Heart Disease Prediction Using Random Forest

1. Introduction

Heart disease is one of the leading causes of death worldwide. Early detection
and accurate prediction can help save lives by enabling timely medical intervention. This project applies
machine learning techniques, specifically the Random Forest algorithm, to predict heart disease based on
patient health data.

2. Objective

The goal of this project is to develop a predictive model that can classify whether
a patient has heart disease or not based on their medical attributes. By using an advanced machine
learning approach, we aim to improve diagnostic accuracy and assist healthcare professionals in
decision-making.

3. Methodology

3.1 Dataset Description

- Dataset Used: Heart Disease UCI dataset
- Number of Records: 303
- Number of Features: 13 + Target variable (Presence/Absence of heart disease)
- Target Variable: target (1 = Disease present, 0 = No disease)

3.2 Data Preprocessing

- Handling Missing Values: Checked and replaced missing values if necessary.
- Feature Scaling: Standardized numerical features to improve model performance.
- Feature Selection: Used feature importance methods to select relevant attributes.

3.3 Model Selection & Training

- Algorithm Used: Random Forest Classifier
- Splitting Data: 80% training, 20% testing

- Evaluation Metrics: Accuracy, Precision, Recall, F1-score

4. Results & Observations

The Random Forest model achieved high accuracy in predicting heart disease.

Performance was evaluated using standard classification metrics, and the model demonstrated reliable predictions.

Graphs and visualizations were used to interpret results effectively.

5. Conclusion

This project successfully demonstrated the application of machine learning in medical

diagnostics. The Random Forest model provided accurate predictions, which can assist healthcare professionals

in early detection of heart disease. Future improvements may include hyperparameter tuning and integration

with real-time healthcare systems.

6. References

- UCI Machine Learning Repository: Heart Disease Dataset
- Machine Learning Textbooks & Research Papers
- Online documentation for Random Forest implementation