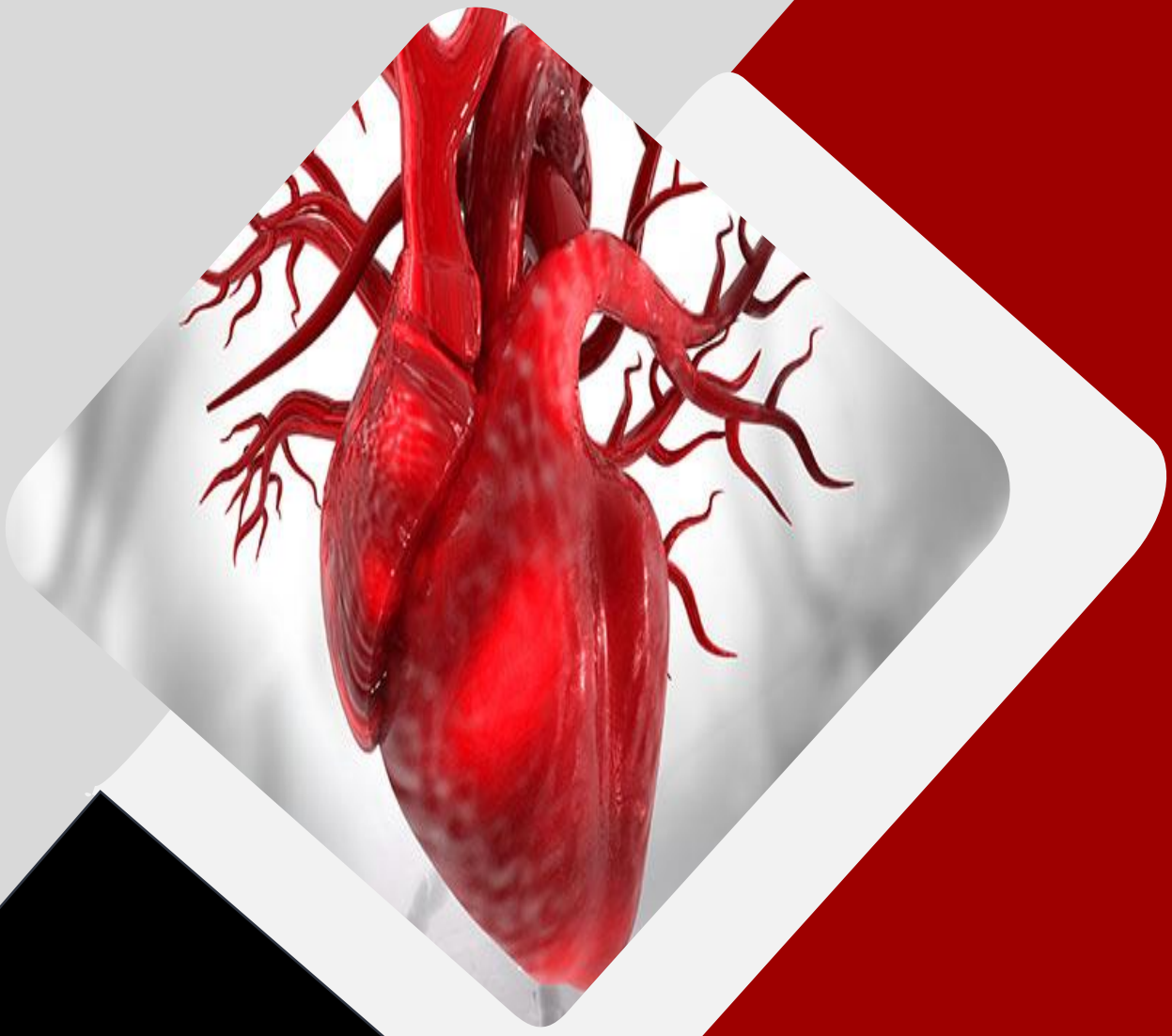


Heart Disease Detection

By: Dalal Qallajo

Description: This project focuses on predicting the presence of heart disease in patients using six machine learning classifiers:

#Scale-insensitive classifiers

Random Forest classifier - Naive Bayes classifier - Gradient boosting classifier

#Scale-sensitive classifiers

K-nearest neighbors (KNN) - Logistic regression - Support vector classifier (SVC)

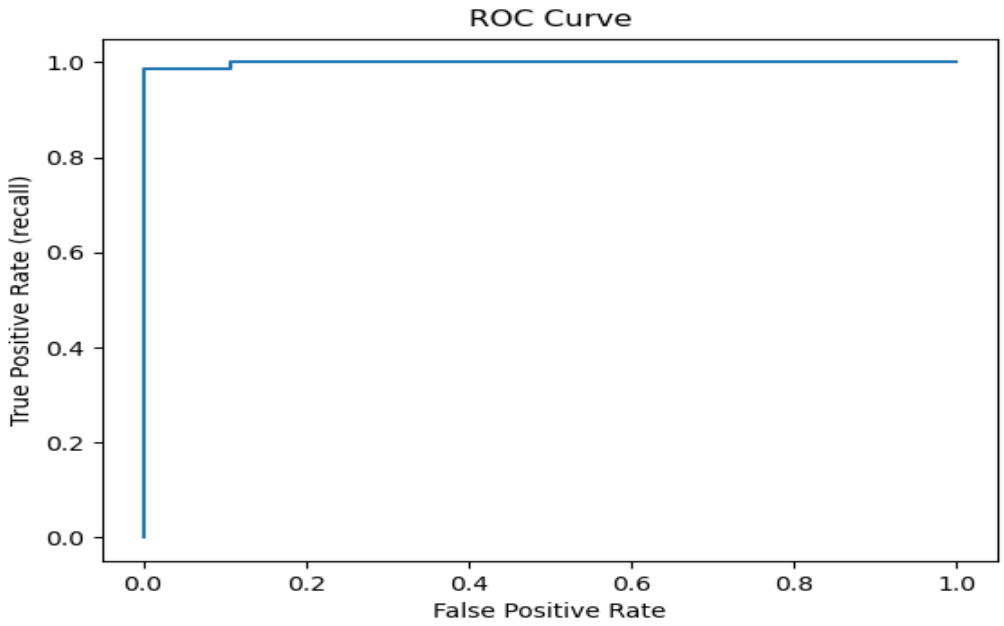
The main objectives and functionalities of the project include:

- Data Acquisition: Utilizes a publicly available heart disease dataset from Kaggle to analyze patient features and their correlation with heart disease outcomes.
- Data Preprocessing: Involves splitting the dataset into training and testing sets, handling feature scaling for algorithms sensitive to data scale, and preparing the data for model training.
- Model Implementation: Implements multiple machine learning classifiers
- Model Evaluation: Evaluates model performance using accuracy scores and recall metrics, providing insights into the effectiveness of each classifier in predicting heart disease. The Random Forest model was specifically chosen for hyperparameter tuning due to its highest accuracy among the six models evaluated.
- Hyperparameter Tuning: Utilizes GridSearchCV to optimize the hyperparameters of the Random Forest model, enhancing its predictive accuracy and ensuring the best possible performance.

Visualization:

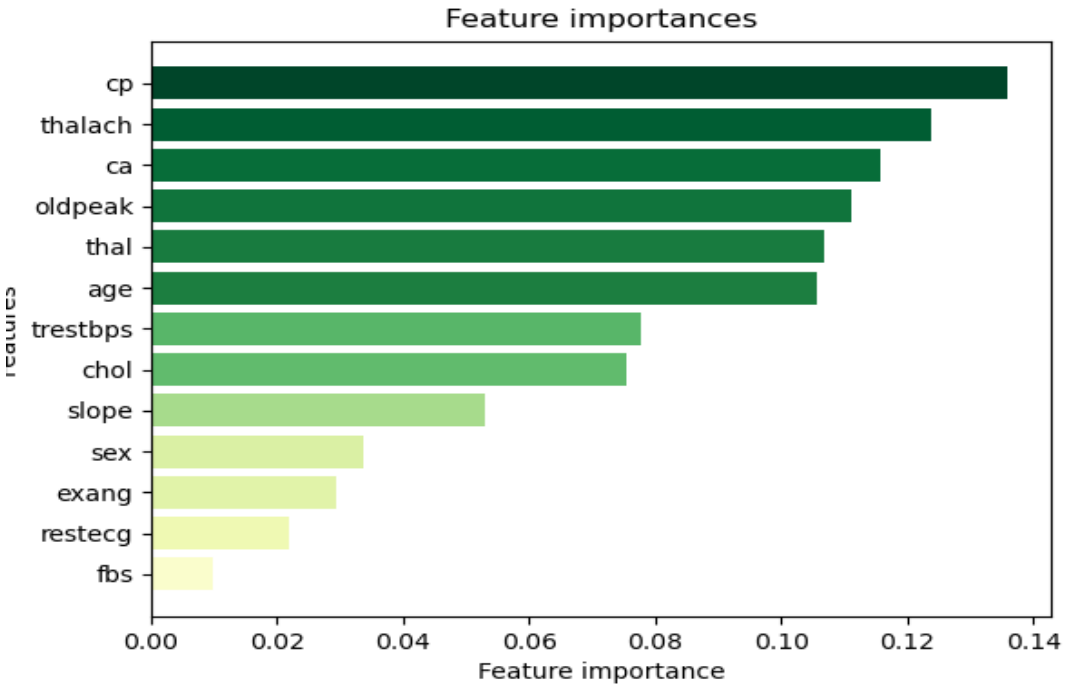
ROC Curve

Generates ROC curves to visualize the trade-off between true positive rates and false positive rates, along with calculating the area under the curve (AUC) for model performance assessment.



Feature Importance

Analyzes and visualizes feature importance to identify which patient characteristics are most influential in predicting heart disease.



Code link on GitHub:

https://github.com/Dalalkaljo/Deep-Learning/blob/main/Heart%20Disease%20Prediction%20Using%20Machine%20Learning