Heart Disease Data Analysis & Prediction Report

# Task 1: Data Analysis Report

This report presents a comprehensive analysis of heart disease data based on multiple patient parameters. Below are the key findings from exploratory data analysis (EDA):

- Target class distribution is relatively balanced (44% positive cases).

- Heart disease is more common in patients over 50 years old.

- Males are more prone to heart disease in this dataset.

- Lower max heart rate and higher ST depression are significant indicators of heart issues.

- Features like `oldpeak\_eq\_st\_depression`, `number\_of\_major\_vessels`, and `thal` show strong correlation with disease.

# Task 2: Machine Learning Model

We trained and evaluated multiple machine learning models to predict heart disease presence. The key steps were:

- Encoded categorical feature `thal`.

- Standardized all features for uniformity.

- Split the dataset into 80% training and 20% testing.

- Trained Logistic Regression and Random Forest classifiers.

Random Forest generally provides better accuracy and feature importance insights.

# Task 3: Hospital Recommendations

Based on the patterns identified, we recommend the following actions:

- Implement treadmill stress tests for patients over 50.

- Encourage fitness programs to boost max heart rate capacity.

- Regular screening for patients with ST depression or abnormal thalassemia (`thal`) readings.

- Focus on male-targeted heart health awareness programs.

- Deploy an ML-based online heart disease risk calculator.

- Monthly workshops and app notifications for lifestyle improvements.

Here’s a realistic high-risk sample input that may indicate potential heart disease:

* Age: 63
* Sex (1=Male, 0=Female): 1
* Chest Pain Type (0-3): 3 (asymptomatic — high risk)
* Resting BP: 145
* Cholesterol: 233
* Fasting Blood Sugar > 120? (1=Yes, 0=No): 1
* Resting EKG (0=Normal, 1=ST-T, 2=LVH): 2 (left ventricular hypertrophy)
* Max Heart Rate: 150
* Exercise Angina (1=Yes, 0=No): 1
* Oldpeak: 2 (significant ST depression)
* Slope (0=Upsloping, 1=Flat, 2=Downsloping): 2
* Major Vessels (0-3): 2
* Thal (1=Normal, 2=Fixed Defect, 3=Reversible): 3

## Prediction: ⚠️ You may have heart disease.

**Low-Risk Example Input**

* **Age: 38**
* **Sex (1=Male, 0=Female): 0**
* **Chest Pain Type (0-3): 1 *(non-anginal pain)***
* **Resting BP: 115**
* **Cholesterol: 180**
* **Fasting Blood Sugar > 120? (1=Yes, 0=No): 0**
* **Resting EKG (0=Normal, 1=ST-T, 2=LVH): 0**
* **Max Heart Rate: 175**
* **Exercise Angina (1=Yes, 0=No): 0**
* **Oldpeak: 0.2**
* **Slope (0=Upsloping, 1=Flat, 2=Downsloping): 0**
* **Major Vessels (0-3): 0**
* **Thal (1=Normal, 2=Fixed Defect, 3=Reversible): 1**

### Prediction :✅ You are at low risk

Medium Risk ⚠️ Please consult a doctor.

* **Age:** 58
* **Sex:** 1 *(Male)*
* **Chest Pain Type:** 2 *(Atypical Angina)*
* **Resting BP:** 140
* **Cholesterol:** 250
* **Fasting Blood Sugar > 120?:** 1 *(No)*
* **Resting EKG:** 1 *(ST-T wave abnormality)*
* **Max Heart Rate:** 150
* **Exercise Angina:** 0 *(No)*
* **Oldpeak:** 1.5
* **Slope:** 1 *(Downsloping)*
* **Major Vessels:** 1
* **Thal:** 2 *(Reversible Defect)*

### Prediction : Medium Risk ⚠️ Please consult a doctor.