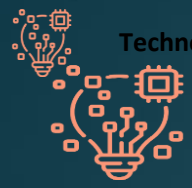
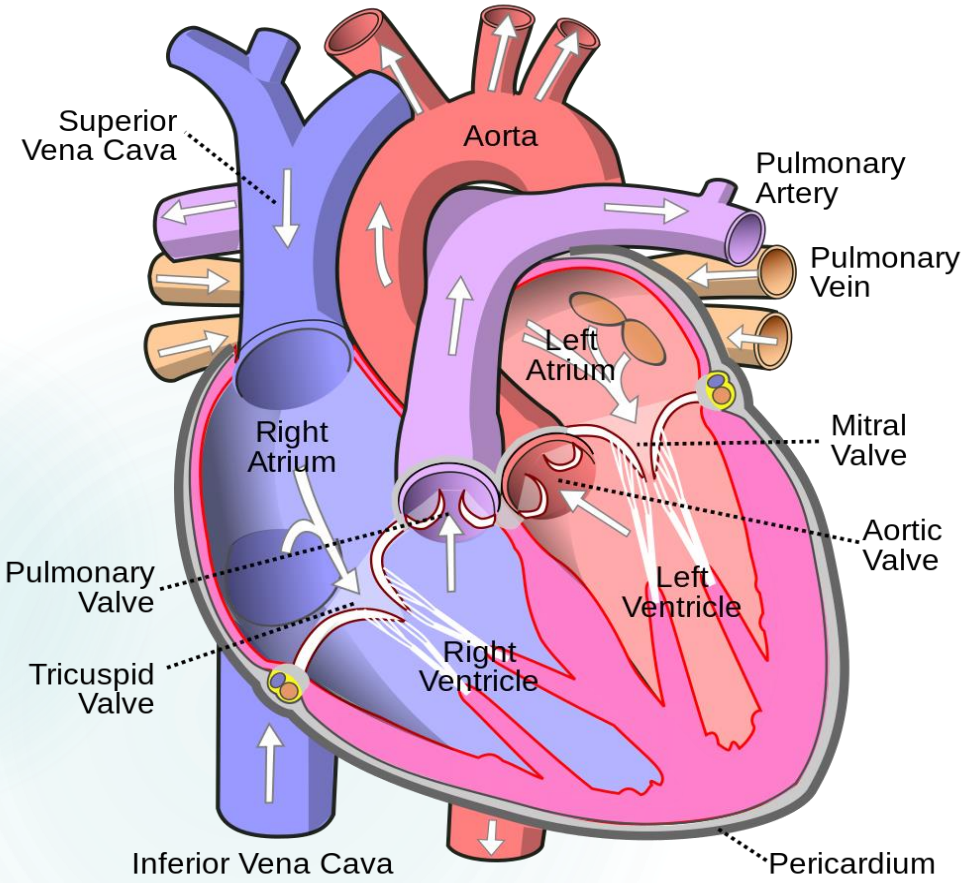


HEALTH MATTERS

HEART ANALYSIS

CORONARY HEART DISEASE ANALYSIS

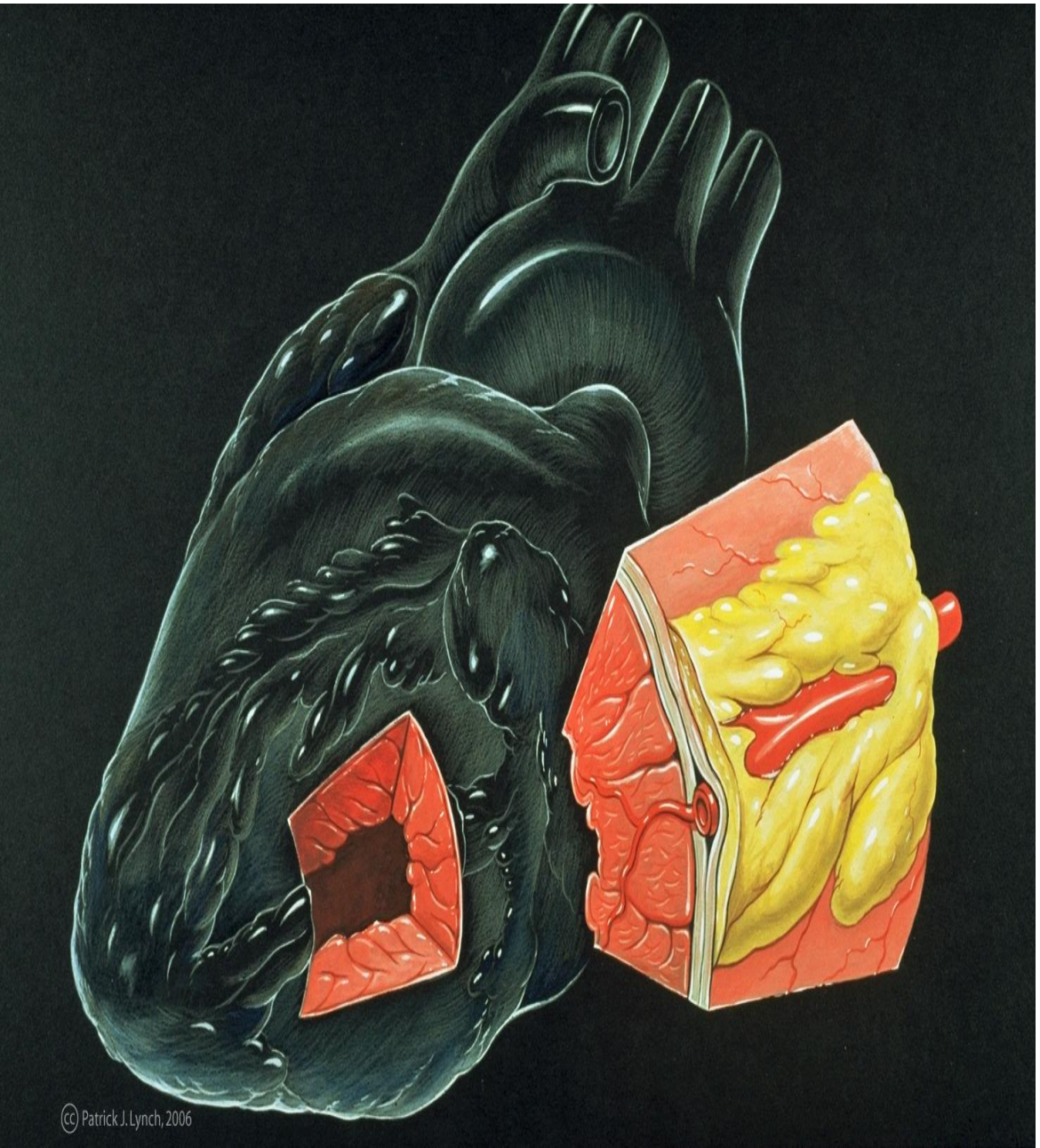


Introduction

Cardiovascular (heart) diseases are a leading cause of death globally, taking an estimated 17.9 million lives each year and imposing a significant burden on public health systems (WHO, 2023). Cardiometabolic, behavioral, environmental, and social risk factors are major drivers of heart diseases.
According to WHO, timely and accurate prediction of heart diseases likelihood can play a crucial role in improving patient outcomes and optimizing healthcare resources. There are four main types of cardiovascular diseases i.e.,: coronary heart disease, stroke, peripheral arterial disease and aortic disease (NHS, 2023).

Problem statement

The project aims to tackle the global rise in heart diseases by focusing on early detection in primary care. Unhealthy behaviors like poor diet, inactivity, tobacco use, and excessive alcohol contribute to elevated risks such as high blood pressure and obesity. With 4 out of 5 heart-related deaths happening prematurely, especially impacting those under 70, urgent intervention is crucial. Positive changes in behavior show potential in lowering risks. Using key indicators like age and cholesterol levels from a dataset, our goal is to build models that identify individuals at higher risk, allowing for timely interventions and contributing to reducing heart disease worldwide



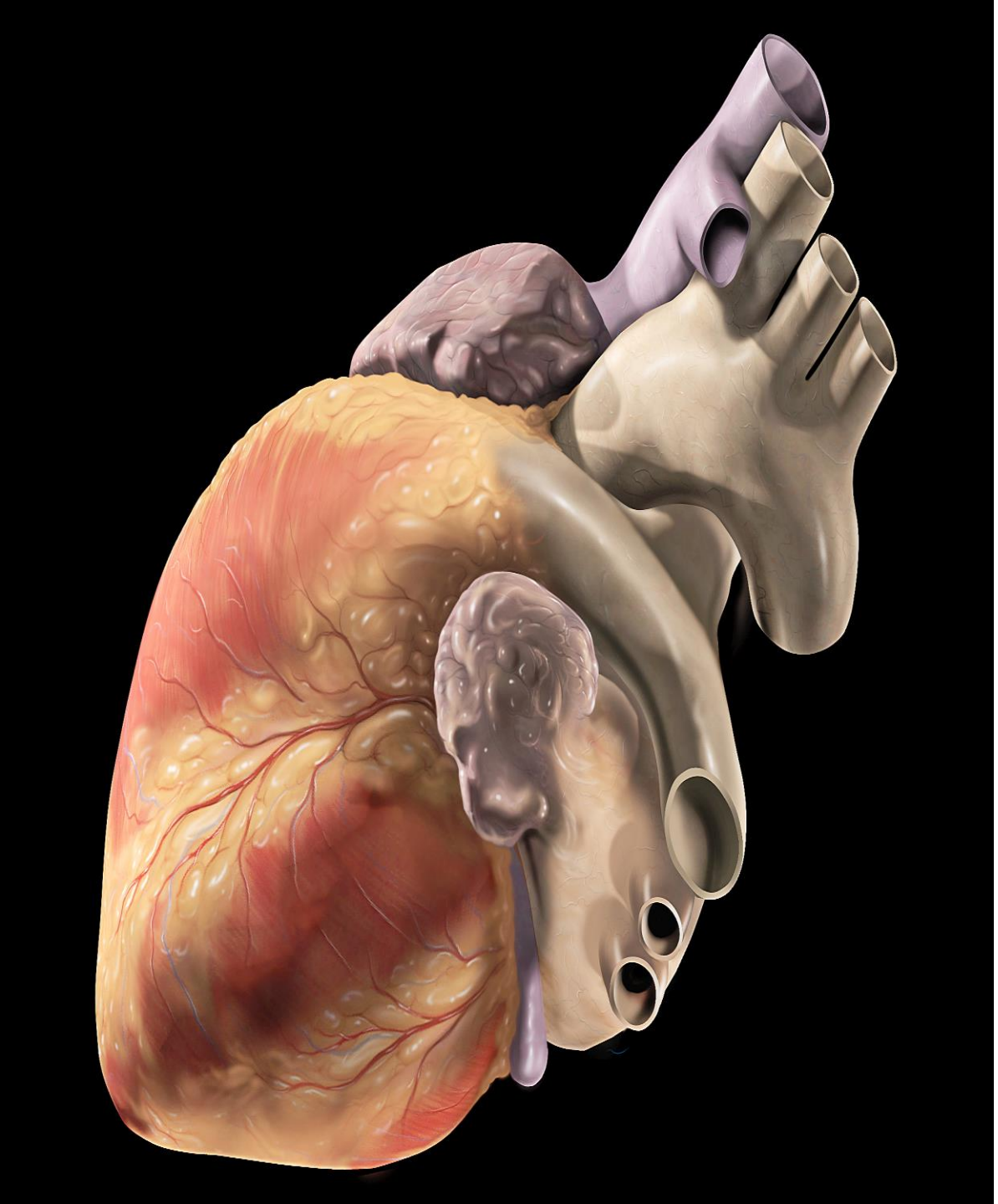
Project Objectives

- **Comprehensive Prediction System:**
  Build a robust CHD prediction system aiding early intervention.

- **Pattern Identification:**
  Uncover patterns and risk factors, enabling targeted preventive measures.

- **Healthcare Assistance:**
  Assist healthcare professionals in early detection and proactive CHD management
  .

- **Machine Learning Models:**
  Create predictive models for a 10-year CHD risk, aiding prioritized preventive interventions.



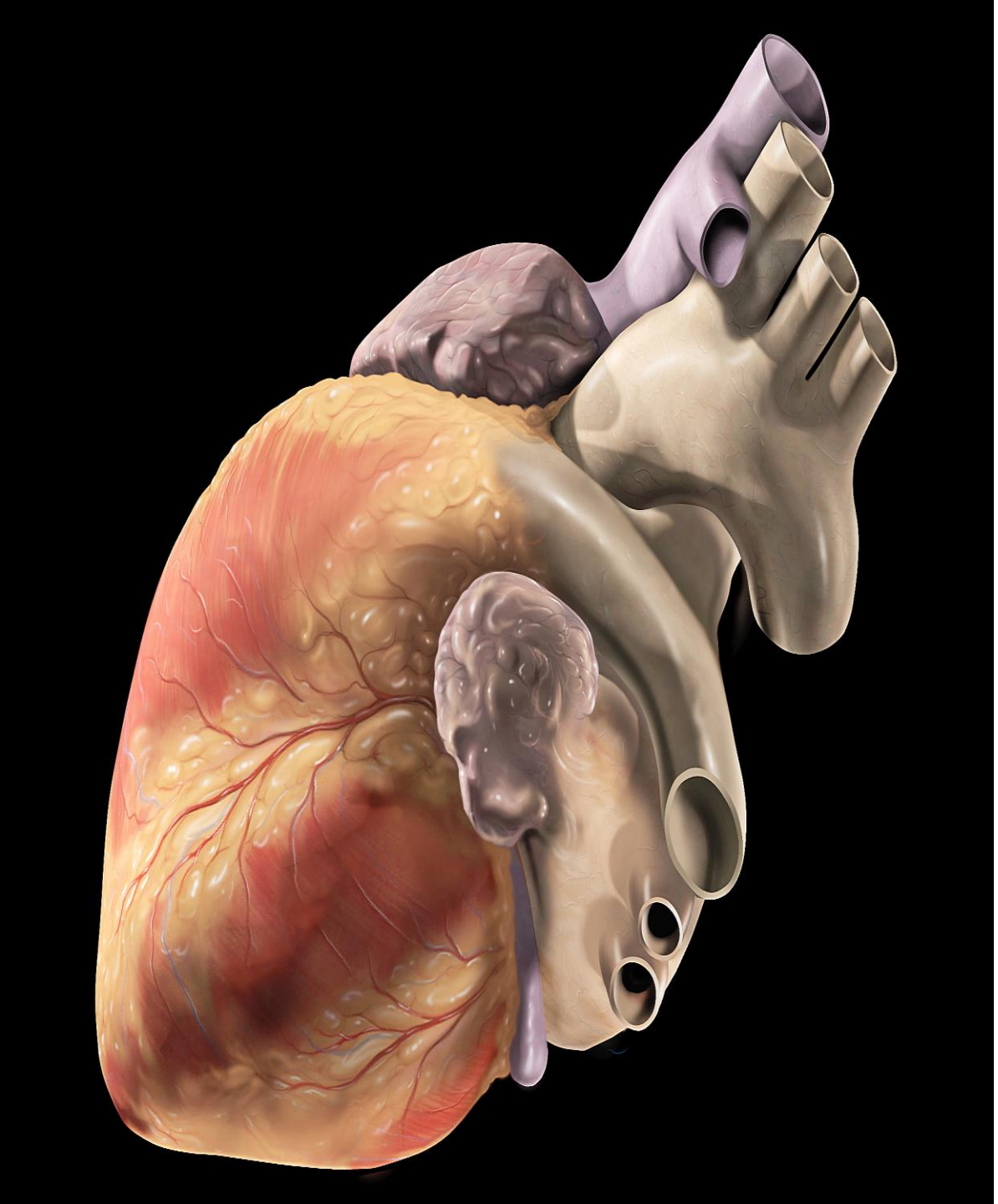
Data understanding

The dataset was sourced from Kaggle datasets.

It contains 15 features used for predicting the 10-year risk of CHD (binary: "1" - Yes, "0" - No).

The main columns used for prediction are:

- Age
- Life style factor
- medical indicators
- BMI



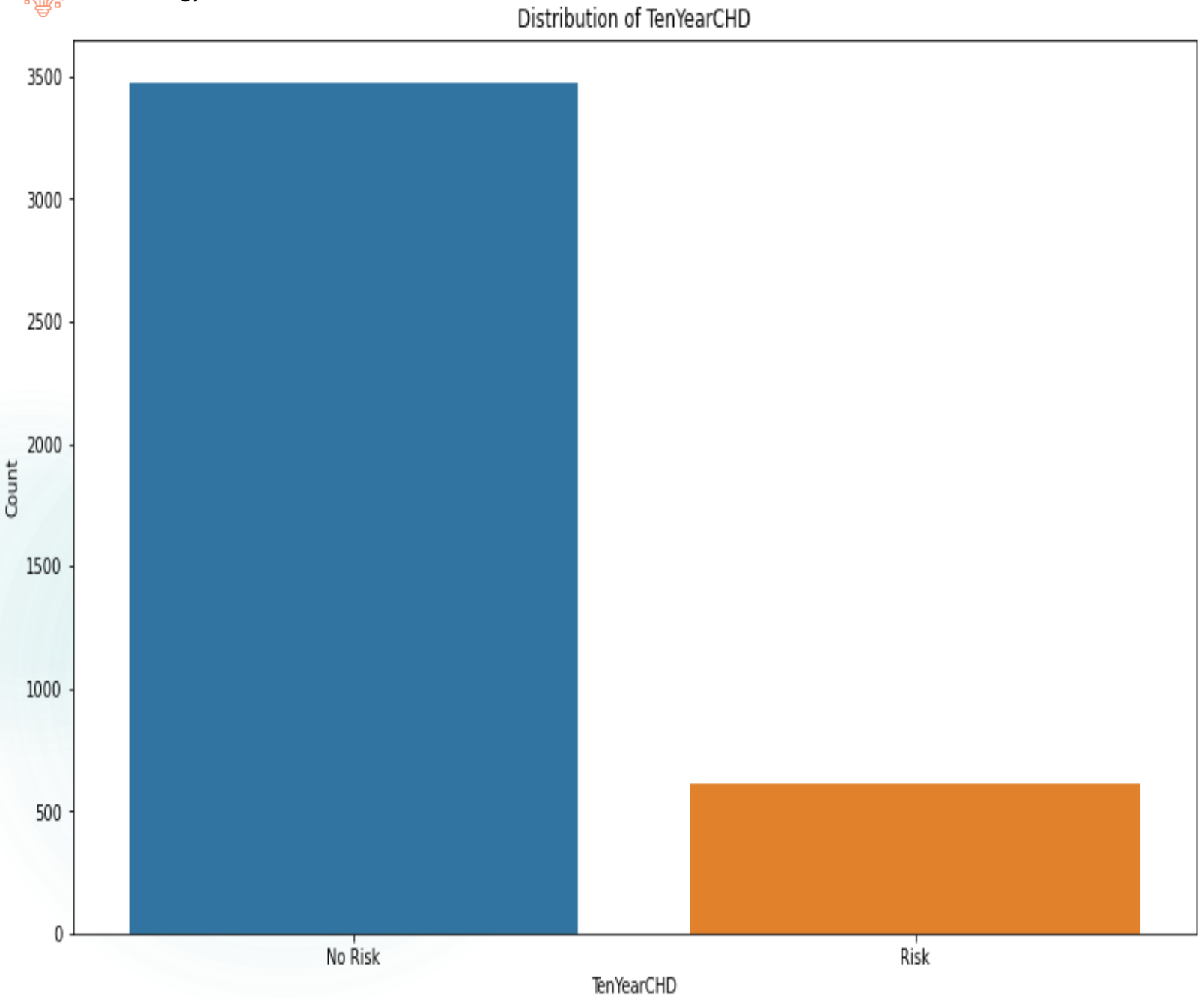
Data preprocessing

- **Handling Missing Values:**
  Dropped columns with missing values < 2%.
  Imputed glucose and education columns with mean and mode, respectively.

- **Outlier Handling:**
  Applied Winsorization to address outliers in selected columns.

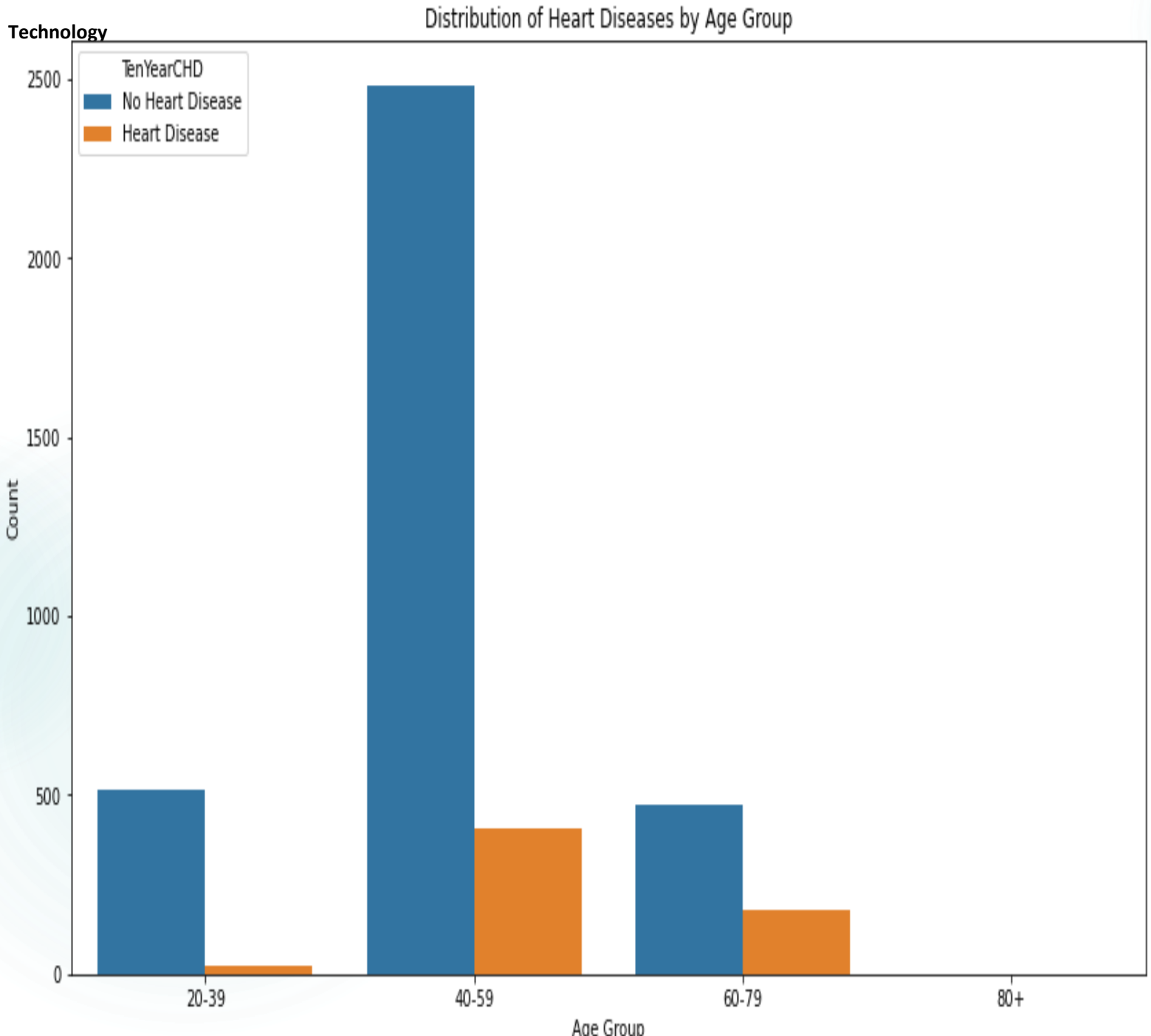
- **Feature Engineering:**
  Created BMI categories and age groups for additional insights.

Exploratory Data Analysis

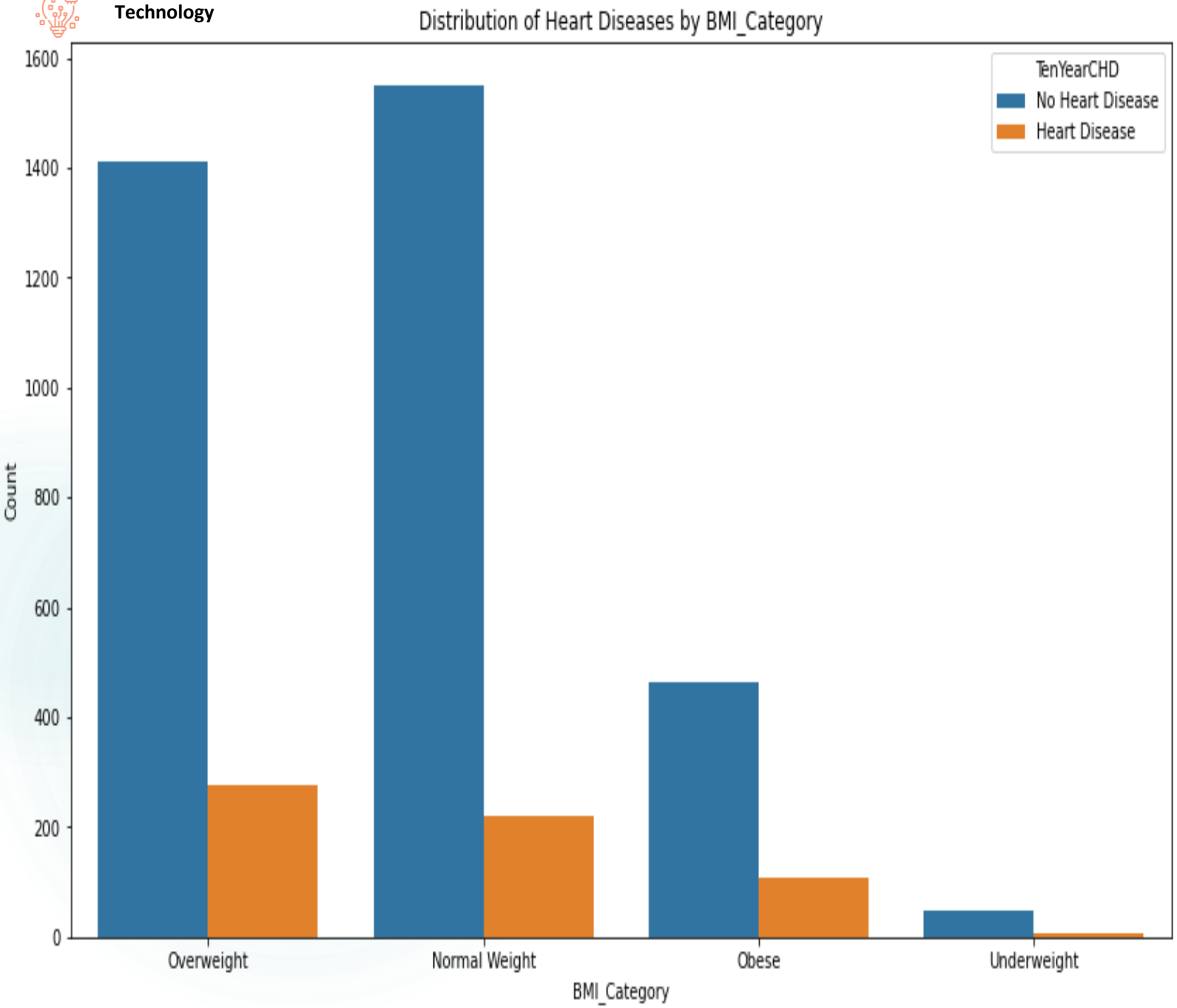


Distribution of Target variable

Technology

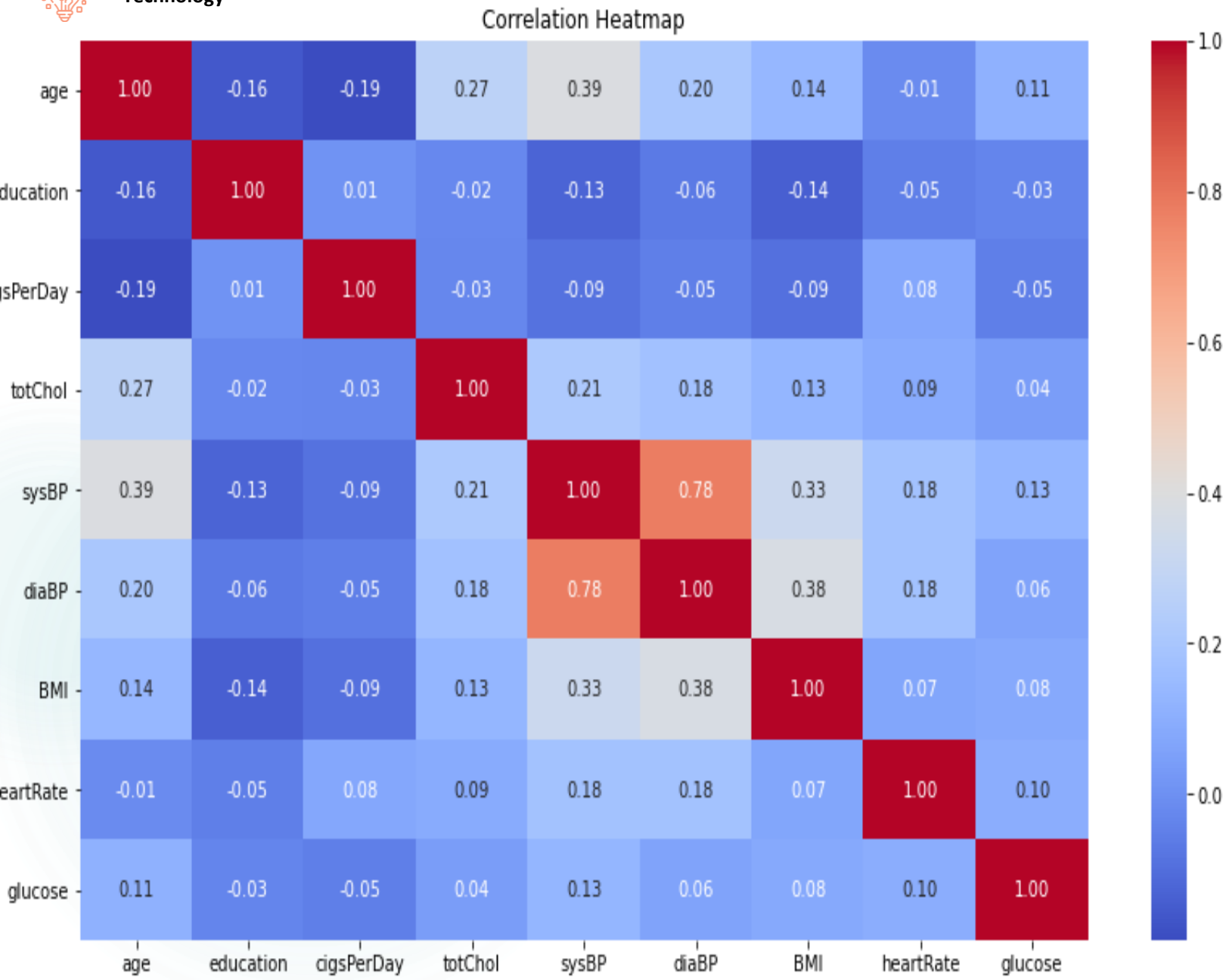


Distribution of Disease by Age



Distribution by BMI

Correlation Heatmap



Modeling

	CLASS	PRECISION	RECALL	F1 SCORE
BASELINE MODEL	0	72%	69%	70%
	1	67%	70%	
DECISION TREE	0	92%	77%	85%
	1	79%	93%	
RANDOM FOREST	0	94%	86%	90%
	1	86%	94%	

Best Performing metric

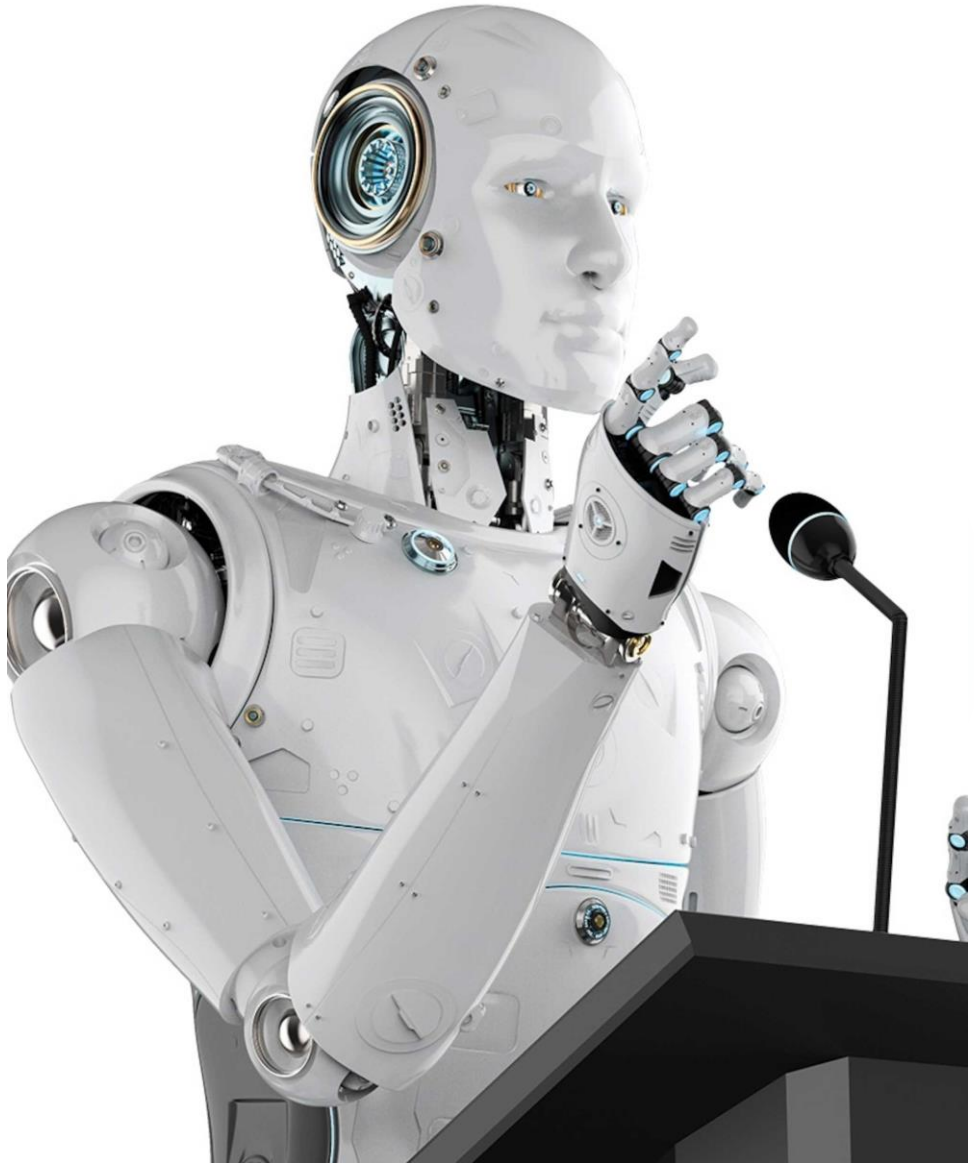
The decision to employ a Random Forest model for generating the provided classification metrics is well-justified given the achieved results. With an impressive accuracy of 89.7%, Random Forest, as an ensemble learning method, excels in handling diverse health indicators like age, blood pressure, and cholesterol levels. The model demonstrates a remarkable balance between precision and recall for both classes (0 and 1), hovering around 90%, indicating its proficiency in correctly identifying instances of heart-related complications

Recommendation

- **Age and Blood Pressure Emphasis:**
  Prioritize individuals aged 40-59 and consider systolic/diastolic blood pressure.
- **Weight Management Focus:**
  Highlight the significance of weight management, especially for overweight individuals.
- **Comprehensive Risk Assessment:**
  Encourage thorough evaluations considering multiple risk factors.

Next steps

- **Promote Heart-Healthy Behaviors:**
  Launch public health campaigns promoting exercise, balanced diets, and smoking cessation.
- **Collaboration with Medical Professionals:**
  Integrate predictive models into healthcare systems for informed decision-making.
- **Continual Improvement:**
  Regular model updates, evaluations, and feedback loops ensure ongoing accuracy.



Thank You.