



Heart Failure Prediction Data Analysis

Christina Reeder



Stakeholder Introduction

- The stakeholder is hospitals and doctors treating heart disease in patients.
- The goal of this model is to predict if a patient will develop a heart disease based on a variety of health-related attributes the patient exhibits

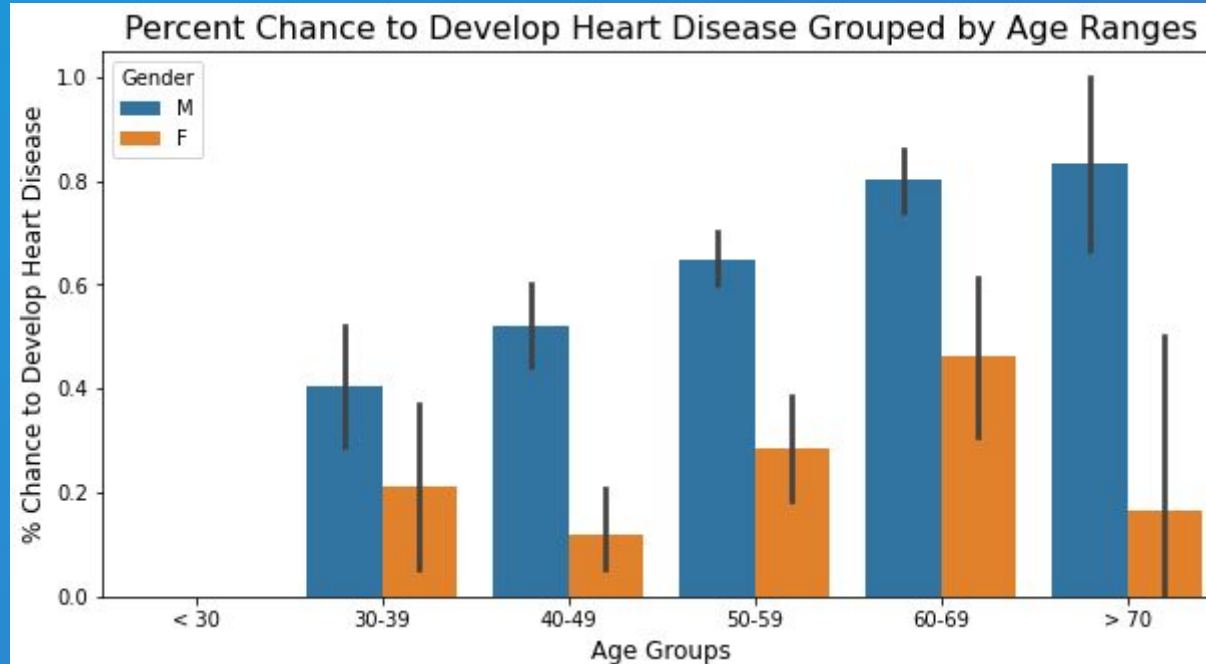
Heart Failure Dataset

Attribute Information

1. Age: age of the patient [years]
2. Sex: sex of the patient [M: Male, F: Female]
3. ChestPainType: chest pain type [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic]
4. RestingBP: resting blood pressure [mm Hg]
5. Cholesterol: serum cholesterol [mm/dl]
6. FastingBS: fasting blood sugar [1: if FastingBS > 120 mg/dl, 0: otherwise]
7. RestingECG: resting electrocardiogram results [Normal: Normal, ST: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV), LVH: showing probable or definite left ventricular hypertrophy by Estes' criteria]
8. MaxHR: maximum heart rate achieved [Numeric value between 60 and 202]
9. ExerciseAngina: exercise-induced angina [Y: Yes, N: No]
10. Oldpeak: oldpeak = ST [Numeric value measured in depression]
11. ST_Slope: the slope of the peak exercise ST segment [Up: upsloping, Flat: flat, Down: downsloping]
12. HeartDisease: output class [1: heart disease, 0: Normal]



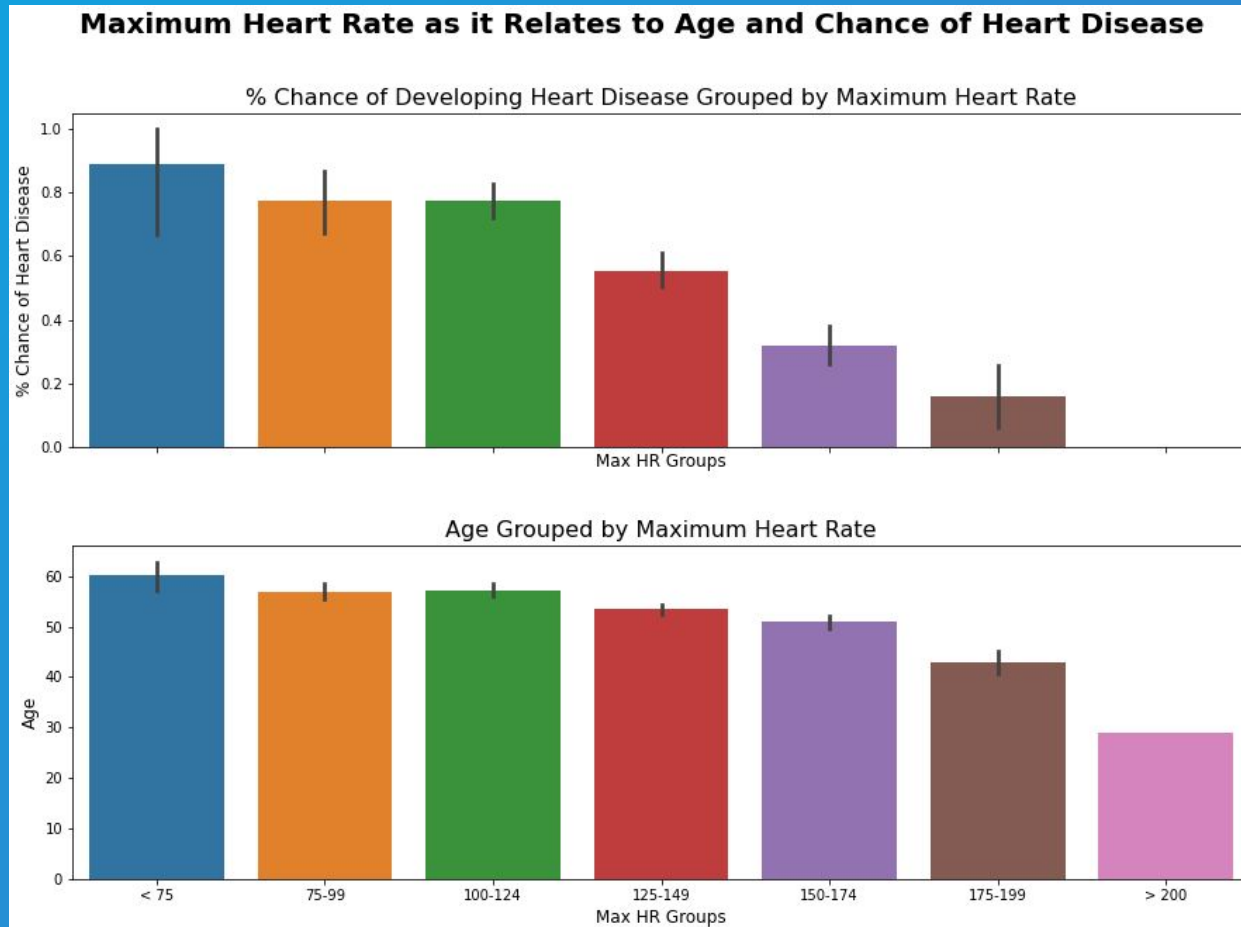
Exploring Chance of Heart Disease Based on Age and Gender



There is a significantly higher percent chance for male patients to develop a heart disease than female patients. The percent chance to develop heart disease increases as the age of the male patients increases. However, the chance of heart disease is highest in females in the age range of 60-69 and lowest in females in the age range of 40-49.



Explore How Maximum Heart Rate Effects Percent Chance of Heart Disease and Changes with Age



- The top plot shows that the chance a patient will develop a heart disease decreases as the maximum heart rate achieved by that patient increases, meaning that the faster the heart is able to beat correlates to how healthy that heart is.
- The bottom plot shows that the maximum heart rate achieved by a patient decreases with the age of the patient.
- This leads me to the conclusion that older patients are at a higher risk for heart disease than younger patients on average.



Effects of False Positives and False Negatives on the Stakeholder

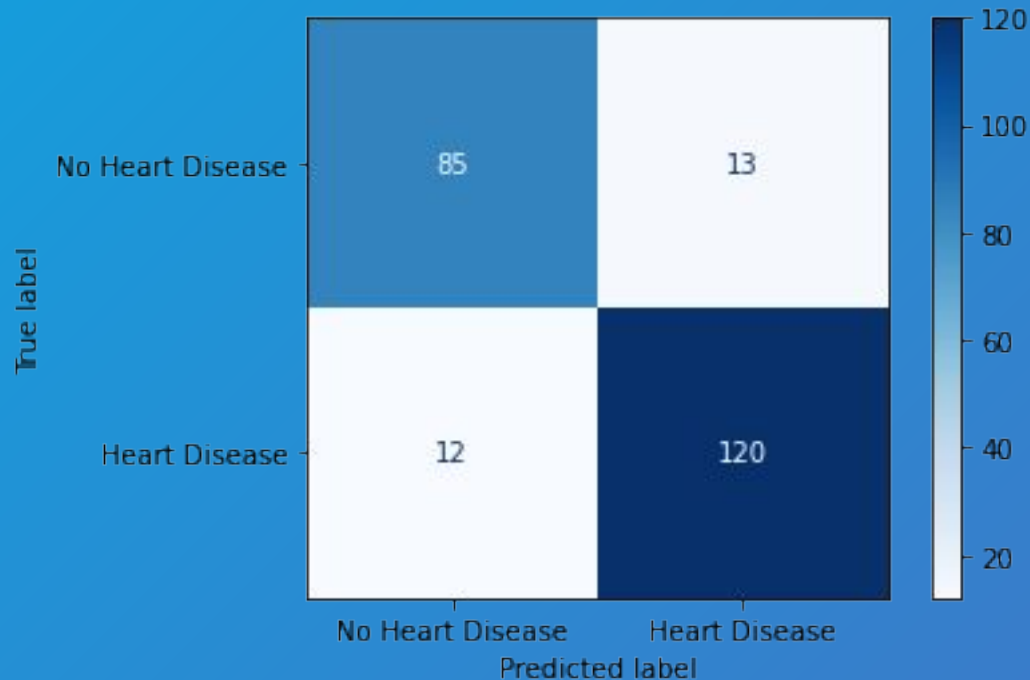
For this model:

- 0: No heart disease in the patient
- 1: Heart disease in the patient

I prioritized minimizing errors in which a patient with heart disease is predicted to not have a heart disease (false negative prognosis).

A patient being given a false negative prognosis will preclude them from receiving treatment they may otherwise need to save their life.

Best Model Strengths and Limitations



- The production model was able to predict heart disease with an accuracy of 89%
- There was still 12 patients that were given a false negative prognosis
- There were 13 patients that were given a false positive prognosis



Final Recommendations

- It is my recommendation that more data points be collected on these patients' health history to better train the model for predicting the risk of a patient developing a heart disease.
- As it is, I do not believe this model is sufficient be utilized to predict anything relating to a patient's health or treatment. The number of false positive and false negative predictions made by the model are far too high when someone's life is on the line.
- I would require an accuracy in a model around 98-99% to even begin considering commercial use.