Predictions of Metabolic Syndrome

Kris Barbier

Stakeholders and Business Case

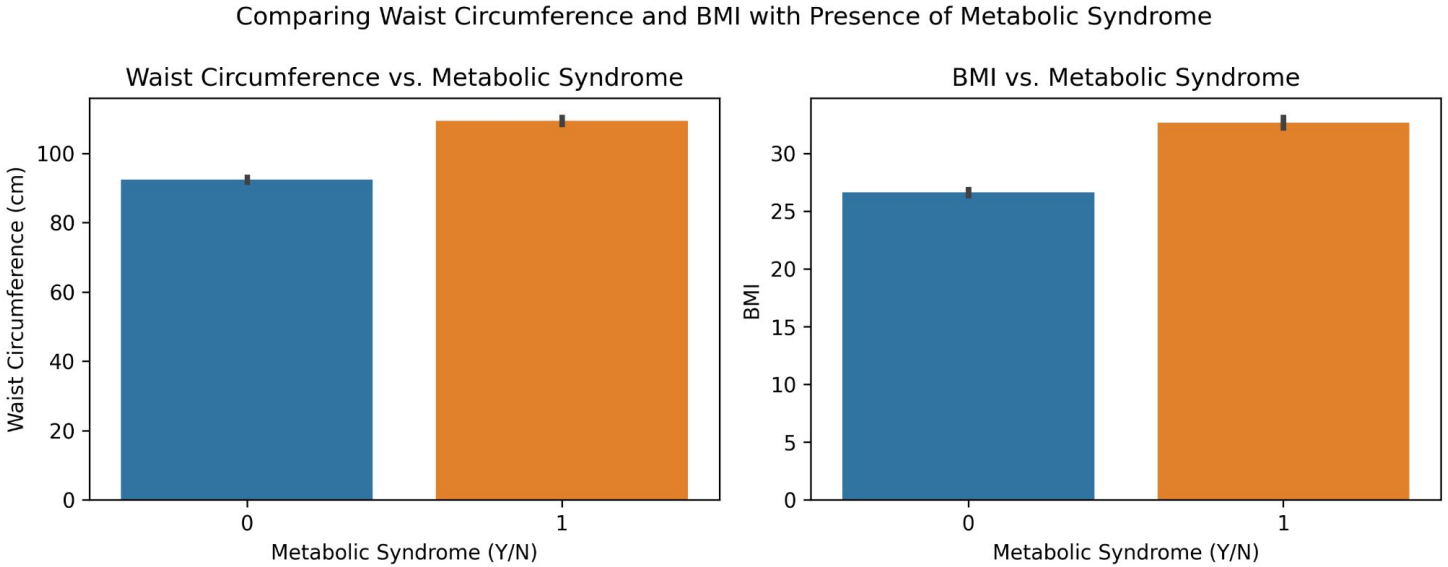
- Project Overview:
  - Stakeholders: Doctors/researchers looking into metabolic syndrome
  - This project aims to discover what health conditions are the most likely to lead to a diagnosis of metabolic syndrome in a patient, and what patients can do to lower their risk of developing metabolic syndrome.
- What is metabolic syndrome?
  - Metabolic syndrome is a group of health conditions that raise a person's risk of developing heart disease, diabetes, and other adverse health conditions.
  - Metabolic syndrome is formally diagnosed when a patient has 5 different markers: excess body fat around the waist, high blood pressure, high triglycerides, high blood glucose, and low HDL levels.

The Data Used

- The data used can be accessed here: [Metabolic Syndrome Prediction - dataset by informatics-edu | data.world]
- This data contained the following features:

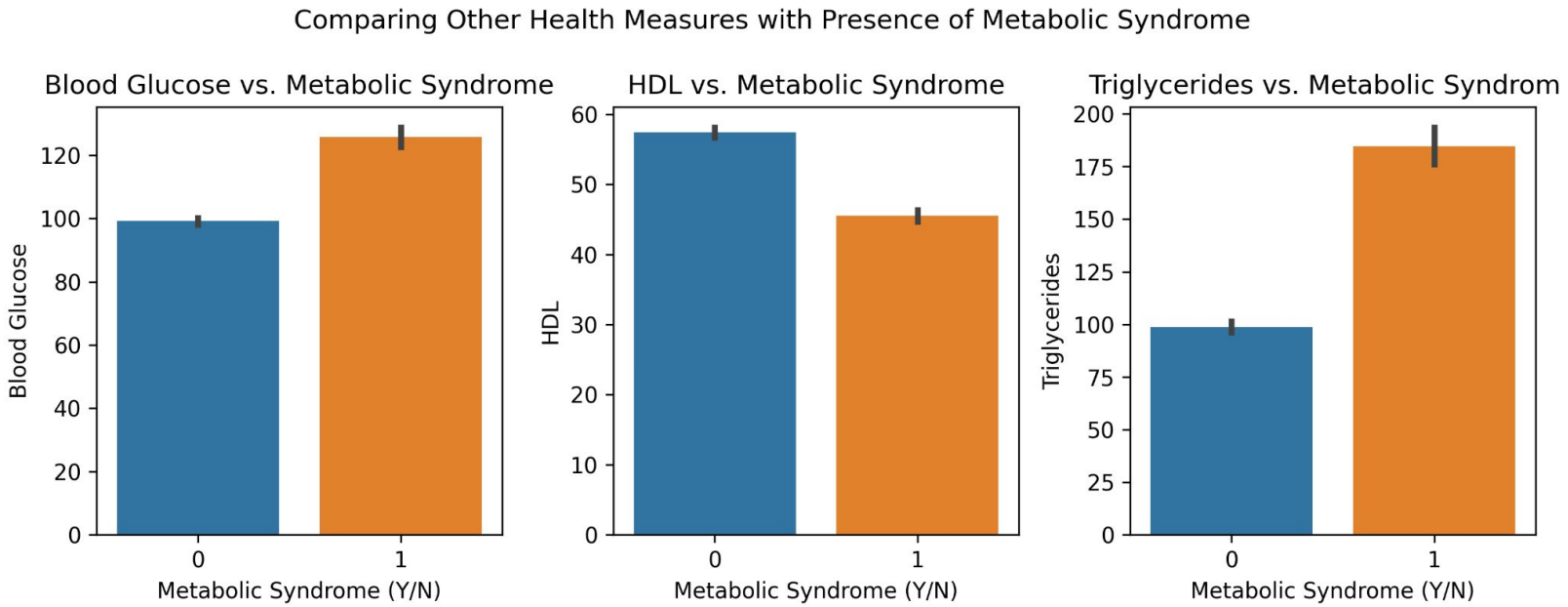
Feature	Description
Age	The patient's age.
Sex	The patient's gender.
Marital	The patient's marital status.
Income	The patient's monthly income.
Race	The patient's ethnicity.
WaistCirc	The patient's waist circumference.
BMI	The patient's BMI (Body Mass Index).
Albuminuria	The patient's Albuminuria stage.
UrAlbCr	The patient's Urine Albumin-Creatinine ratio.
UricAcid	The patient's uric acid level.
BloodGlucose	The patient's blood glucose level.
HDL	The patient's HDL (High-Density Lipoprotein) level.
Triglycerides	The patient's triglycerides level.
MetabolicSyndrome (target)	Whether the patient suffers from metabolic syndrome or not.

Visual 1



- In this visual, we see that, on average, patients with a higher average waist circumference and BMI were also more likely to be diagnosed with metabolic syndrome.

Visual 2



- In this visual, we see that the only marker that does not lead to a higher instance of metabolic syndrome is HDL levels.

Prediction Model

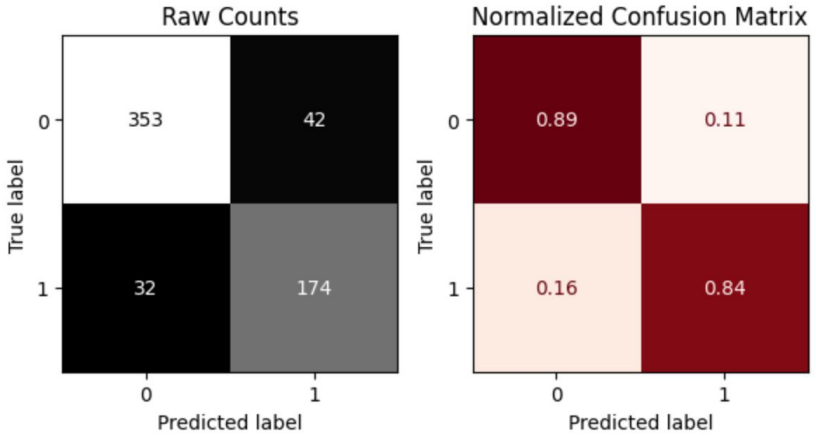
- The strongest prediction model was a tuned random forest model.
  - Strengths- The strengths of this model are that it was able to yield the highest accuracy rate with the least number of false negative predictions.
  - Weaknesses- The weaknesses of this model are that although the accuracy and false negative rates were better than other models, it still had 32 false negatives.
- False Negative Rates: In this case, false negatives need to be minimized as much as possible because it means that someone with metabolic syndrome was predicted not to have it, and they may go without proper treatment. This can lead to them developing much worse health conditions later on.

Tuned Random Forest Metrics

```
------------------------------------------------------------------------
 Classification Metrics: Test Data
------------------------------------------------------------------------
              precision    recall  f1-score   support

           0       0.92      0.89      0.91       395
           1       0.81      0.84      0.82       206

    accuracy                           0.88       601
   macro avg       0.86      0.87      0.86       601
weighted avg       0.88      0.88      0.88       601
```



- In the classification report, we find an accuracy rating of 88%, which means the model correctly identified 88% of cases correctly in the test data.
- The false negative count (far bottom left corner) is 32, which is still far from ideal for this case.
- Part of the cause of this could be due to imbalanced class numbers.

Further Steps

- In order to further improve the predictions from this project, I would try the following in the future:
  - Research and possibly remove outliers
  - Lower the decision threshold for the model
  - Try different class balancing strategies, such as random over/under sampling and SMOTE

Final Results

- In conclusion, we see that from this data, that the health measures most associated with developing metabolic syndrome include: high waist circumference, high BMI, high triglycerides, high blood glucose, and low HDL levels.
- For patients at risk of developing metabolic syndrome, it is advised that they maintain a healthy lifestyle through proper diet and exercise in order to keep these measure at the appropriate levels.

Further Information

- For further information, please contact me at:
  - krisbarbier02@gmail.com
- For a full view of this project, please visit:
  - https://github.com/krisbarbier/Prediction-of-Metabolic-Syndrome/tree/main