostic tools, and emerging trends, we can gain valuable insights into the evolving landscape of heart disease prediction, paving the way for more proactive and personalized healthcare strategies.

## Dataset Attributes :

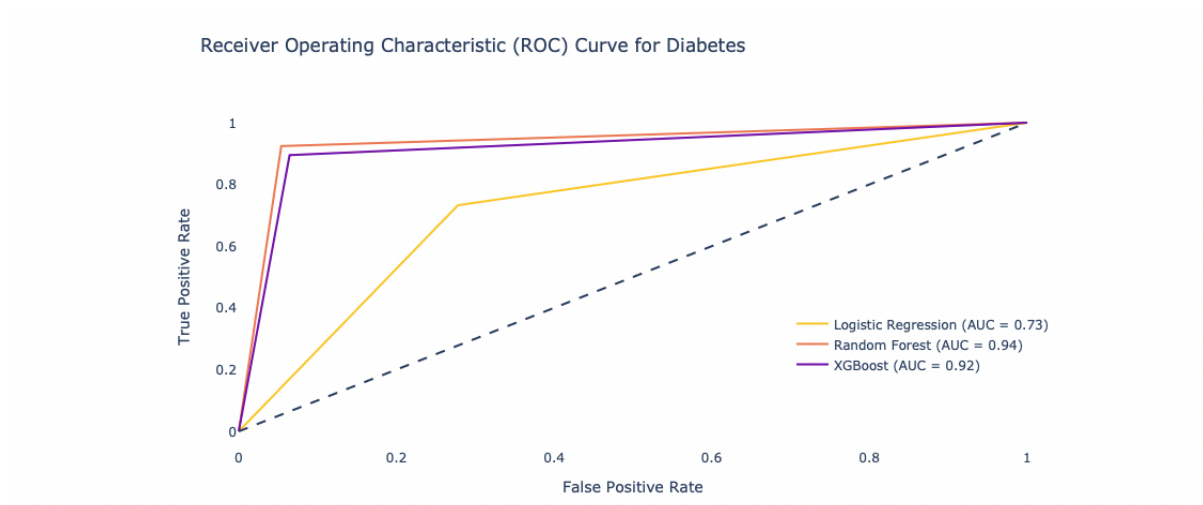
- General\_Health
- Checkup
- Exercise
- Heart\_Disease
- Skin\_Cancer
- Other\_Cancer
- Depression
- Diabetes
- Arthritis
- Sex
- Age\_Category
- Height\_(cm)
- Weight\_(kg)
- BMI
- Smoking\_History
- Alcohol\_Consumption
- Fruit\_Consumption
- Green\_Vegetables\_Consumption
- FriedPotato\_Consumption

## PROJECT OBJECTIVES:

- Analyze the data
- Extract the hidden trends and patterns
- Group the reviewers based upon health conditions and risk
- Analyze Each clusters and find out the nature for better medical recommendations
- Examine which Machine Learning Algorithm is most suitable for the dataset

## Work Done:

- ☐ Data Analysis completed
- ☐ Trends and patterns with varying columns of the dataset
- ☐ Feature engineering, demographic analysis and correlation analysis done
- ☐ Data Pre-processing:
  - One-hot encoding
  - Label encoding
  - Training and testing on data
  - Feature scaling
- ☐ Implemented Logistic regression, random forest classifier and XGBoost classifier



**Conflicts:** None so far apart from dummy data creation for EDA and deploying the project.

**Next Steps:** Deployment of the project to create a user-friendly interface for CVD risk assessment.

## GITHUB link:

[https://github.com/bhavyakalra/Cardiovascular\\_prediction/blob/main/MVP\\_kalra\\_bhavya.ipynb](https://github.com/bhavyakalra/Cardiovascular_prediction/blob/main/MVP_kalra_bhavya.ipynb)