

Heart Disease Prediction – Detailed Project Report

1. Introduction

Heart disease prediction using machine learning to classify patients as having heart disease or not.

2. Problem Statement

Predict binary output (0/1) for heart disease using medical attributes.

3. Dataset Overview

Dataset: Heart Failure Prediction Dataset (Kaggle)

Rows: 918, Columns: 12, Target: HeartDisease

4. Methodology

Steps include data collection, cleaning, EDA, feature engineering, model training, evaluation, and saving.

5. Data Cleaning

Removed duplicates, handled missing values, encoded categorical columns.

6. EDA

Analyzed distributions, correlations, risk factors.

7. Modeling

Random Forest Classifier selected for its high accuracy and robustness.

8. Results

Accuracy achieved: ~86–90%, high precision and recall.

9. Future Work

Deployment using Flask/FastAPI, UI, dashboards, hypertuning.

10. Conclusion

Successfully completed end-to-end ML pipeline with strong performance.