

# Predicting Metabolic Syndrome

**Raymond Garcia**

This is a Machine Learning case study with a dataset from data.world. I am attempting to predict Metabolic Syndrome with data gathered from multiple databases. This was a classification problem and was solved using various Statistical Learning algorithms.



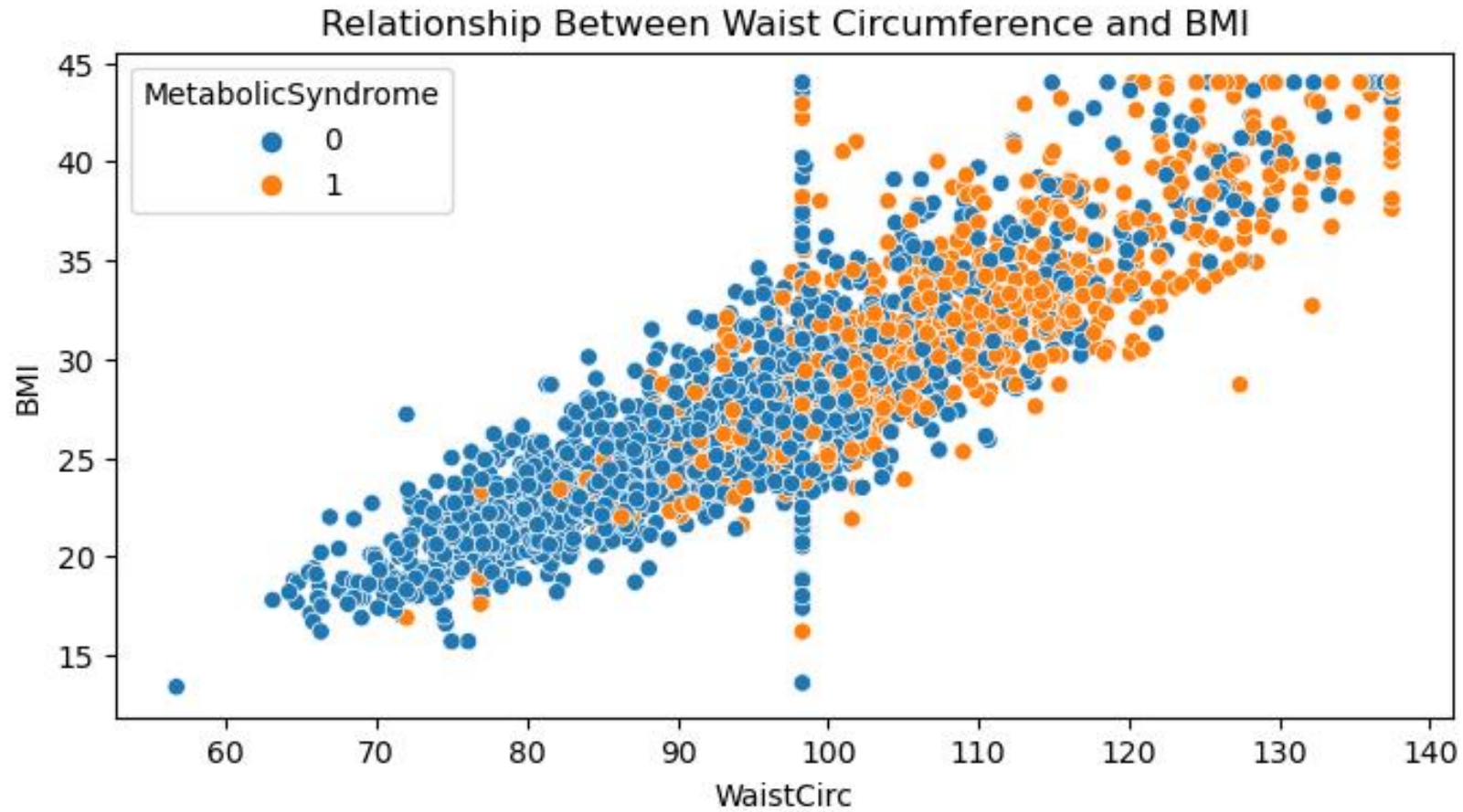
# Data Overview

---

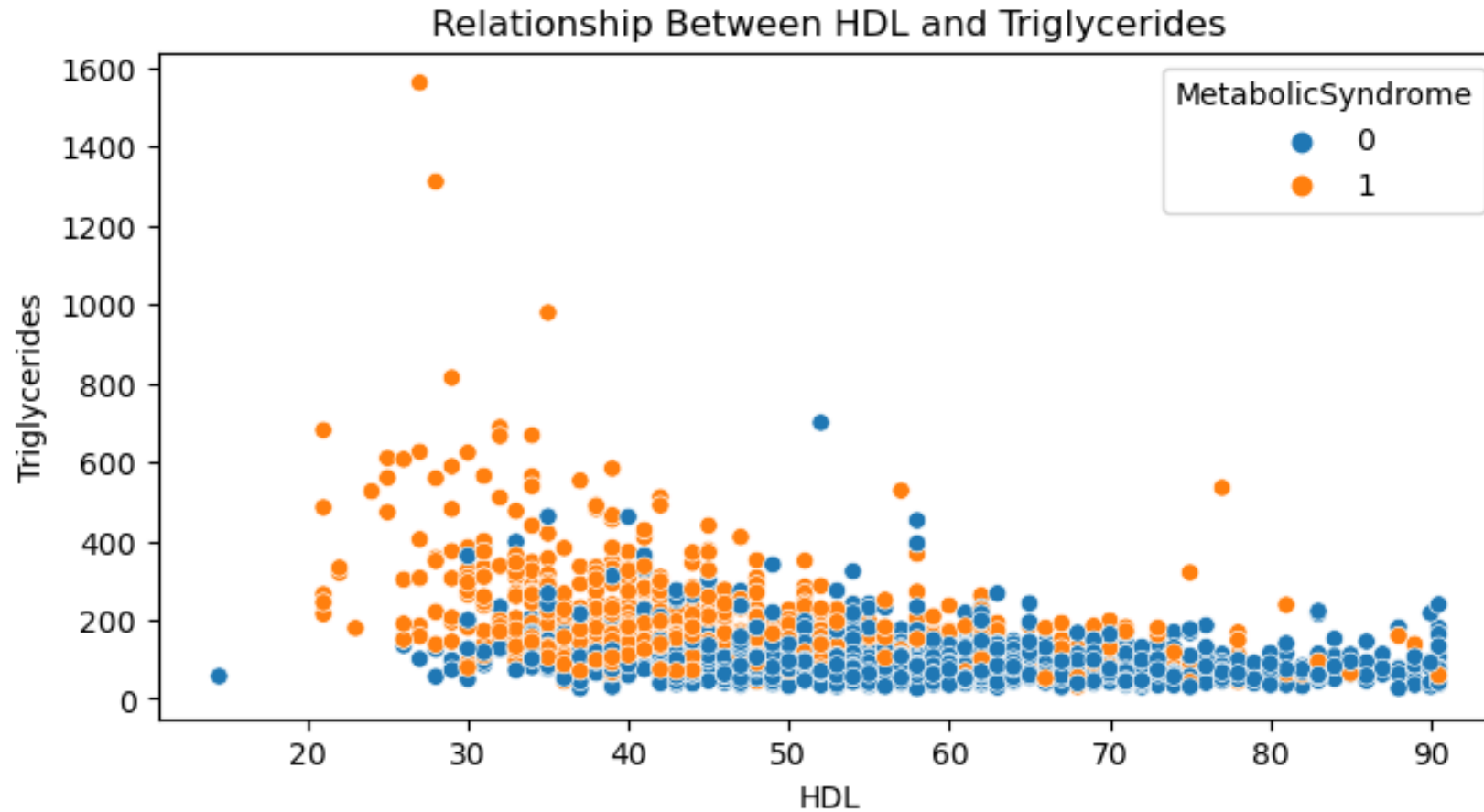
- The dataset has a variety of features for a patient that could be used to predict Metabolic Syndrome. The features are as follows:
  - Age
  - Sex
  - Marital Status
  - Income
  - Race
  - Waist Circumference
  - Body Mass Index
  - Albuminuria Categories
  - Urine Albumin-Creatinine Ratio
  - Uric Acid
  - Blood Glucose
  - HDL
  - Triglycerides
- It is composed of 13 columns and 2401 records. Each record represents a person and their corresponding bio markers.
- The Target was Metabolic Syndrome (yes/no)



# Waist Circumference, BMI, and Metabolic Syndrome



# HDL, Triglycerides, and Metabolic Syndrome







# Models

---

- Once Preprocessed, 4 models were fit to the feature matrix:
  - Logistic Regression
  - Random Forest Classifier
  - Light GBM Classifier
  - Light GBM Classifier with PCA and Feature Engineering



# Results

---

- Because this was a medical case study, I wanted to maximize Recall and Accuracy. Recall is important because it penalizes false negatives, and I don't want to let the target go undiagnosed.
- Best model was the Light GBM Classifier with:
  - Max Depth: 4
  - Number of Estimators: 200
  - Lambda Regularization: 20
- The model produced the following scores:
  - Accuracy: .8885
  - Recall: .8786
  - F1: .8438
  - AUC: .97



# Recommendations

---

- Implement the LightGBM model when diagnosing patients, or screening for patients at risk of developing Metabolic Syndrome.
- This model could also be implemented on routine checkups once lab work is turned in.
- Because of the high correlation between Waist Circumference, BMI, and Metabolic Syndrome, I suggest that any physician work with patients who have a high BMI and Waist Circumference to make lifestyle changes in order to reduce their odds of developing Metabolic Syndrome.