



Module 3 – Heart Disease

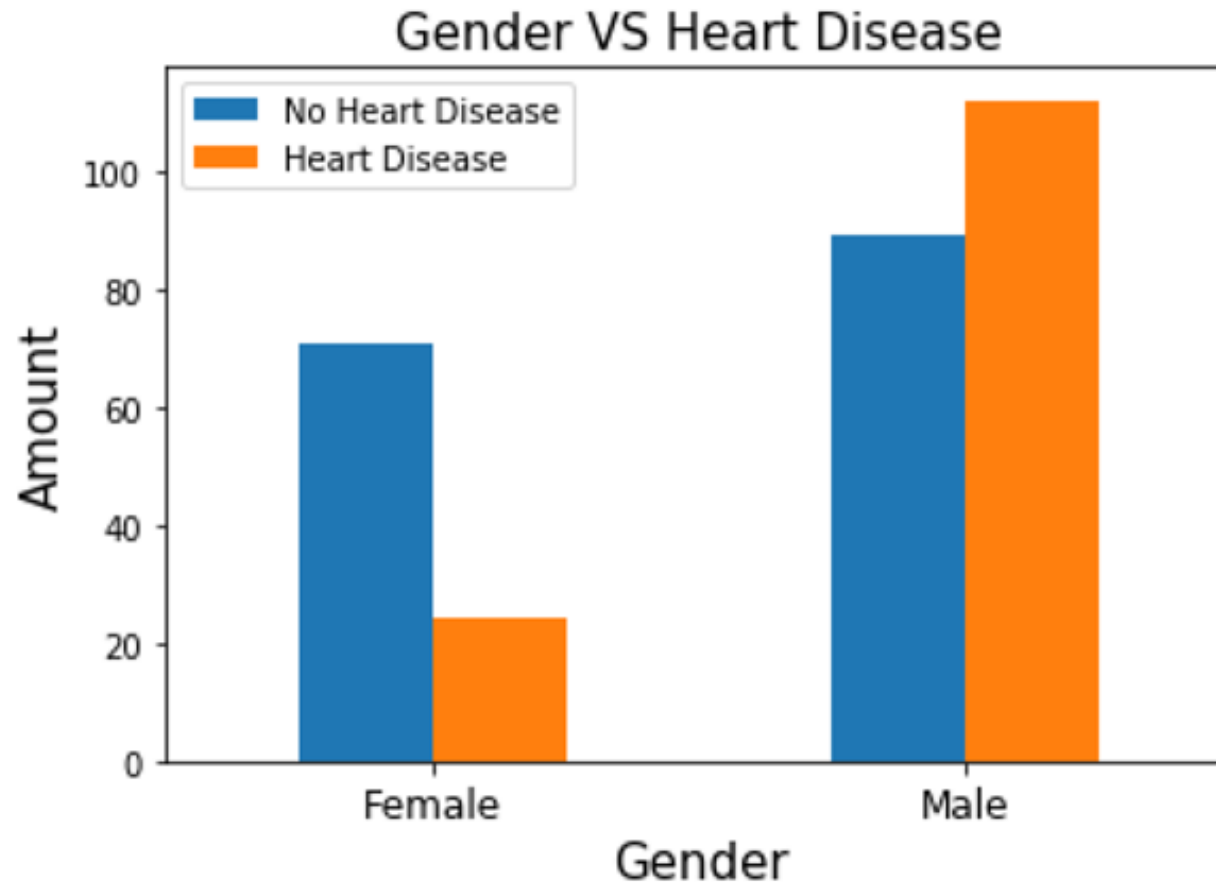
Andy Peng



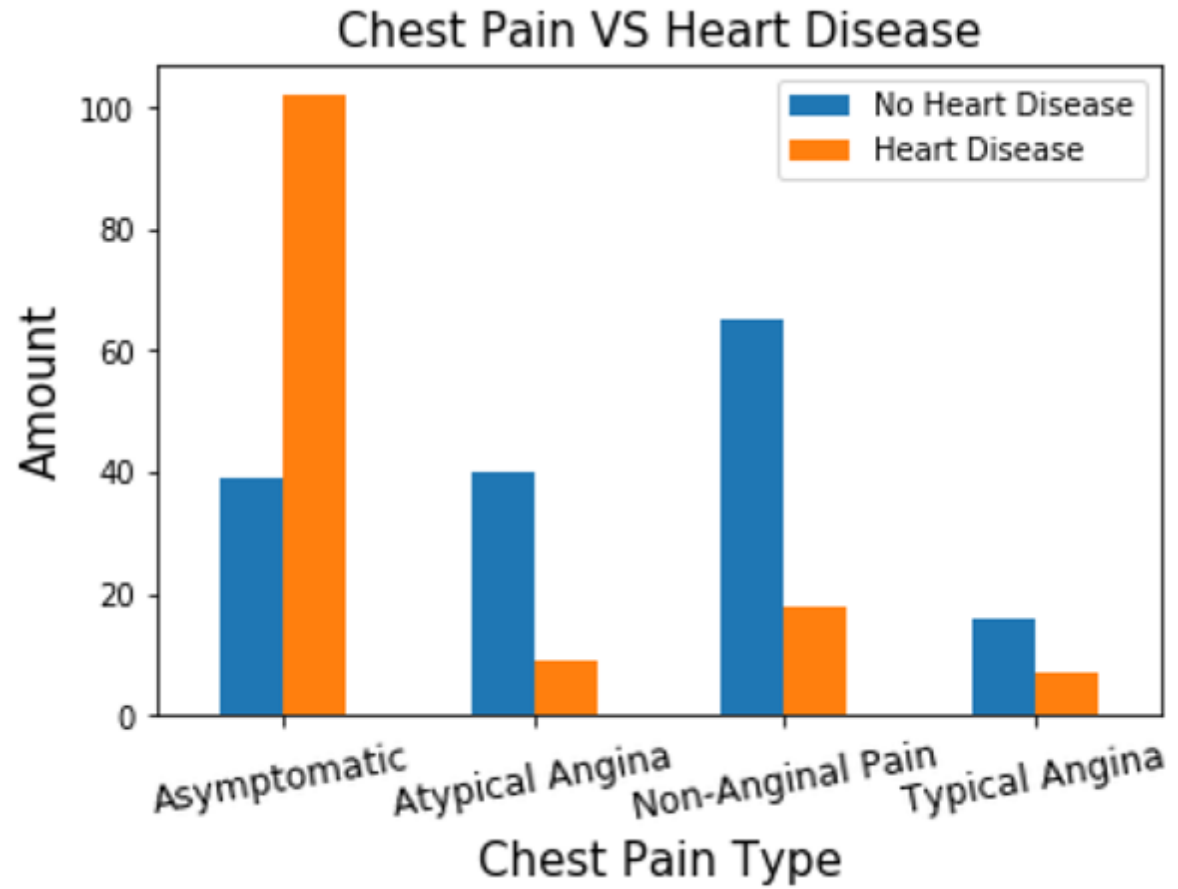
Overview

- Relationship to heart disease
 - Gender
 - Thallium Stress Result
 - Age
- Modeling

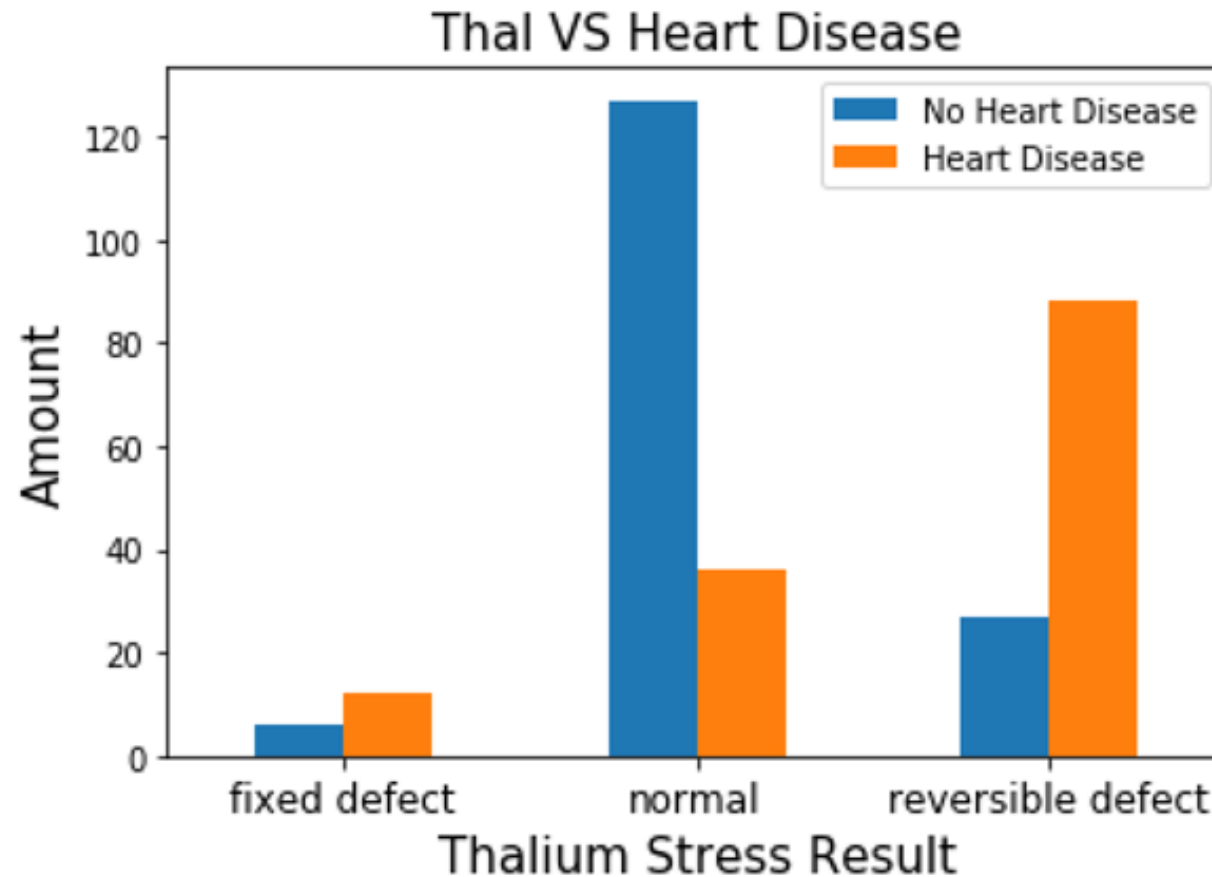
Does Gender affect Heart Diseases?



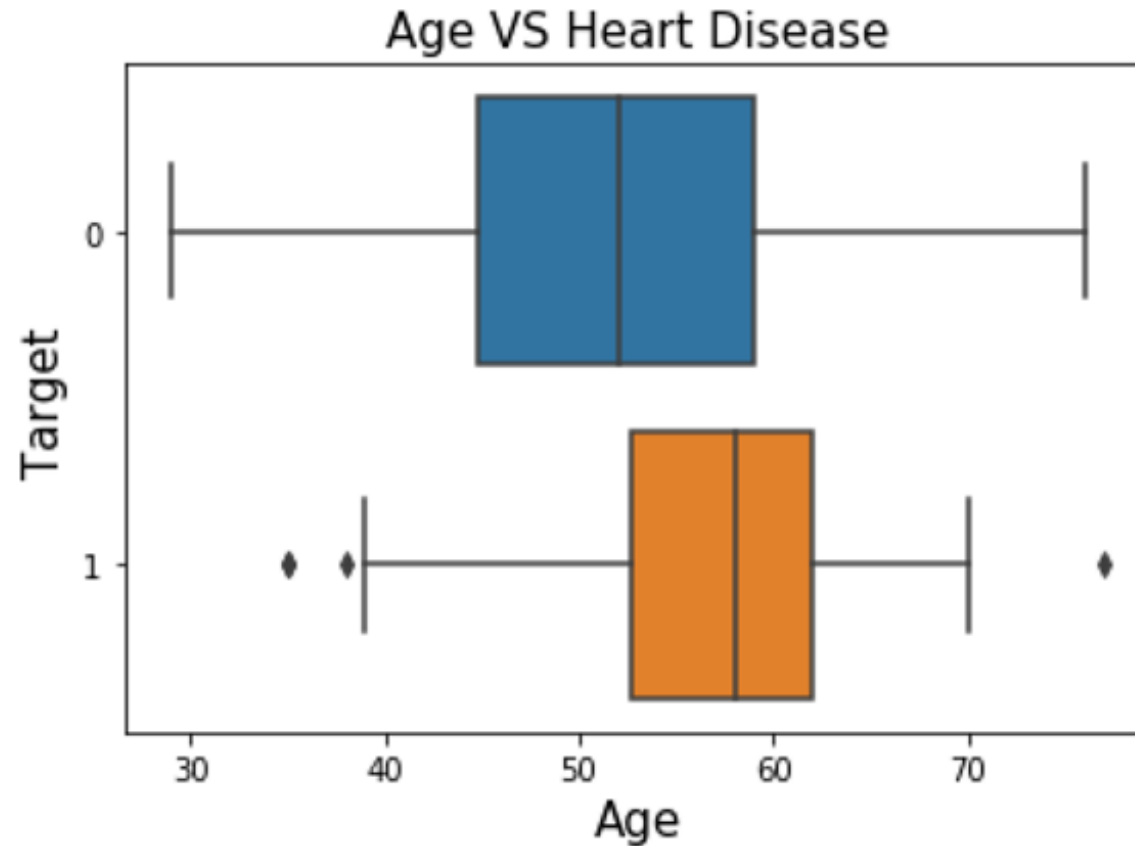
Chest Pain



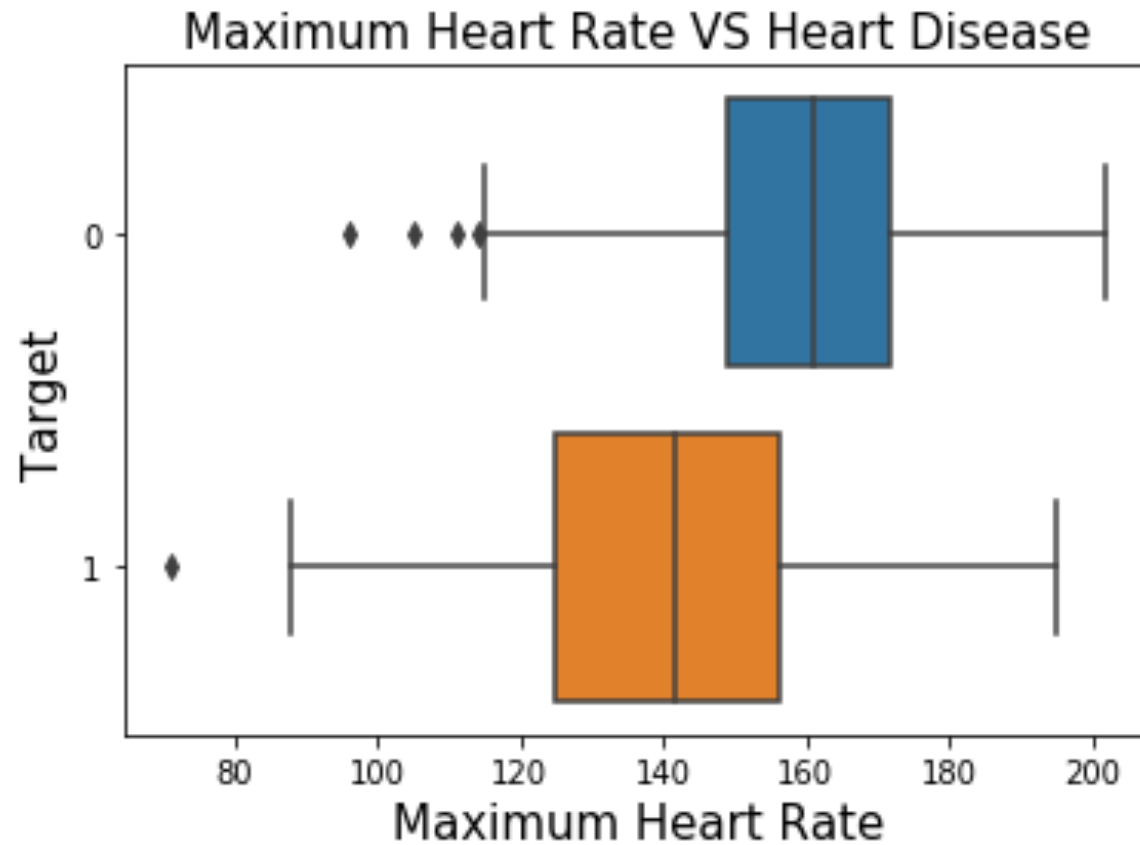
Thallium Stress Results



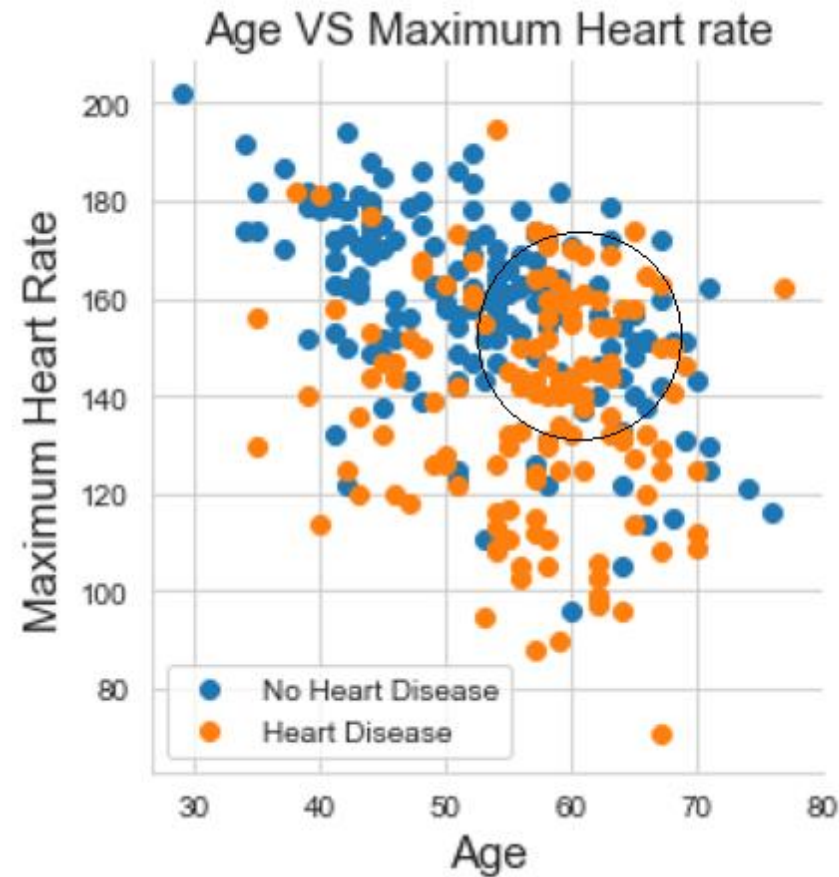
Does age affect Heart Diseases?



Maximum Heart Rate



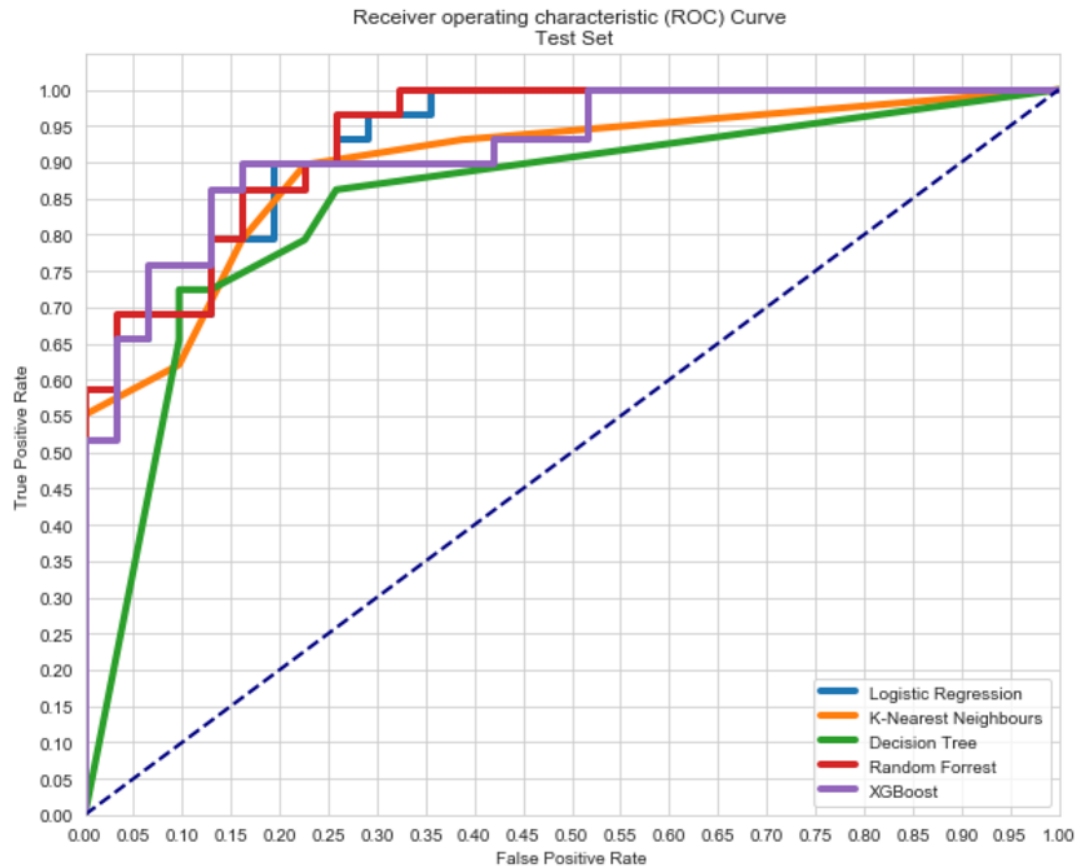
Maximum Heart Rate



Modeling

- Logistic Regression
- K Nearest Neighbors
- Decision Tree
- Random Forest
- XGBoost
- GridSearchCV

Results



- AUC - Random Forest with GridSearchCV
- Accuracy/ F1 Score/ Recall - XGBoost
- Precision - Decision Tree with GridSearchCV

Recommendations

- XGBoost for Recall Score

Base Model

Model	Test Accuracy	Test F1	Test Precision	Test Recall	AUC Score
0 Logistic Regression	0.83	0.82	0.85	0.79	0.931034
1 K-Nearest Neighbours	0.80	0.79	0.81	0.76	0.894883
2 Decision Tree	0.75	0.74	0.75	0.72	0.749166
3 Random Forrest	0.80	0.78	0.84	0.72	0.927697
4 XGBoost	0.87	0.86	0.86	0.86	0.924360

Model Tuning

Model	Test Accuracy	Test F1	Test Precision	Test Recall	AUC Score
0 Logistic Regression	0.83	0.82	0.85	0.79	0.931034
1 K-Nearest Neighbours	0.82	0.81	0.82	0.79	0.895996
2 Decision Tree	0.80	0.77	0.87	0.69	0.845940
3 Random Forrest	0.78	0.75	0.83	0.69	0.935484
4 XGBoost	0.85	0.84	0.86	0.83	0.919911

Recommendations

- Gender
- Asymptomatic Chest Pain
- Reversible Defect
- Age & Maximum Heart Rate

Next Steps

- Features to Explore

- Genetic Disorders
- Body Weight Percentages
- Patient's Diet

- Gather More Data



Thank You