Model Quality and Improvements

Project Deliverables

You will be required to submit:

- A GitHub repository with your project written in Python or R.

Instructions

Problem Statement

As a data professional working for a pharmaceutical company, you need to develop a model that predicts whether a patient will be diagnosed with diabetes. The model needs to have an accuracy score greater than 0.85.

You will be required to document the following steps:

- Data Importation
- Data Exploration
- Data Cleaning
- Data Preparation
- Data Modeling *(Using Decision Trees, Random Forest and Logistic Regression)*
- Model Evaluation
- Hyparameter Tuning
- Findings and Recommendations

Dataset

- Dataset URL: https://bit.ly/DiabetesDS

Project Source: https://bit.ly/3CU4b7d