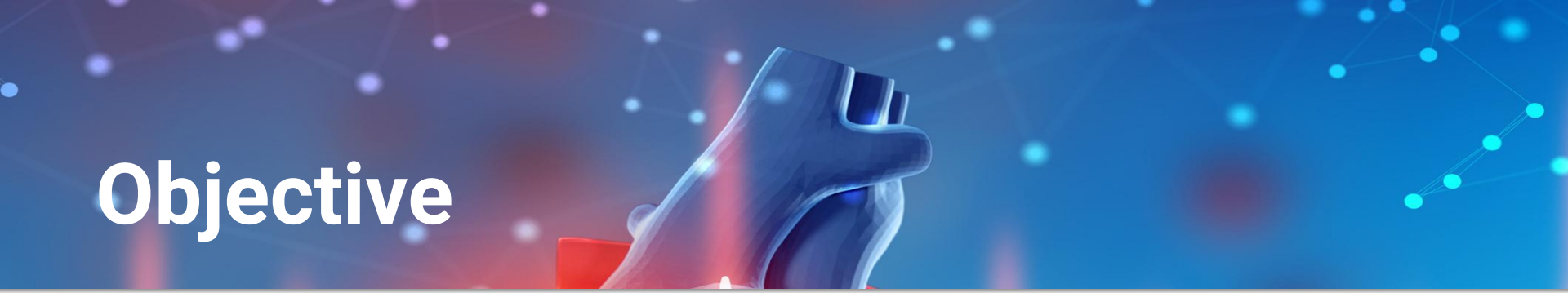
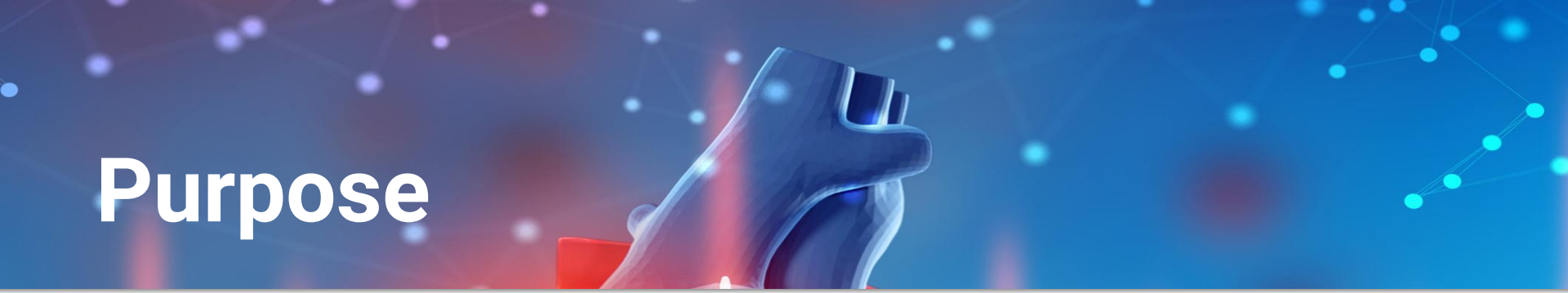
# Heart Attack Prediction

By: Audrey Emeribe



# Objective

**To predict the likelihood of a patient getting a heart attack from patient attributes and medical data**



# Purpose

Quickly note a patient’s susceptibility to a heart attack

# Dataset Contents

**Target Variable = Output(Heart attack or not)**

**Features:**

- Age
- Sex(Male=1, Female=0)
- Chest pain type
- Resting blood pressure
- Cholesterol level
- Fasting blood sugar
- Resting ECG results
- Maximum heart rate achieved
- Old peak
- Slope
- Number of major vessels
- Thallium stress test result
- Exercised induced angina

# Data Cleaning

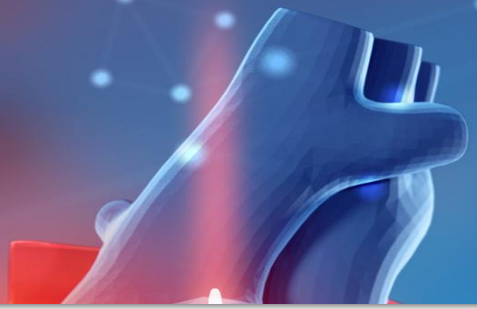
- Started with 3500+ rows & 15 columns

- Dropped o2Saturation column as it made dataset irregular

- Clean data had 303 rows with 14 column

# Output: 1= high chance of heart attack, 0= low chance of heart attack
# Sex: 1= Male, 0= Female

# Insights



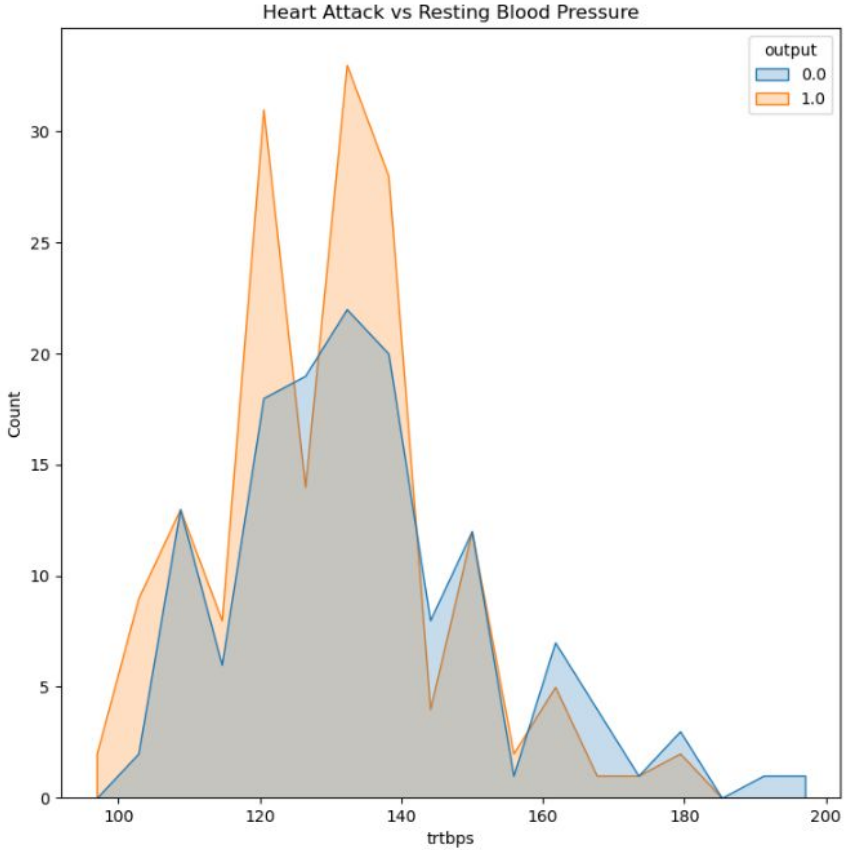
FEMALES:

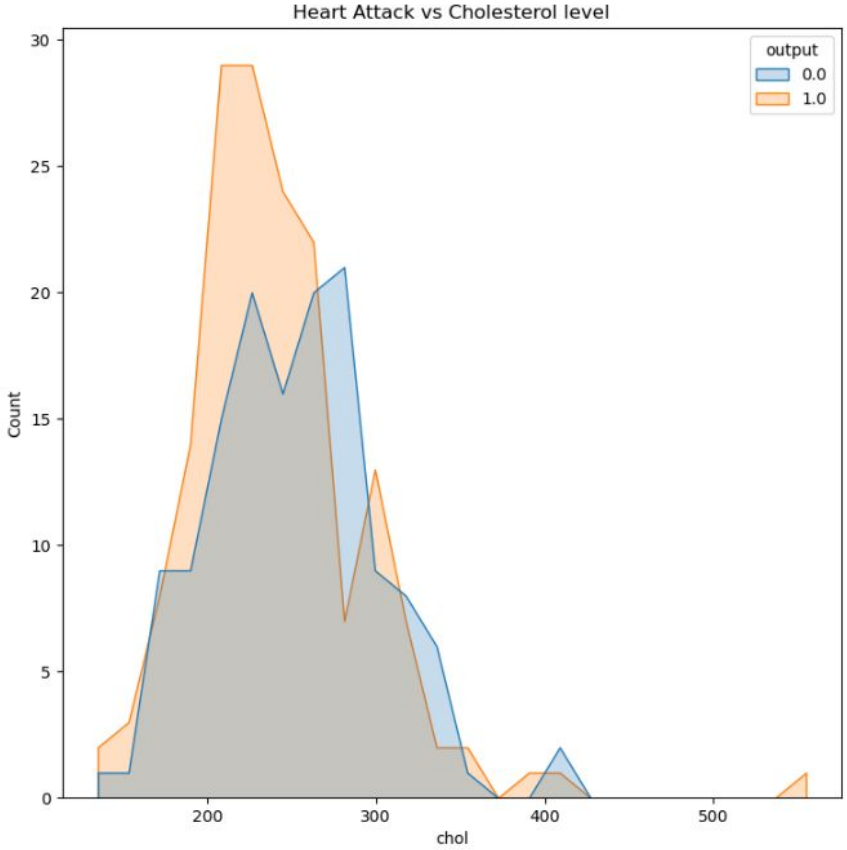
- Less likely to have heart attack
- Median age of 54 to experience heart attack
- Heart attack more likely in later years

MALES:

- More likely to have heart attack
- Median age of 52 to experience heart attack
- Heart attack occurence more likely in earlier years

# More Insights

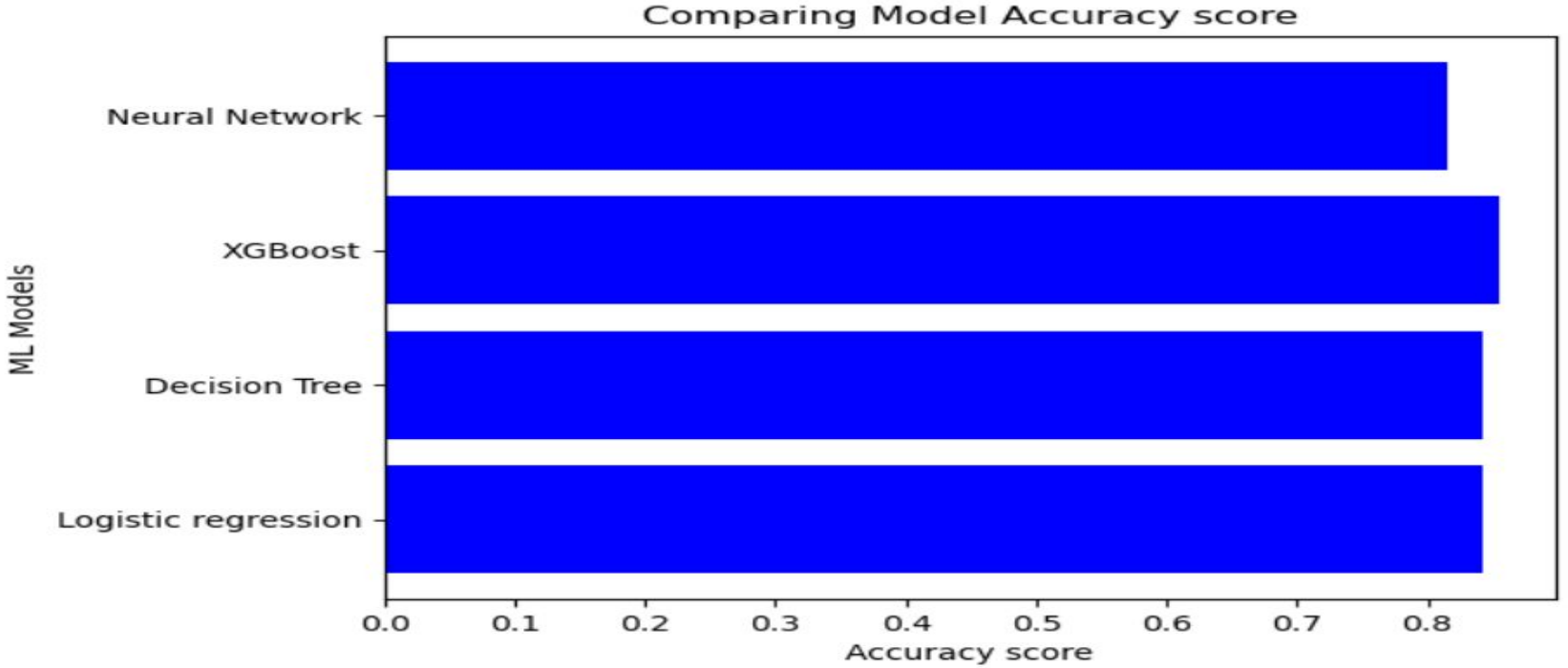


# More Insights

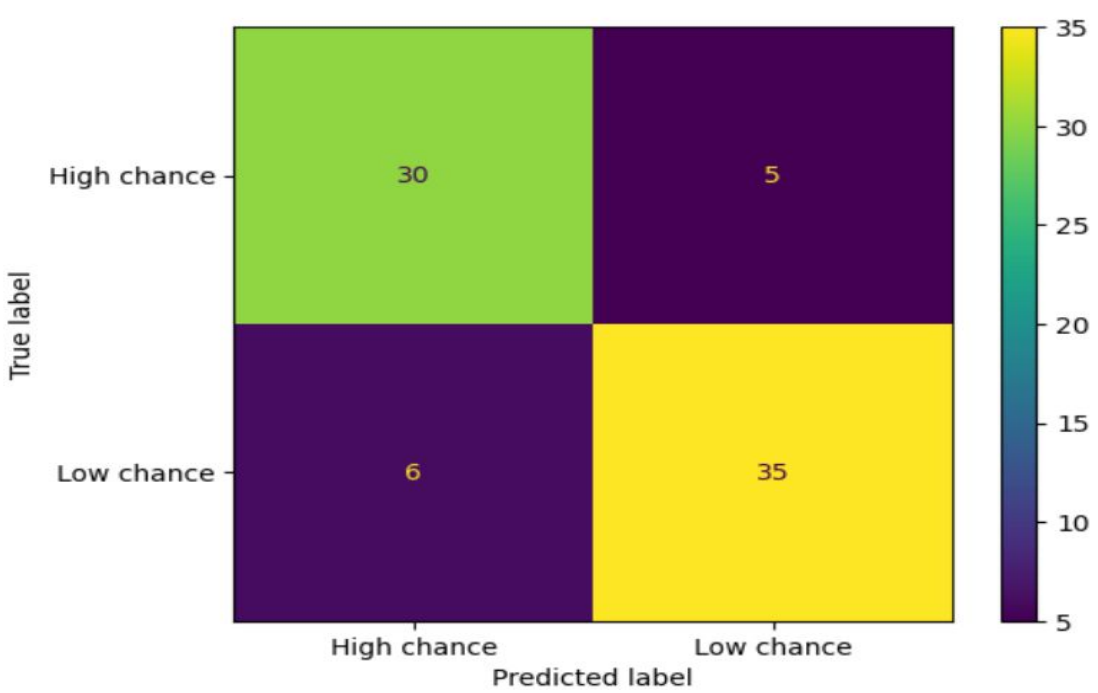
- Resting blood pressure between 120 - 140 had higher heart attack chances
- Increased cholesterol rates also increased the chances of a heart attack

# ML Models Fitted



# XGBOOST MODEL

- Accuracy of 85.5%
- Higher recall of 86% for high chance of heart attack
- High True Positive and true negative predictions in the confusion matrix

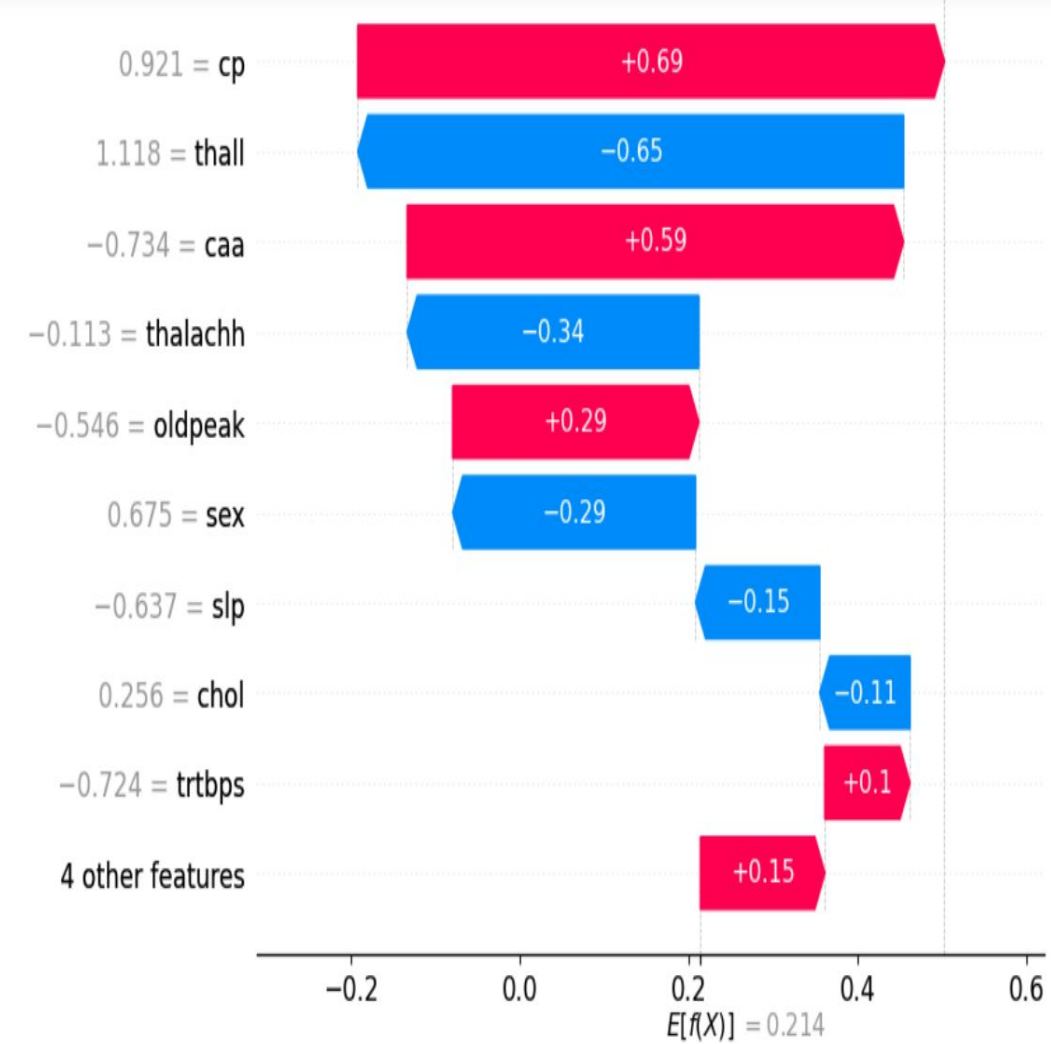



| | precision | recall | f1-score | support |
|---|---|---|---|---|
| High chance | 0.83 | 0.86 | 0.85 | 35 |
| Low chance | 0.88 | 0.85 | 0.86 | 41 |
| accuracy | | | 0.86 | 76 |
| macro avg | 0.85 | 0.86 | 0.85 | 76 |
| weighted avg | 0.86 | 0.86 | 0.86 | 76 |

# Model Interpretation using SHAP values(Patient 1) - 50% chance of Heart attack

## Features that increase heart attack chances for patient 1

- Chest pain
- Number of major vessels
- Old peak
- Resting blood pressure



# IMPORTANT FEATURES NOTED

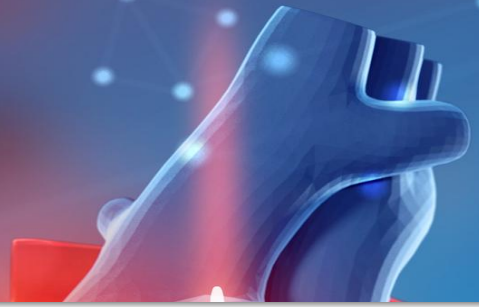
- Chest pain type
- Sex of the patient
- Thallium stress test result
- Number of major vessels

# Deployment

- Create a web app using the FLASK framework

- Because of data size, store on github all dependencies and deploy on Heroku
- Test and monitor model performance using New Relic add-on on Heroku platform

# Next Steps



- Sample deployed app : http://ds-heart-attack.herokuapp.com/
- For larger datasets, I would opt for cloud platform options such as GCP with TensorFlow Extended(TFX) or AWS Sagemaker with the help and contributions of software engineers and UX manager on the team

# Custom Question

- With 10 minutes of continuous ECG data, there won’t be change in my model prediction as ECG doesn’t contribute to prediction. I will either change my model to a Logistic regression model which included ECG feature as a prediction contributor
- Or retrain my model with more patient data as the previous train set is too small

# Thank You

**Github repo:**
https://github.com/audreyemeribe/Heart-Attack-NB