

Heart Disease Health Indicators

Abstract

The goal of this project is to use survey responses from the (BRFSS) Behavioral Risk Factor Surveillance System for predicting heart disease risk, preventive health practices and risk behaviors that are linked to chronic diseases in the adult population. Heart Disease is among the most prevalent chronic diseases in the United States, impacting millions of Americans each year and exerting a significant financial burden on the economy. In the United States alone, heart disease claims roughly 647,000 lives each year — making it the leading cause of death.

<https://github.com/maryam0003/Data>

Design

The Behavioral Risk Factor Surveillance System (BRFSS) is the nation's premier system of health-related telephone surveys that collect state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services. Use survey responses from the (BRFSS) Behavioral Risk Factor Surveillance System for predicting heart disease risk, preventive health practices and risk behaviors that are linked to chronic diseases in the adult population.

Data

This dataset contains 253,680 survey responses from BRFSS. Contains 22 features (High Blood Pressure, High Cholesterol, Cholesterol Check, Smoker, Stroke, Body Mass Indicator (BMI), Diabetes, Phys Activity, Fruits, Any Health care, No Docbc Cost, General Health, Mental Health, Diff Walk, Sex, Education, Income).

Algorithm

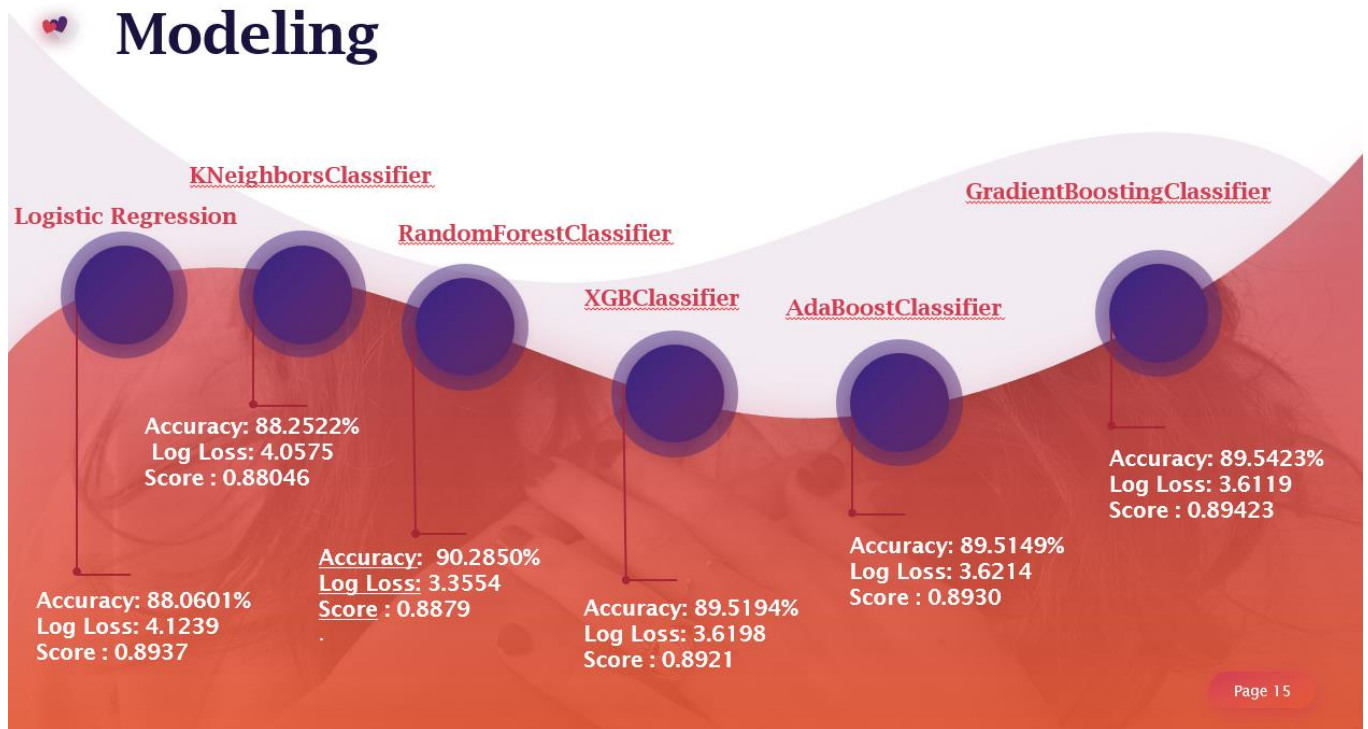
Model

- LogisticRegression
- KNeighborsClassifier
- RandomForestClassifier
- XGBClassifier

- AdaBoostClassifier
- GradientBoostingClassifier

Model Evaluation and Selection

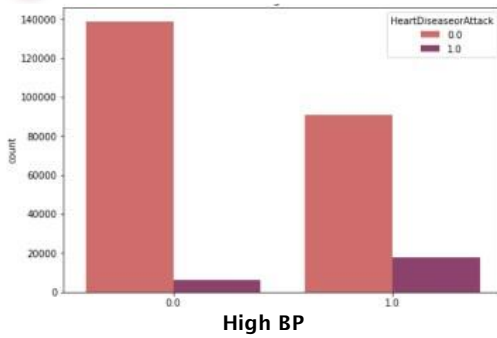
- The entire training dataset of 253,680 records was split into 80/20 train
- Random Forest Classifier gave me the best results and high accuracy



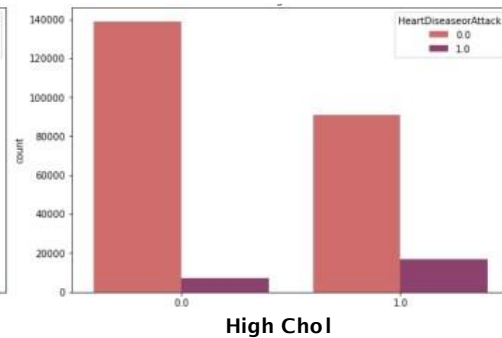
Tools

- Numpy and Pandas for data manipulation
- Scikit-learn for modeling
- Matplotlib and Seaborn for plotting
- Tableau for interactive visualizations

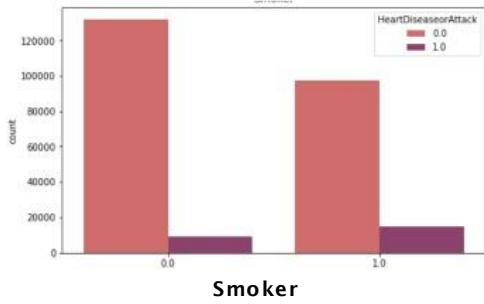
Communication



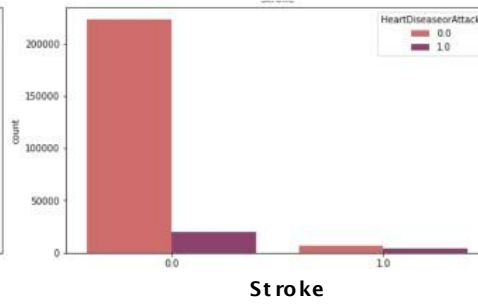
High BP



From the graph, it is clear to us from the sample of people that the rate of heart disease increases in people who suffer from high blood pressure and high cholesterol

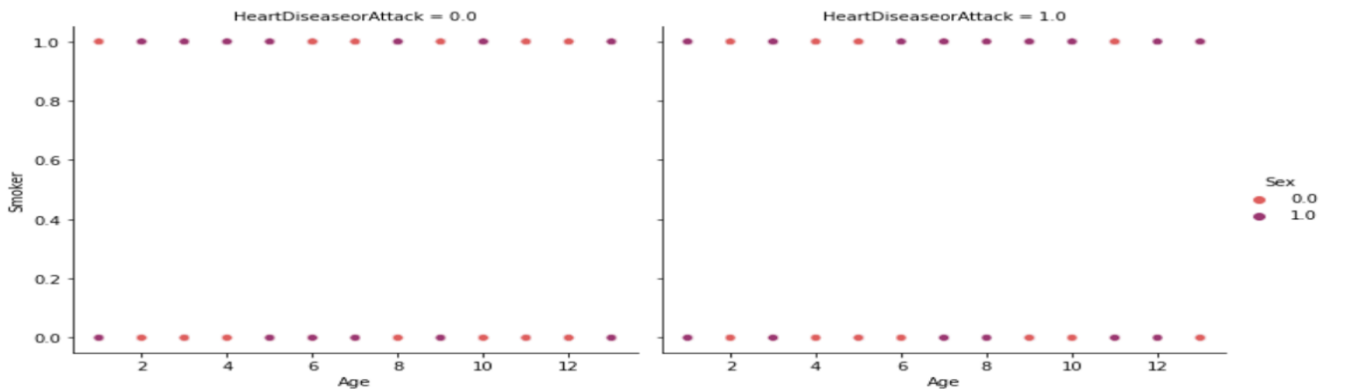


Smoker



Stroke

From the graph, the incidence of heart disease increases among smokers, in addition that the majority of people with strokes have heart disease.



From the graph heart disease is higher in men who smoke

