

Detecting Diabetes with Machine Learning



Ian Butler
In Association with the Center for Disease Control

Objective - Increase Accessibility to Diabetes Testing

Complications

- Heart Disease
- Vision Loss
- Lower-Limb Amputation
- Kidney Disease
- Stroke

Early Diagnosis

- Lifestyle Changes
- More Effective Treatment



Data - CDC's Behavioral Risk Factor Surveillance System

250,000 Records

- Reported Fields
 - Diet
 - Exercise
 - General Health
- Calculated Fields
 - BMI
 - High Cholesterol
 - High Blood Pressure



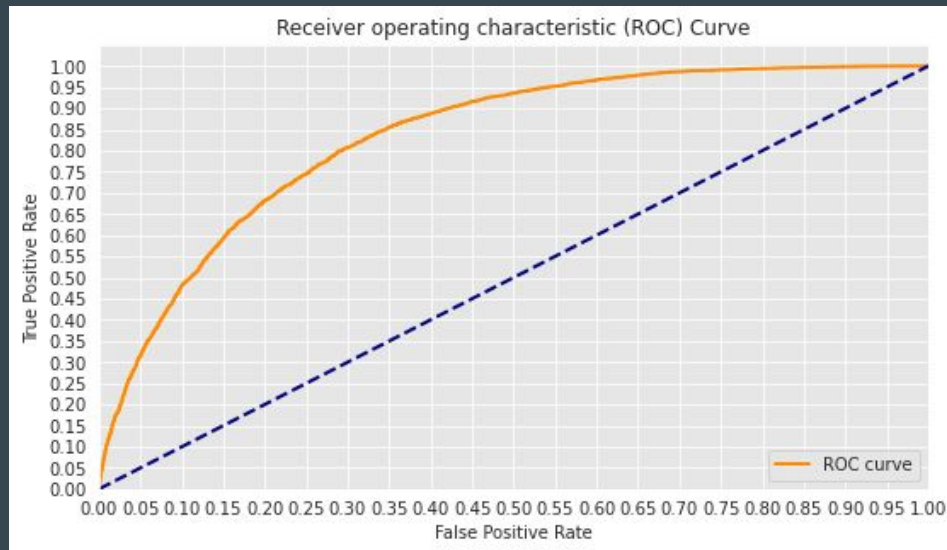
Results - Final Model and Performance

GBC 2

- Gradient Boosting Classifier
- Ensemble Method
- Regression Trees

Performance

- Accuracy: 75%
- Precision: 74%
- Recall: 80%
- F1: 76%
- AUC: 83%



Deployment - Streamlit

Share ☆ ☰

Are You Concerned About Having Diabetes?

After answering a few questions, this application will be able to provide you with a prediction of whether you have diabetes, as well as the probability that this prediction is correct.

Answer a few questions about yourself:

- ☐ Have you ever been told you have high blood pressure by a doctor, nurse, or other health professional?
- ☐ Have you ever been told you have high blood cholesterol by a doctor, nurse, or other health professional?
- ☐ Have you had your cholesterol checked within the last five years?

How tall are you?

Feet

4.00 4.00 6.00

Inches

0.00 0.00 11.00

How much do you weigh?

Pounds

0.00

< Manage app

Moving Forward - Improving the Application

Next steps:

- Reduce Dimensionality
 - Less Questions
 - Ease of Use
- Multiple Models
 - Variety of Questions
 - Flexibility



Thank you!

