



Heart Disease Predictor

Created by Leah Nagy

FUNCTIONAL MEDICINE

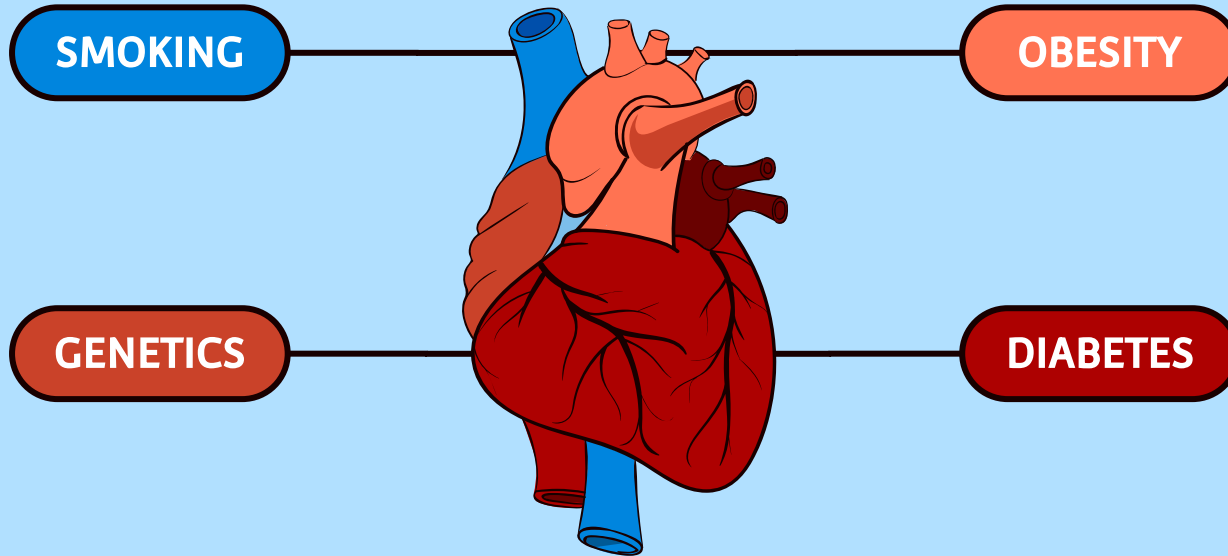


- In 2015 - diagnosed with an incurable, disabling disease
- **Functional Medicine*** found the root cause
- Created individualized treatment plan
- Today - healthy, disease-free!!!

* **Functional Medicine** is a systems biology-based approach that focuses on identifying and addressing the root cause of disease.

CAUSES OF HEART DISEASE

1 Leading Cause of Death in the United States



80% of heart disease is preventable

PROJECT GOALS



Predictive Model

using SKLearn, Pandas, and other libraries to determine probability of current heart disease



App Prototype

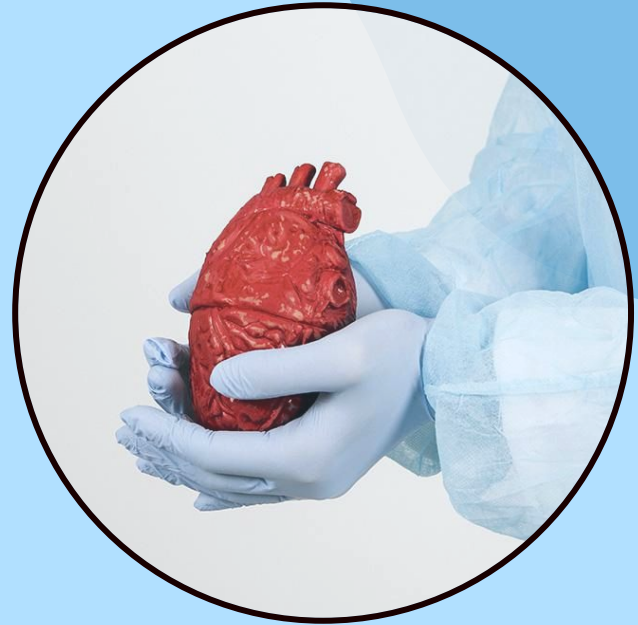
using StreamLit for patients to answer lifestyle questions to be shared with their doctor



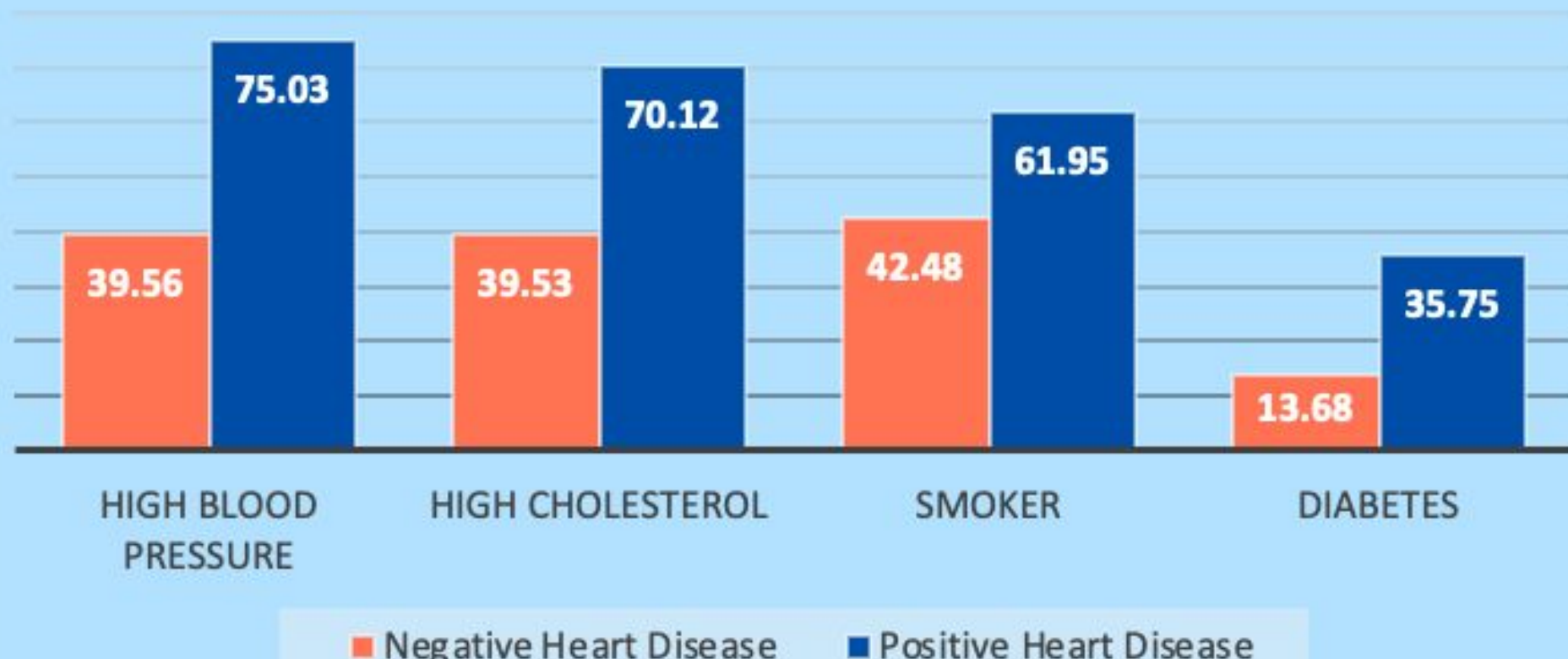
Streamlit

DATASET

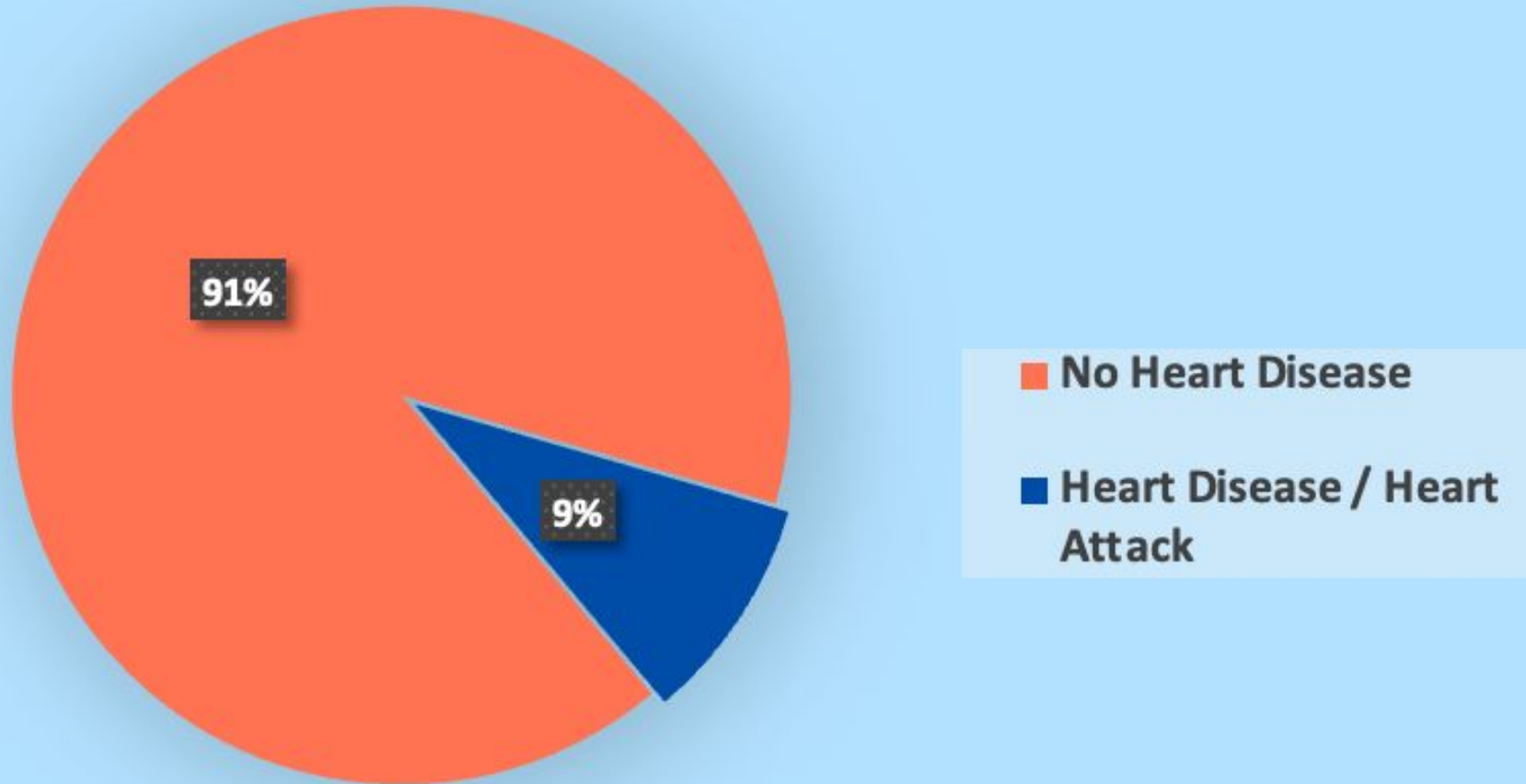
- Data Source: The Behavioral Risk Factor Surveillance survey conducted by the CDC
- Sample Size: 253,680
- Target: Presence/History of Heart Disease or Attack
- Features: 21 Health & Lifestyle Features (such as: age, sex, general health, diabetes, blood pressure, smoker, activity level, etc.)



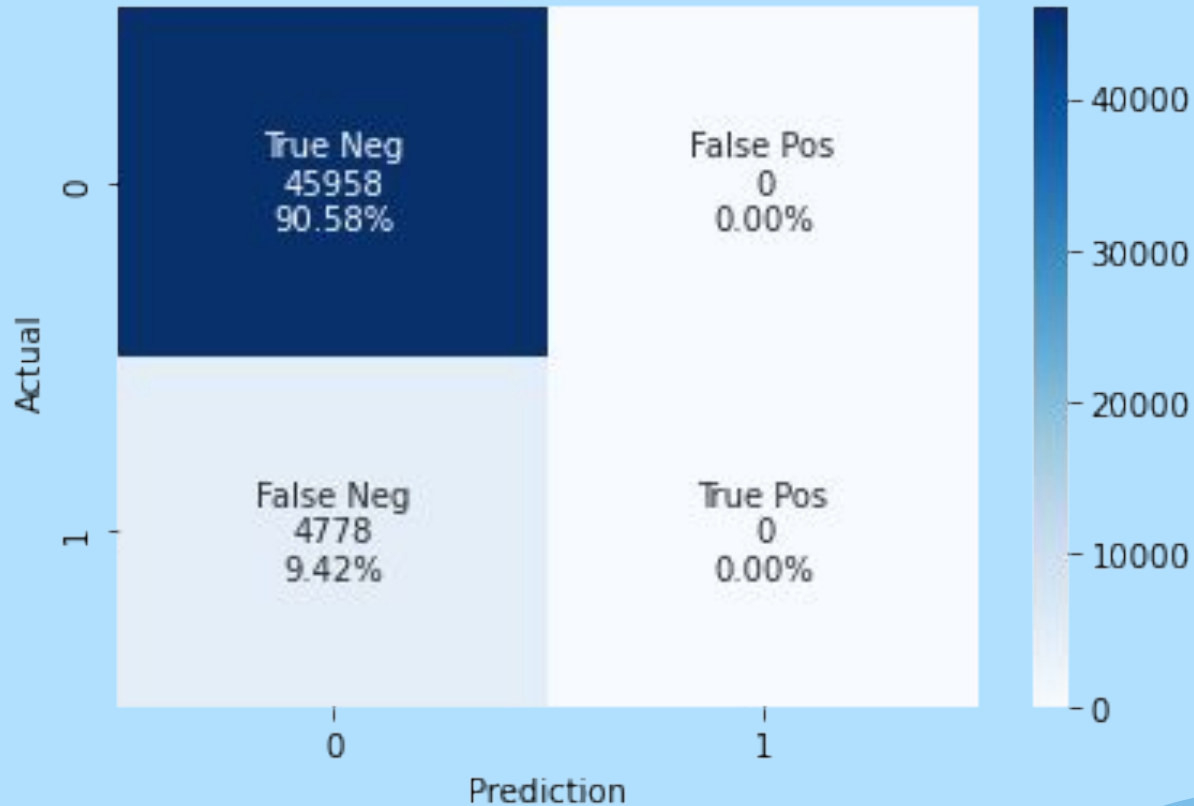
Risk Factors & Presence of Heart Disease



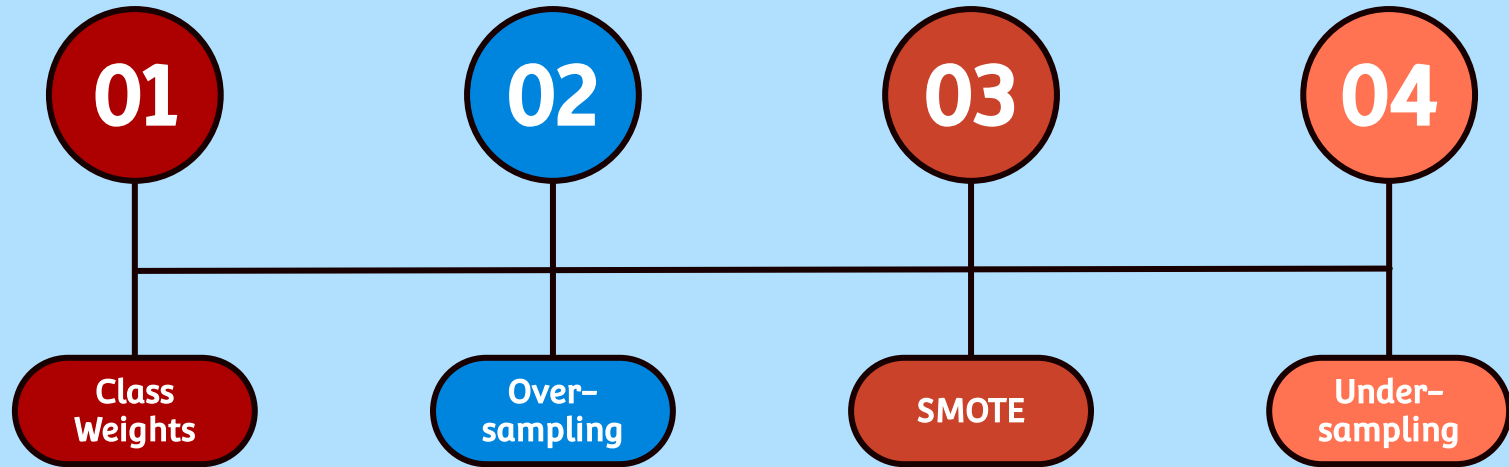
Distribution of Target Variable



Confusion Matrix – Simple Logistic Regression

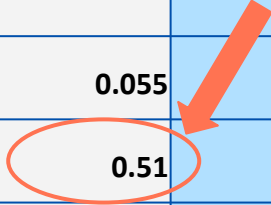


Class Imbalance Solutions



Class Imbalance Solutions

Logistic Regression Model	Accuracy	Recall	Precision	F2	Brier Loss
Baseline	0.905	0.045	0.474	0.055	0.905
Class Weights	0.7	0.792	0.21	0.51	0.7
Oversampling	0.724	0.756	0.22	0.508	0.724
SMOTE	0.726	0.75	0.22	0.506	0.726
Undersampling	0.725	0.755	0.22	0.508	0.725

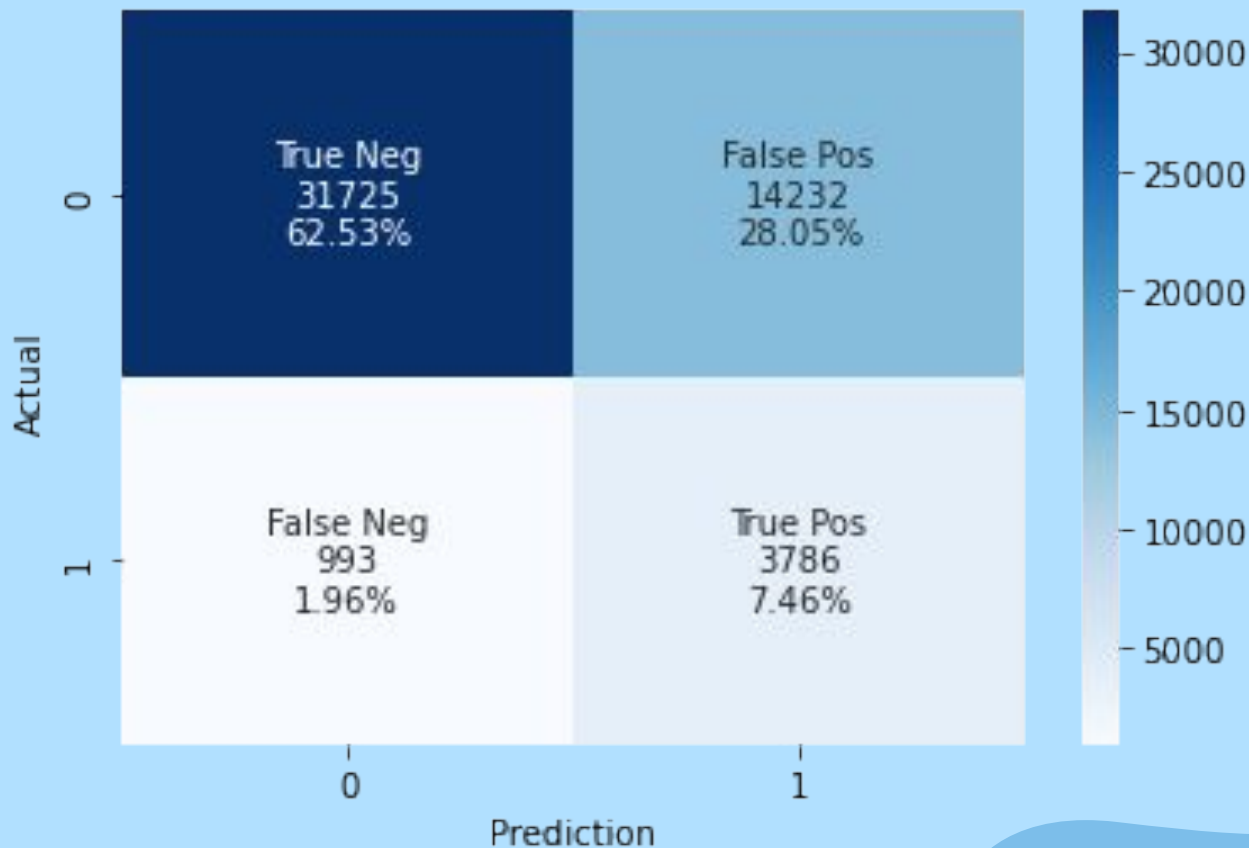


Logistic Regression w/ Class Weights

Cross Validation

Regularization:
C = 10

Class Weights:
0: 0.09, 1: 0.91

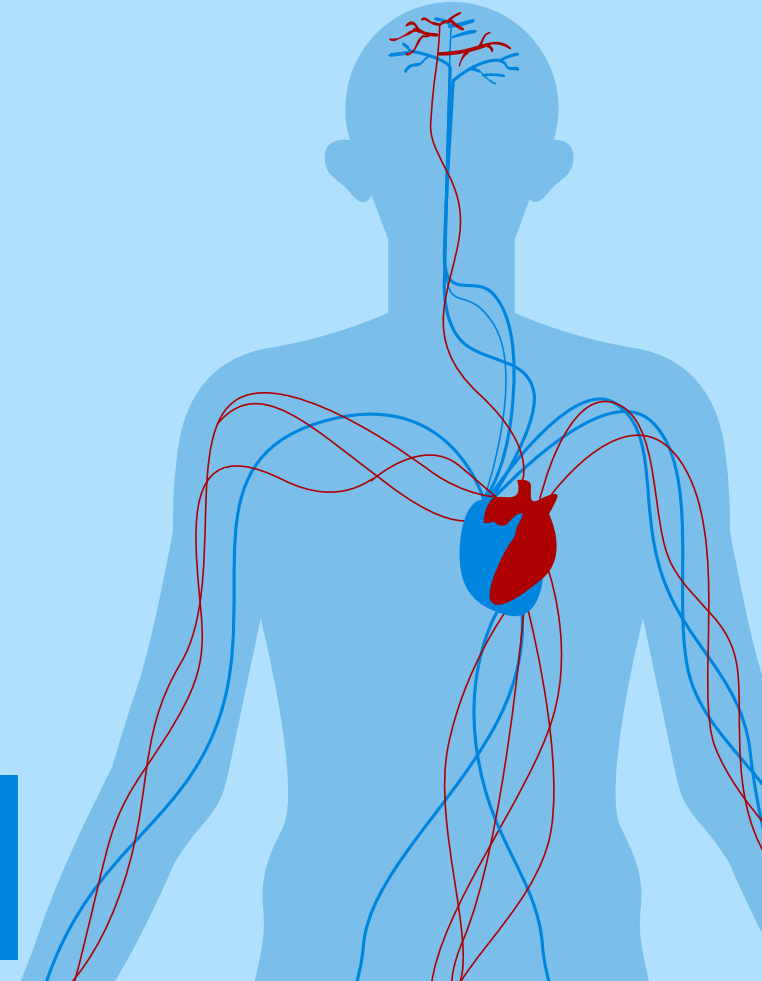


Thresholds

Thresholds	False Negatives	False Positives
$P \geq 0.25$	195 (0.38%)	27,619 (54.44%)
$P \geq 0.5$	993 (1.96%)	14,232 (28.05%)
$P \geq 0.75$	2,733 (5.39%)	4,659 (9.18%)

Lower Threshold = Decrease FN, Increase FP

** This lower threshold could help PREVENT future cases. If used for diagnostics, a higher threshold would be utilized.



StreamLit App

User Input:

Features to be used
in the Logistic
Regression model

Results:

Probability of heart
disease predicted
from inputs.

User Health Input

Have you been told by a health practitioner that you have high blood pressure?

no

Have you been told by a health practitioner that you have high cholesterol?

no

Have you had your cholesterol checked within the last 5 years?

yes

What is your BMI?

1 32 99

Have you smoked at least 100 cigarettes in your entire life?

yes

Have you been told by a health practitioner that you had a stroke?

yes

Have you been told by a health practitioner that you have diabetes?

yes

Heart Disease Prediction

This app predicts the probability of the presence of heart disease.

Instructions:

Answer the questions in the sidebar. Your probability of heart disease will be displayed.

Prediction

You may be at risk for heart disease. Please speak to your healthcare practitioner to discuss how to reduce your risk.


Prediction Probability

	0
0	0.6556

Important

The probability is not indicative of a diagnosis. This number is used as part of a Heart Disease prevention program to be established with your healthcare practitioner."

Made with Streamlit



To view and use the app please visit or follow the QR code:

https://share.streamlit.io/leahnagy/streamlit-heart-disease/main/proj_streamlit2.py



THANKS!

The code, slides, and more information can be found on my **GitHub** site at:

<https://github.com/leahnagy>

Connect on LinkedIn:

<https://www.linkedin.com/in/leah-j-nagy/>

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**