

Heart Disease ML Project Report

■ Dataset:

The dataset contains heart disease patient records. After preprocessing and cleaning, we applied PCA to reduce dimensions from 34 features to 15 principal components.

■■ Project Steps:

1. Data Preprocessing (handling missing values, cleaning).
2. Dimensionality Reduction using PCA.
3. Training multiple models (Logistic Regression, Random Forest, XGBoost).
4. Model selection and saving the best model.
5. Evaluation on test data (classification report, accuracy, confusion matrix).
6. Visualization of results.

■ Model Results:

LogReg: 0.5435
RandomForest: 0.5543
XGBoost: 0.5489

Best Model: RandomForest with accuracy 0.5543



