**Experiment No 6**

**Title:** To Implement Decision tree classifier

**Tools Required: Anaconda Navigator**

**Concept :**

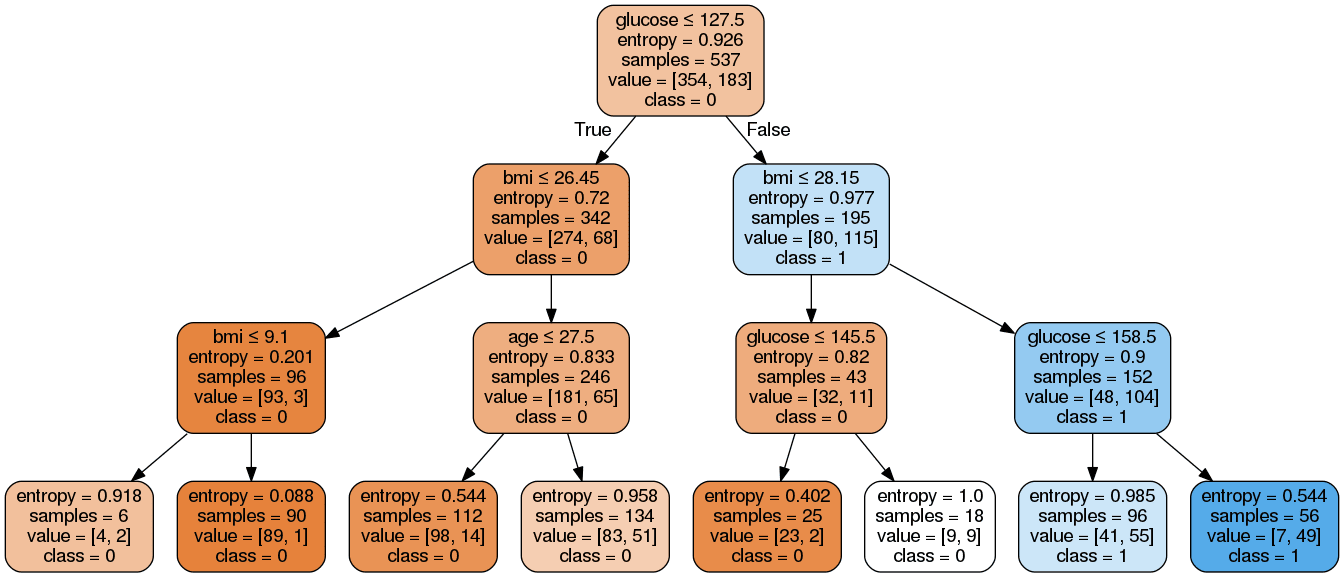
A decision tree is a simple representation for classifying examples. It is a supervised machine learning technique where the data is continuously split according to a certain parameter. Decision tree analysis can help solve both classification & regression problems. The decision tree algorithm breaks down a dataset into smaller subsets; while during the same time, an associated decision tree is incrementally developed.

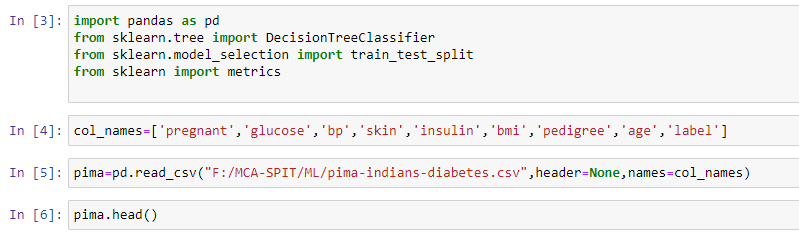
**Example Problem**

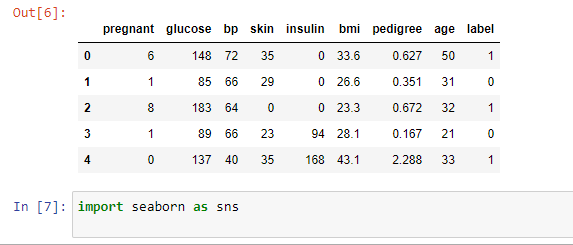
For our analysis, we have chosen a very relevant, and unique dataset which is applicable in the field of medical sciences, that will help predict whether or not a patient has diabetes, based on the variables captured in the dataset. This information has been sourced from the National Institute of Diabetes, Digestive and Kidney Diseases and includes predictor variables like a patient’s BMI, pregnancy details, insulin level, age, etc.

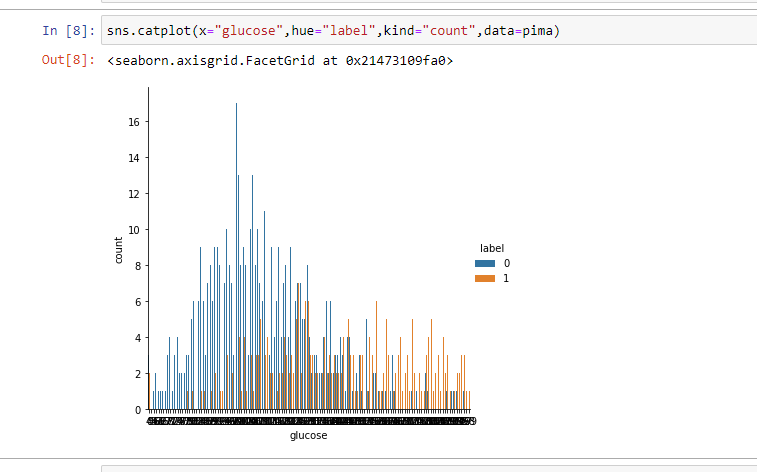
**Steps:**

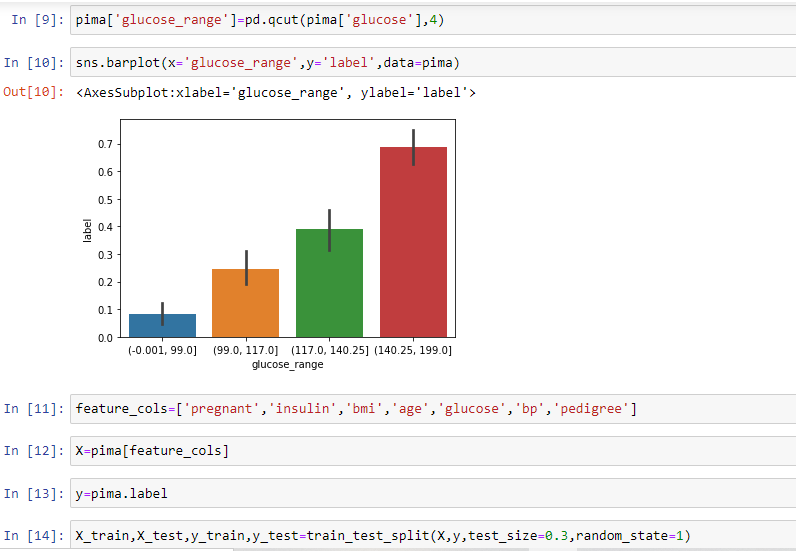
1. We import the required libraries for our decision tree analysis & pull in the required data
2. Perform EDA on the dataset.
3. After loading the data, we understand the structure & variables, determine the target & feature variables (dependent & independent variables respectively)
4. Splitting of data into training & testing sets in the ratio of 70:30.
5. Performing The decision tree analysis using scikit learn
   1. Create Decision Tree classifier object
   2. Train Decision Tree Classifier
   3. Predict the response for test dataset
6. We should estimate how accurately the classifier predicts the outcome. The accuracy is computed by comparing actual test set values and predicted values.
7. Visualize the decision tree - We have created a decision tree, let’s see what it looks like when we visualize it. The Scikit-learn’s export\_graphviz function can help visualize the decision tree
8. Optimization- Pruning/shortening a tree help in optimization. In our case, we will be varying the maximum depth of the tree as a control variable for pre-pruning, by trying max\_depth=3.

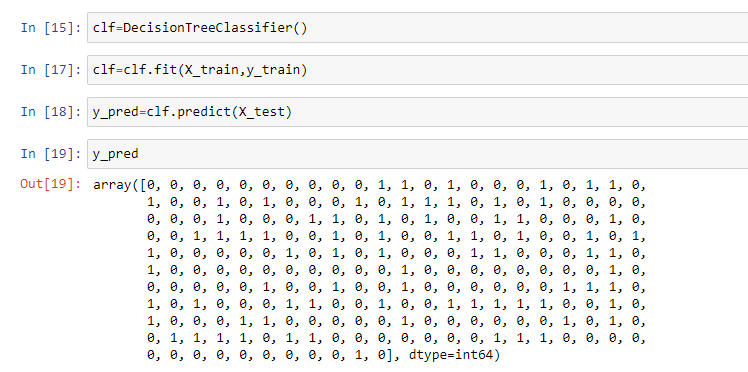


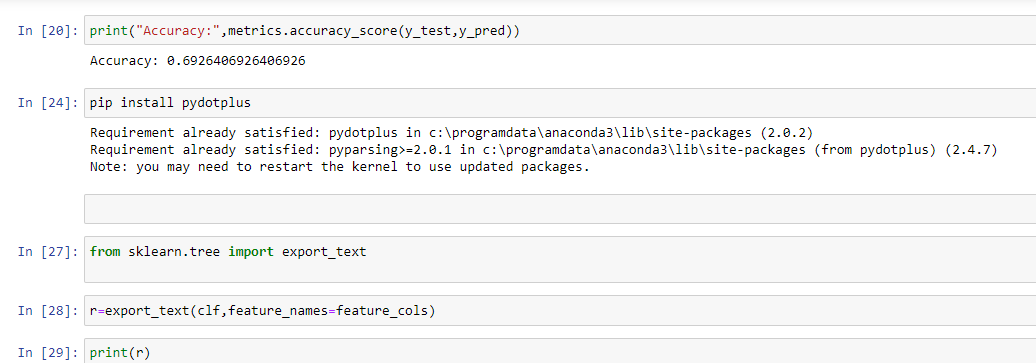












|--- glucose <= 129.50

| |--- bmi <= 26.30

| | |--- bmi <= 9.10

| | | |--- age <= 28.00

| | | | |--- class: 0

| | | |--- age > 28.00

| | | | |--- class: 1

| | |--- bmi > 9.10

| | | |--- pedigree <= 0.67

| | | | |--- class: 0

| | | |--- pedigree > 0.67

| | | | |--- pedigree <= 0.71

| | | | | |--- class: 1

| | | | |--- pedigree > 0.71

| | | | | |--- class: 0

| |--- bmi > 26.30

| | |--- age <= 27.50

| | | |--- bmi <= 45.40

| | | | |--- bp <= 12.00

| | | | | |--- class: 1

| | | | |--- bp > 12.00

| | | | | |--- pregnant <= 7.00

| | | | | | |--- pedigree <= 1.27

| | | | | | | |--- bmi <= 30.95

| | | | | | | | |--- class: 0

| | | | | | | |--- bmi > 30.95

| | | | | | | | |--- bp <= 53.00

| | | | | | | | | |--- pregnant <= 2.50

| | | | | | | | | | |--- pedigree <= 0.57

| | | | | | | | | | | |--- class: 0

| | | | | | | | | | |--- pedigree > 0.57

| | | | | | | | | | | |--- class: 1

| | | | | | | | | |--- pregnant > 2.50

| | | | | | | | | | |--- class: 1

| | | | | | | | |--- bp > 53.00

| | | | | | | | | |--- pedigree <= 0.65

| | | | | | | | | | |--- insulin <= 36.50

| | | | | | | | | | | |--- truncated branch of depth 5

| | | | | | | | | | |--- insulin > 36.50

| | | | | | | | | | | |--- class: 0

| | | | | | | | | |--- pedigree > 0.65

| | | | | | | | | | |--- insulin <= 65.50

| | | | | | | | | | | |--- class: 0

| | | | | | | | | | |--- insulin > 65.50

| | | | | | | | | | | |--- truncated branch of depth 2

| | | | | | |--- pedigree > 1.27

| | | | | | | |--- bmi <= 29.50

| | | | | | | | |--- class: 0

| | | | | | | |--- bmi > 29.50

| | | | | | | | |--- class: 1

| | | | | |--- pregnant > 7.00

| | | | | | |--- class: 1

| | | |--- bmi > 45.40

| | | | |--- age <= 22.50

| | | | | |--- class: 0

| | | | |--- age > 22.50

| | | | | |--- class: 1

| | |--- age > 27.50

| | | |--- pedigree <= 0.56

| | | | |--- glucose <= 101.50

| | | | | |--- bp <= 27.00

| | | | | | |--- class: 1

| | | | | |--- bp > 27.00

| | | | | | |--- age <= 42.50

| | | | | | | |--- class: 0

| | | | | | |--- age > 42.50

| | | | | | | |--- pedigree <= 0.38

| | | | | | | | |--- glucose <= 76.00

| | | | | | | | | |--- class: 0

| | | | | | | | |--- glucose > 76.00

| | | | | | | | | |--- age <= 57.50

| | | | | | | | | | |--- class: 1

| | | | | | | | | |--- age > 57.50

| | | | | | | | | | |--- class: 0

| | | | | | | |--- pedigree > 0.38

| | | | | | | | |--- class: 0

| | | | |--- glucose > 101.50

| | | | | |--- bp <= 67.00

| | | | | | |--- bp <= 58.00

| | | | | | | |--- bmi <= 30.05

| | | | | | | | |--- bp <= 25.00

| | | | | | | | | |--- class: 1

| | | | | | | | |--- bp > 25.00

| | | | | | | | | |--- class: 0

| | | | | | | |--- bmi > 30.05

| | | | | | | | |--- class: 0

| | | | | | |--- bp > 58.00

| | | | | | | |--- pedigree <= 0.43

| | | | | | | | |--- class: 1

| | | | | | | |--- pedigree > 0.43

| | | | | | | | |--- class: 0

| | | | | |--- bp > 67.00

| | | | | | |--- bmi <= 43.10

| | | | | | | |--- glucose <= 102.50

| | | | | | | | |--- class: 1

| | | | | | | |--- glucose > 102.50

| | | | | | | | |--- bmi <= 34.60

| | | | | | | | | |--- age <= 35.00

| | | | | | | | | | |--- class: 0

| | | | | | | | | |--- age > 35.00

| | | | | | | | | | |--- glucose <= 123.50

| | | | | | | | | | | |--- truncated branch of depth 5

| | | | | | | | | | |--- glucose > 123.50

| | | | | | | | | | | |--- class: 1

| | | | | | | | |--- bmi > 34.60

| | | | | | | | | |--- class: 0

| | | | | | |--- bmi > 43.10

| | | | | | | |--- class: 1

| | | |--- pedigree > 0.56

| | | | |--- pregnant <= 8.50

| | | | | |--- bp <= 87.00

| | | | | | |--- glucose <= 97.00

| | | | | | | |--- class: 0

| | | | | | |--- glucose > 97.00

| | | | | | | |--- glucose <= 116.50

| | | | | | | | |--- pedigree <= 1.39

| | | | | | | | | |--- class: 1

| | | | | | | | |--- pedigree > 1.39

| | | | | | | | | |--- class: 0

| | | | | | | |--- glucose > 116.50

| | | | | | | | |--- insulin <= 78.00

| | | | | | | | | |--- class: 0

| | | | | | | | |--- insulin > 78.00

| | | | | | | | | |--- glucose <= 123.50

| | | | | | | | | | |--- age <= 33.50

| | | | | | | | | | | |--- truncated branch of depth 2

| | | | | | | | | | |--- age > 33.50

| | | | | | | | | | | |--- class: 0

| | | | | | | | | |--- glucose > 123.50

| | | | | | | | | | |--- class: 1

| | | | | |--- bp > 87.00

| | | | | | |--- class: 0

| | | | |--- pregnant > 8.50

| | | | | |--- class: 1

|--- glucose > 129.50

| |--- bmi <= 27.85

| | |--- glucose <= 145.50

| | | |--- age <= 59.50

| | | | |--- class: 0

| | | |--- age > 59.50

| | | | |--- bmi <= 25.25

| | | | | |--- class: 1

| | | | |--- bmi > 25.25

| | | | | |--- class: 0

| | |--- glucose > 145.50

| | | |--- age <= 25.50

| | | | |--- class: 0

| | | |--- age > 25.50

| | | | |--- age <= 62.00

| | | | | |--- bmi <= 27.00

| | | | | | |--- bmi <= 22.15

| | | | | | | |--- class: 0

| | | | | | |--- bmi > 22.15

| | | | | | | |--- class: 1

| | | | | |--- bmi > 27.00

| | | | | | |--- bp <= 32.00

| | | | | | | |--- class: 1

| | | | | | |--- bp > 32.00

| | | | | | | |--- class: 0

| | | | |--- age > 62.00

| | | | | |--- class: 0

| |--- bmi > 27.85

| | |--- glucose <= 158.50

| | | |--- age <= 30.50

| | | | |--- bp <= 23.00

| | | | | |--- class: 1

| | | | |--- bp > 23.00

| | | | | |--- bp <= 88.00

| | | | | | |--- bp <= 72.00

| | | | | | | |--- bmi <= 33.75

| | | | | | | | |--- bmi <= 31.05

| | | | | | | | | |--- bp <= 58.00

| | | | | | | | | | |--- class: 0

| | | | | | | | | |--- bp > 58.00

| | | | | | | | | | |--- pregnant <= 4.50

| | | | | | | | | | | |--- class: 1

| | | | | | | | | | |--- pregnant > 4.50

| | | | | | | | | | | |--- class: 0

| | | | | | | | |--- bmi > 31.05

| | | | | | | | | |--- class: 0

| | | | | | | |--- bmi > 33.75

| | | | | | | | |--- pedigree <= 0.53

| | | | | | | | | |--- class: 1

| | | | | | | | |--- pedigree > 0.53

| | | | | | | | | |--- bmi <= 35.45

| | | | | | | | | | |--- class: 1

| | | | | | | | | |--- bmi > 35.45

| | | | | | | | | | |--- class: 0

| | | | | | |--- bp > 72.00

| | | | | | | |--- glucose <= 157.50

| | | | | | | | |--- bmi <= 43.40

| | | | | | | | | |--- class: 0

| | | | | | | | |--- bmi > 43.40

| | | | | | | | | |--- pregnant <= 5.00

| | | | | | | | | | |--- class: 1

| | | | | | | | | |--- pregnant > 5.00

| | | | | | | | | | |--- class: 0

| | | | | | | |--- glucose > 157.50

| | | | | | | | |--- class: 1

| | | | | |--- bp > 88.00

| | | | | | |--- class: 1

| | | |--- age > 30.50

| | | | |--- bmi <= 34.05

| | | | | |--- bp <= 75.00

| | | | | | |--- class: 0

| | | | | |--- bp > 75.00

| | | | | | |--- glucose <= 143.00

| | | | | | | |--- age <= 47.50

| | | | | | | | |--- bmi <= 33.00

| | | | | | | | | |--- insulin <= 147.50

| | | | | | | | | | |--- class: 1

| | | | | | | | | |--- insulin > 147.50

| | | | | | | | | | |--- pedigree <= 0.20

| | | | | | | | | | | |--- class: 1

| | | | | | | | | | |--- pedigree > 0.20

| | | | | | | | | | | |--- class: 0

| | | | | | | | |--- bmi > 33.00

| | | | | | | | | |--- class: 0

| | | | | | | |--- age > 47.50

| | | | | | | | |--- class: 0

| | | | | | |--- glucose > 143.00

| | | | | | | |--- class: 1

| | | | |--- bmi > 34.05

| | | | | |--- pedigree <= 1.09

| | | | | | |--- insulin <= 306.50

| | | | | | | |--- pedigree <= 0.22

| | | | | | | | |--- pregnant <= 6.50

| | | | | | | | | |--- class: 0

| | | | | | | | |--- pregnant > 6.50

| | | | | | | | | |--- class: 1

| | | | | | | |--- pedigree > 0.22

| | | | | | | | |--- class: 1

| | | | | | |--- insulin > 306.50

| | | | | | | |--- pregnant <= 6.50

| | | | | | | | |--- class: 1

| | | | | | | |--- pregnant > 6.50

| | | | | | | | |--- class: 0

| | | | | |--- pedigree > 1.09

| | | | | | |--- insulin <= 147.50

| | | | | | | |--- class: 0

| | | | | | |--- insulin > 147.50

| | | | | | | |--- class: 1

| | |--- glucose > 158.50

| | | |--- pedigree <= 1.16

| | | | |--- pedigree <= 0.34

| | | | | |--- age <= 48.50

| | | | | | |--- glucose <= 177.00

| | | | | | | |--- class: 1

| | | | | | |--- glucose > 177.00

| | | | | | | |--- pregnant <= 8.00

| | | | | | | | |--- glucose <= 187.50

| | | | | | | | | |--- class: 1

| | | | | | | | |--- glucose > 187.50

| | | | | | | | | |--- insulin <= 65.00

| | | | | | | | | | |--- class: 1

| | | | | | | | | |--- insulin > 65.00

| | | | | | | | | | |--- class: 0

| | | | | | | |--- pregnant > 8.00

| | | | | | | | |--- class: 0

| | | | | |--- age > 48.50

| | | | | | |--- bmi <= 34.85

| | | | | | | |--- class: 0

| | | | | | |--- bmi > 34.85

| | | | | | | |--- class: 1

| | | | |--- pedigree > 0.34

| | | | | |--- class: 1

| | | |--- pedigree > 1.16

| | | | |--- pregnant <= 3.50

| | | | | |--- pregnant <= 0.50

| | | | | | |--- class: 0

| | | | | |--- pregnant > 0.50

| | | | | | |--- class: 1

| | | | |--- pregnant > 3.50

| | | | | |--- class: 0

