**ML ASSIGNMENT – 2, Decision Tree**

**Part – A**

Decision Tree for monk data set 1

Diagram

Description automatically generated

TREE

+-- [SPLIT: x4 = 1 True]

| +-- [LABEL = 1]

+-- [SPLIT: x4 = 1 False]

| +-- [SPLIT: x0 = 1 True]

| | +-- [SPLIT: x1 = 1 True]

| | | +-- [LABEL = 1]

| | +-- [SPLIT: x1 = 1 False]

| | | +-- [LABEL = 0]

| +-- [SPLIT: x0 = 1 False]

| | +-- [SPLIT: x1 = 2 True]

| | | +-- [LABEL = 1]

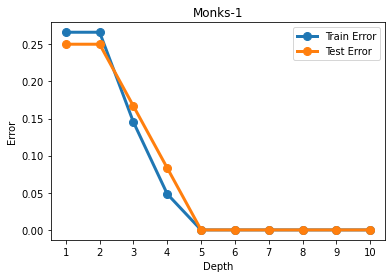
| | +-- [SPLIT: x1 = 2 False]

| | | +-- [LABEL = 0]

Test Error = 16.67%.

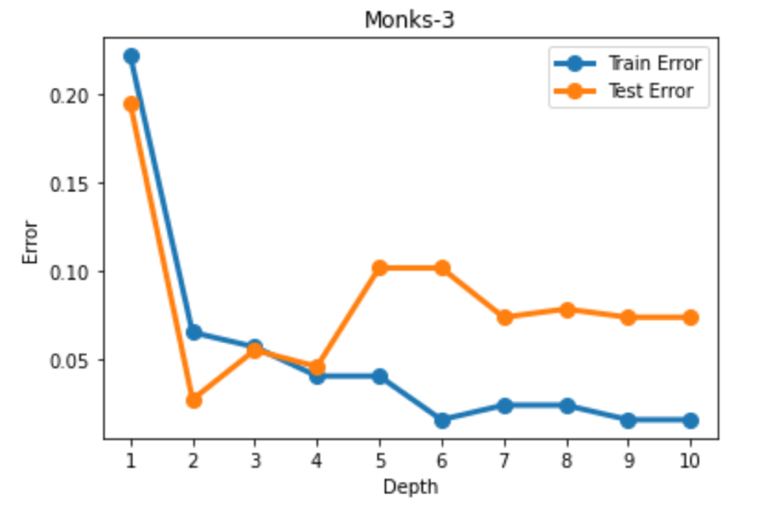
**Part – B**

Training and testing errors for the above data set



Chart, line chart

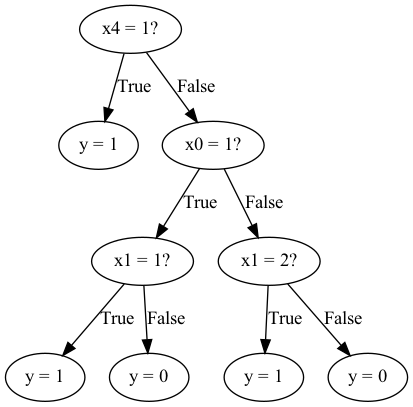
Description automatically generated



**Part – C**

Decision trees for depth 1, 3, and 5 on Monk data set 1

Diagram

Description automatically generated

A picture containing shape

Description automatically generated

Pretty tree for the following depths,

Depth = 1

TREE

+-- [SPLIT: x4 = 1 True]

| +-- [LABEL = 1]

+-- [SPLIT: x4 = 1 False]

| +-- [LABEL = 0]

Test Error = 25.00%.

Confusion Matrix for depth 1

[[216 0]

[108 108]]

Depth = 3

TREE

+-- [SPLIT: x4 = 1 True]

| +-- [LABEL = 1]

+-- [SPLIT: x4 = 1 False]

| +-- [SPLIT: x0 = 1 True]

| | +-- [SPLIT: x1 = 1 True]

| | | +-- [LABEL = 1]

| | +-- [SPLIT: x1 = 1 False]

| | | +-- [LABEL = 0]

| +-- [SPLIT: x0 = 1 False]

| | +-- [SPLIT: x1 = 2 True]

| | | +-- [LABEL = 1]

| | +-- [SPLIT: x1 = 2 False]

| | | +-- [LABEL = 0]

Test Error = 16.67%.

Confusion Matrix for depth 3

[[180 36]

[ 36 180]]

Depth = 5

TREE

+-- [SPLIT: x4 = 1 True]

| +-- [LABEL = 1]

+-- [SPLIT: x4 = 1 False]

| +-- [SPLIT: x0 = 1 True]

| | +-- [SPLIT: x1 = 1 True]

| | | +-- [LABEL = 1]

| | +-- [SPLIT: x1 = 1 False]

| | | +-- [LABEL = 0]

| +-- [SPLIT: x0 = 1 False]

| | +-- [SPLIT: x1 = 2 True]

| | | +-- [SPLIT: x0 = 2 True]

| | | | +-- [LABEL = 1]

| | | +-- [SPLIT: x0 = 2 False]

| | | | +-- [LABEL = 0]

| | +-- [SPLIT: x1 = 2 False]

| | | +-- [SPLIT: x0 = 3 True]

| | | | +-- [SPLIT: x1 = 3 True]

| | | | | +-- [LABEL = 1]

| | | | +-- [SPLIT: x1 = 3 False]

| | | | | +-- [LABEL = 0]

| | | +-- [SPLIT: x0 = 3 False]

| | | | +-- [LABEL = 0]

Test Error = 0.00%.

Confusion Matrix for depth 5

[[216 0]

[ 0 216]]

**Part – D**

Decision trees learned using scikit-learns decision classifier function

Depth = 1

|--- feature\_4 <= 1.50

| |--- class: 1

|--- feature\_4 > 1.50

| |--- class: 0

Test Error = 25.00%.

Confusion Matrix for depth 1

[[216 0]

[108 108]]

Depth = 3

|--- feature\_4 <= 1.50

| |--- class: 1

|--- feature\_4 > 1.50

| |--- feature\_0 <= 1.50

| | |--- feature\_1 <= 1.50

| | | |--- class: 1

| | |--- feature\_1 > 1.50

| | | |--- class: 0

| |--- feature\_0 > 1.50

| | |--- feature\_1 <= 1.50

| | | |--- class: 0

| | |--- feature\_1 > 1.50

| | | |--- class: 1

Test Error = 16.67%.

Confusion Matrix for depth 3

[[144 72]

[ 0 216]]

Depth = 5

|--- feature\_4 <= 1.50

| |--- class: 1

|--- feature\_4 > 1.50

| |--- feature\_0 <= 1.50

| | |--- feature\_1 <= 1.50

| | | |--- class: 1

| | |--- feature\_1 > 1.50

| | | |--- class: 0

| |--- feature\_0 > 1.50

| | |--- feature\_1 <= 1.50

| | | |--- class: 0

| | |--- feature\_1 > 1.50

| | | |--- feature\_3 <= 1.50

| | | | |--- feature\_5 <= 1.50

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

| | | | |--- feature\_5 > 1.50

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

| | | |--- feature\_3 > 1.50

| | | | |--- feature\_5 <= 1.50

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

| | | | |--- feature\_5 > 1.50

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

Test Error = 16.67%.

Confusion Matrix for depth 5

[[168 48]

[ 24 192]]

**Part – E**

New breast cancer data from UCI repository has been used for this part

Depth-1

Diagram

Description automatically generated

Printing decision tree from above written method for depth 1

TREE

+-- [SPLIT: x4 = 0 True]

| +-- [LABEL = 0]

+-- [SPLIT: x4 = 0 False]

| +-- [LABEL = 1]

Test Error = 23.29%.

Printing decision tree by scikitlearns decision tree classifier for depth 1

|--- feature\_4 <= 0.50

| |--- class: 0

|--- feature\_4 > 0.50

| |--- class: 1

Test Error = 23.29%.

Confusion Matrix for depth 1 for decision tree using our own method

[[47 12]

[ 5 9]]

Confusion Matrix for depth 1 for decision tree using scikit learn

[[47 12]

[ 5 9]]

Depth-3

Diagram

Description automatically generated

Printing decision tree from above written method for depth 3

TREE

+-- [SPLIT: x4 = 0 True]

| +-- [SPLIT: x3 = 10 True]

| | +-- [SPLIT: x7 = 0 True]

| | | +-- [LABEL = 0]

| | +-- [SPLIT: x7 = 0 False]

| | | +-- [LABEL = 1]

| +-- [SPLIT: x3 = 10 False]

| | +-- [SPLIT: x1 = 3 True]

| | | +-- [LABEL = 0]

| | +-- [SPLIT: x1 = 3 False]

| | | +-- [LABEL = 0]

+-- [SPLIT: x4 = 0 False]

| +-- [SPLIT: x7 = 1 True]

| | +-- [SPLIT: x0 = 0 True]

| | | +-- [LABEL = 1]

| | +-- [SPLIT: x0 = 0 False]

| | | +-- [LABEL = 0]

| +-- [SPLIT: x7 = 1 False]

| | +-- [SPLIT: x1 = 2 True]

| | | +-- [LABEL = 0]

| | +-- [SPLIT: x1 = 2 False]

| | | +-- [LABEL = 1]

Test Error = 28.77%.

Printing decision tree by scikitlearns decision tree classifier for depth 3

|--- feature\_4 <= 0.50

| |--- feature\_3 <= 7.50

| | |--- feature\_6 <= 0.50

| | | |--- class: 0

| | |--- feature\_6 > 0.50

| | | |--- class: 0

| |--- feature\_3 > 7.50

| | |--- feature\_8 <= 3.50

| | | |--- class: 1

| | |--- feature\_8 > 3.50

| | | |--- class: 0

|--- feature\_4 > 0.50

| |--- feature\_7 <= 0.50

| | |--- feature\_3 <= 4.50

| | | |--- class: 1

| | |--- feature\_3 > 4.50

| | | |--- class: 1

| |--- feature\_7 > 0.50

| | |--- feature\_0 <= 0.50

| | | |--- class: 1

| | |--- feature\_0 > 0.50

| | | |--- class: 0

Test Error = 28.77%.

Confusion Matrix for depth 3 for decision tree using our own method

[[47 12]

[ 9 5]]

Confusion Matrix for depth 3 for decision tree using scikit learn

[[45 14]

[ 7 7]]

Depth-5

Diagram

Description automatically generated

Printing decision tree from above written method for depth 5

TREE

+-- [SPLIT: x4 = 0 True]

| +-- [SPLIT: x3 = 10 True]

| | +-- [SPLIT: x7 = 0 True]

| | | +-- [LABEL = 0]

| | +-- [SPLIT: x7 = 0 False]

| | | +-- [LABEL = 1]

| +-- [SPLIT: x3 = 10 False]

| | +-- [SPLIT: x1 = 3 True]

| | | +-- [SPLIT: x5 = 2 True]

| | | | +-- [SPLIT: x3 = 2 True]

| | | | | +-- [LABEL = 1]

| | | | +-- [SPLIT: x3 = 2 False]

| | | | | +-- [LABEL = 0]

| | | +-- [SPLIT: x5 = 2 False]

| | | | +-- [SPLIT: x0 = 0 True]

| | | | | +-- [LABEL = 0]

| | | | +-- [SPLIT: x0 = 0 False]

| | | | | +-- [LABEL = 0]

| | +-- [SPLIT: x1 = 3 False]

| | | +-- [SPLIT: x3 = 8 True]

| | | | +-- [LABEL = 1]

| | | +-- [SPLIT: x3 = 8 False]

| | | | +-- [SPLIT: x5 = 0 True]

| | | | | +-- [LABEL = 1]

| | | | +-- [SPLIT: x5 = 0 False]

| | | | | +-- [LABEL = 0]

+-- [SPLIT: x4 = 0 False]

| +-- [SPLIT: x7 = 1 True]

| | +-- [SPLIT: x0 = 1 True]

| | | +-- [SPLIT: x1 = 1 True]

| | | | +-- [LABEL = 1]

| | | +-- [SPLIT: x1 = 1 False]

| | | | +-- [SPLIT: x3 = 3 True]

| | | | | +-- [LABEL = 0]

| | | | +-- [SPLIT: x3 = 3 False]

| | | | | +-- [LABEL = 0]

| | +-- [SPLIT: x0 = 1 False]

| | | +-- [SPLIT: x3 = 5 True]

| | | | +-- [LABEL = 0]

| | | +-- [SPLIT: x3 = 5 False]

| | | | +-- [SPLIT: x3 = 7 True]

| | | | | +-- [LABEL = 1]

| | | | +-- [SPLIT: x3 = 7 False]

| | | | | +-- [LABEL = 1]

| +-- [SPLIT: x7 = 1 False]

| | +-- [SPLIT: x1 = 2 True]

| | | +-- [SPLIT: x4 = 1 True]

| | | | +-- [LABEL = 1]

| | | +-- [SPLIT: x4 = 1 False]

| | | | +-- [LABEL = 0]

| | +-- [SPLIT: x1 = 2 False]

| | | +-- [SPLIT: x3 = 6 True]

| | | | +-- [LABEL = 0]

| | | +-- [SPLIT: x3 = 6 False]

| | | | +-- [SPLIT: x3 = 4 True]

| | | | | +-- [LABEL = 1]

| | | | +-- [SPLIT: x3 = 4 False]

| | | | | +-- [LABEL = 1]

Test Error = 20.55%.

Printing decision tree by scikitlearns decision tree classifier for depth 5

|--- feature\_4 <= 0.50

| |--- feature\_3 <= 7.50

| | |--- feature\_6 <= 0.50

| | | |--- feature\_1 <= 2.50

| | | | |--- feature\_8 <= 2.50

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

| | | | |--- feature\_8 > 2.50

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

| | | |--- feature\_1 > 2.50

| | | | |--- class: 0

| | |--- feature\_6 > 0.50

| | | |--- feature\_5 <= 0.50

| | | | |--- class: 1

| | | |--- feature\_5 > 0.50

| | | | |--- feature\_5 <= 1.50

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

| | | | |--- feature\_5 > 1.50

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

| |--- feature\_3 > 7.50

| | |--- feature\_8 <= 3.50

| | | |--- feature\_2 <= 1.00

| | | | |--- feature\_8 <= 2.50

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

| | | | |--- feature\_8 > 2.50

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

| | | |--- feature\_2 > 1.00

| | | | |--- class: 1

| | |--- feature\_8 > 3.50

| | | |--- class: 0

|--- feature\_4 > 0.50

| |--- feature\_7 <= 0.50

| | |--- feature\_3 <= 4.50

| | | |--- feature\_0 <= 0.50

| | | | |--- feature\_3 <= 3.50

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

| | | | |--- feature\_3 > 3.50

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

| | | |--- feature\_0 > 0.50

| | | | |--- class: 1

| | |--- feature\_3 > 4.50

| | | |--- feature\_4 <= 5.50

| | | | |--- feature\_3 <= 6.50

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

| | | | |--- feature\_3 > 6.50

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

| | | |--- feature\_4 > 5.50

| | | | |--- class: 1

| |--- feature\_7 > 0.50

| | |--- feature\_0 <= 0.50

| | | |--- feature\_3 <= 5.50

| | | | |--- feature\_3 <= 4.50

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

| | | | |--- feature\_3 > 4.50

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

| | | |--- feature\_3 > 5.50

| | | | |--- class: 1

| | |--- feature\_0 > 0.50

| | | |--- feature\_1 <= 1.50

| | | | |--- class: 1

| | | |--- feature\_1 > 1.50

| | | | |--- feature\_3 <= 3.50

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

| | | | |--- feature\_3 > 3.50

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

Test Error = 20.55%.

Confusion Matrix for depth 5 for decision tree using our own method

[[50 9]

[ 6 8]]

Confusion Matrix for depth 5 for decision tree using scikit learn

[[50 9]

[ 6 8]]

**Discussion:**

Implementation of Decision Trees using ID3 Algorithm has been done on Monk data set. ID3 algo uses the maximum information gain to choose the best attribute. ID3 works best for categorical variables, whereas sciket-learn does not support categorical variables. The results are similar for ID3 algo and scikit learn with minute differences in error.