

## Programming Assignment 2

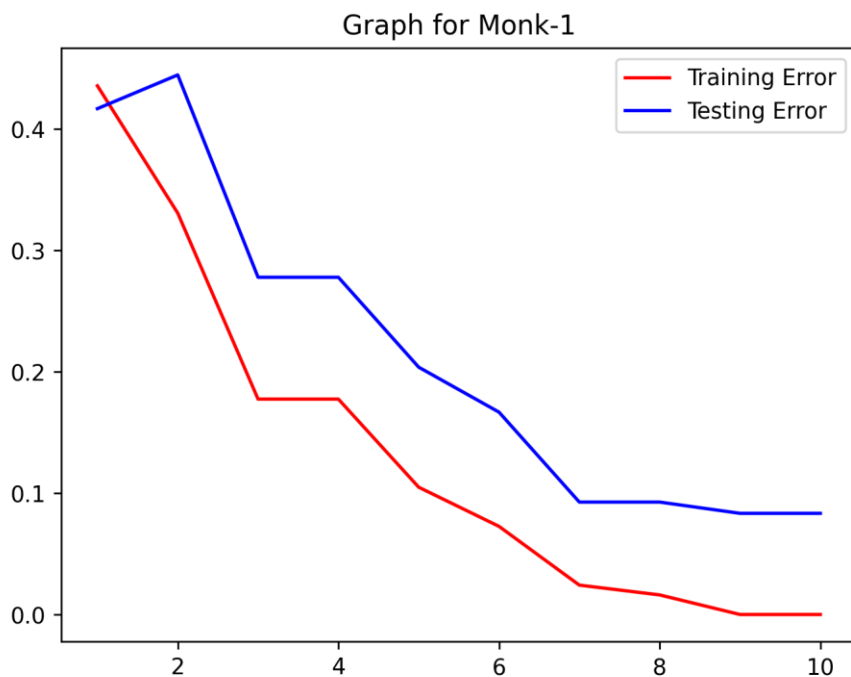
Members:

Erik Hale (netid: emh170004)

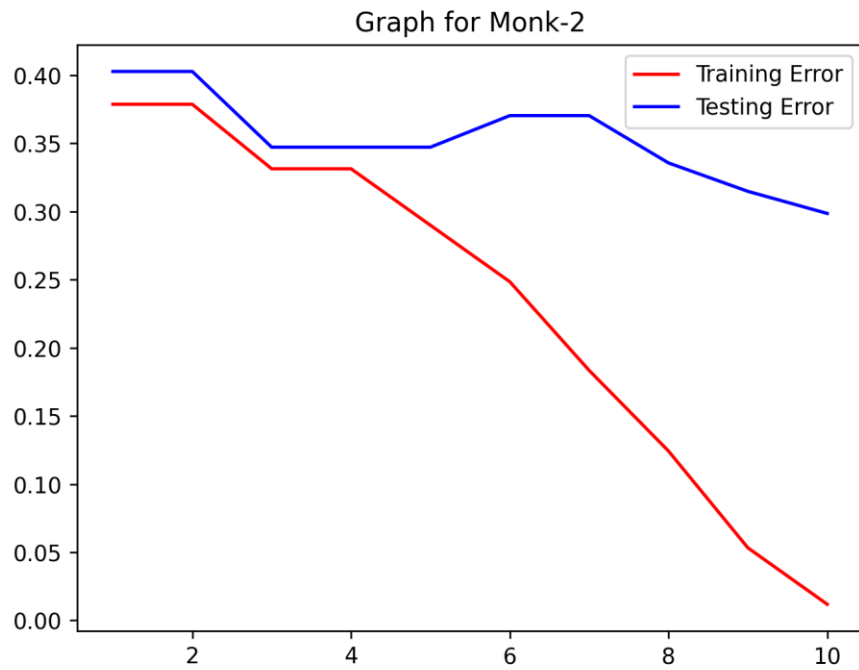
Shiva Kumar( netid: sak220007)

Part a)

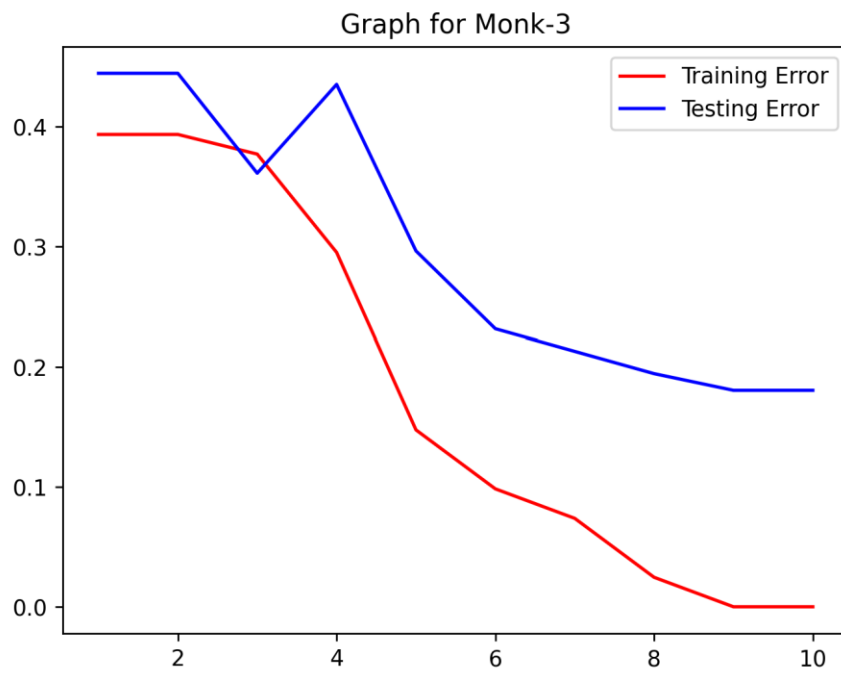
Monk-1:



## Monk-2



# Monk-3



## Part b)

### Monk-1 Depth 1

```
Monk 1 Dataset
TREE
+-- [SPLIT: x4 = 3]
|   +-- [LABEL = 1]
+-- [SPLIT: x4 = 3]
|   +-- [LABEL = 0]
TP: 180 | FN: 36
-----
FP: 144 | TN: 72

TPR: 0.8333333333333334 % | FNR: 0.16666666666666666 %
-----
FPR: 0.6666666666666666 % | TNR: 0.3333333333333333 %
```

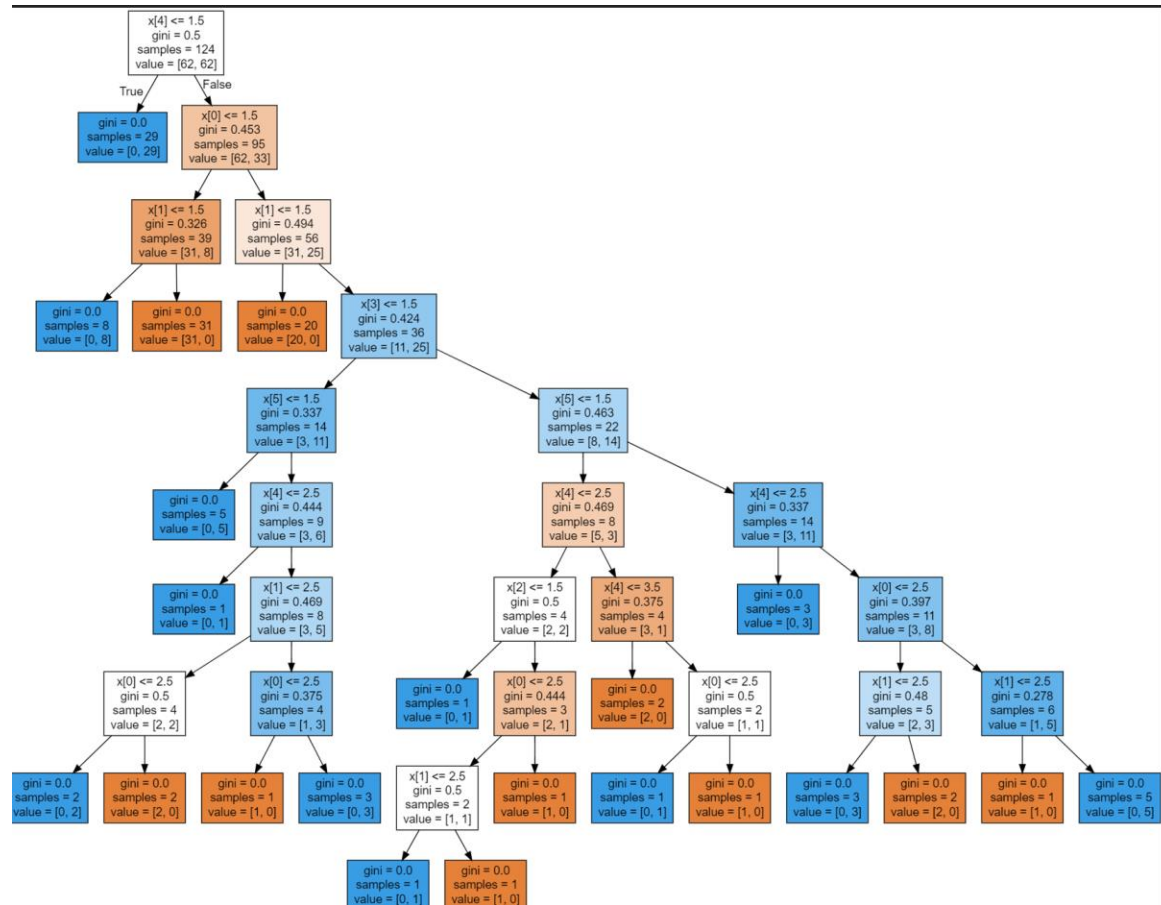
### Monk-2 Depth 2

```
TREE
+-- [SPLIT: x4 = 3]
|   +-- [SPLIT: x0 = 1]
|   |   +-- [LABEL = 1]
|   |   +-- [SPLIT: x0 = 1]
|   |   |   +-- [LABEL = 0]
+-- [SPLIT: x4 = 3]
|   +-- [SPLIT: x0 = 1]
|   |   +-- [LABEL = 0]
|   |   +-- [SPLIT: x0 = 1]
|   |   |   +-- [LABEL = 0]
TP: 120 | FN: 96
-----
FP: 96 | TN: 120

TPR: 0.5555555555555556 % | FNR: 0.4444444444444444 %
-----
FPR: 0.4444444444444444 % | TNR: 0.5555555555555556 %
```

Note: png image created by graphviz are in graphviz\_dataset\_[dataset name] folders.

## Monk-1 Scikit -learn Tree:



## Monk-1 Scikit -learn Confusion Matrix:

```
Test Error for SciKit Learn: 10.19%.
TP:  204  |  FN:  12
-----
FP:  32   |  TN: 184

TPR:  0.9444444444444444 % |  FNR:  0.0555555555555555 %
-----
FPR:  0.14814814814814814 % |  TNR:  0.8518518518518519 %
```

Part d)

Datasets 2 and 3 were found from the UCI Repository.

Dataset 2:

Given 4 attributes, determine whether an individual donated blood:

SPECT Heart Data Set

<https://archive.ics.uci.edu/ml/datasets/SPECT+Heart>

Dataset 2: Part b)

Depth 1

```
SPEC Data Set
TREE
+-- [SPLIT: x0 = 0]
|   +-- [LABEL = 1]
+-- [SPLIT: x0 = 0]
|   +-- [LABEL = 0]
TP:  87 | FN:  85
-----
FP:   3 | TN:  12

TPR:  0.5058139534883721 % | FNR:  0.4941860465116279 %
-----
FPR:  0.2 % | TNR:  0.8 %
```

## Depth 2

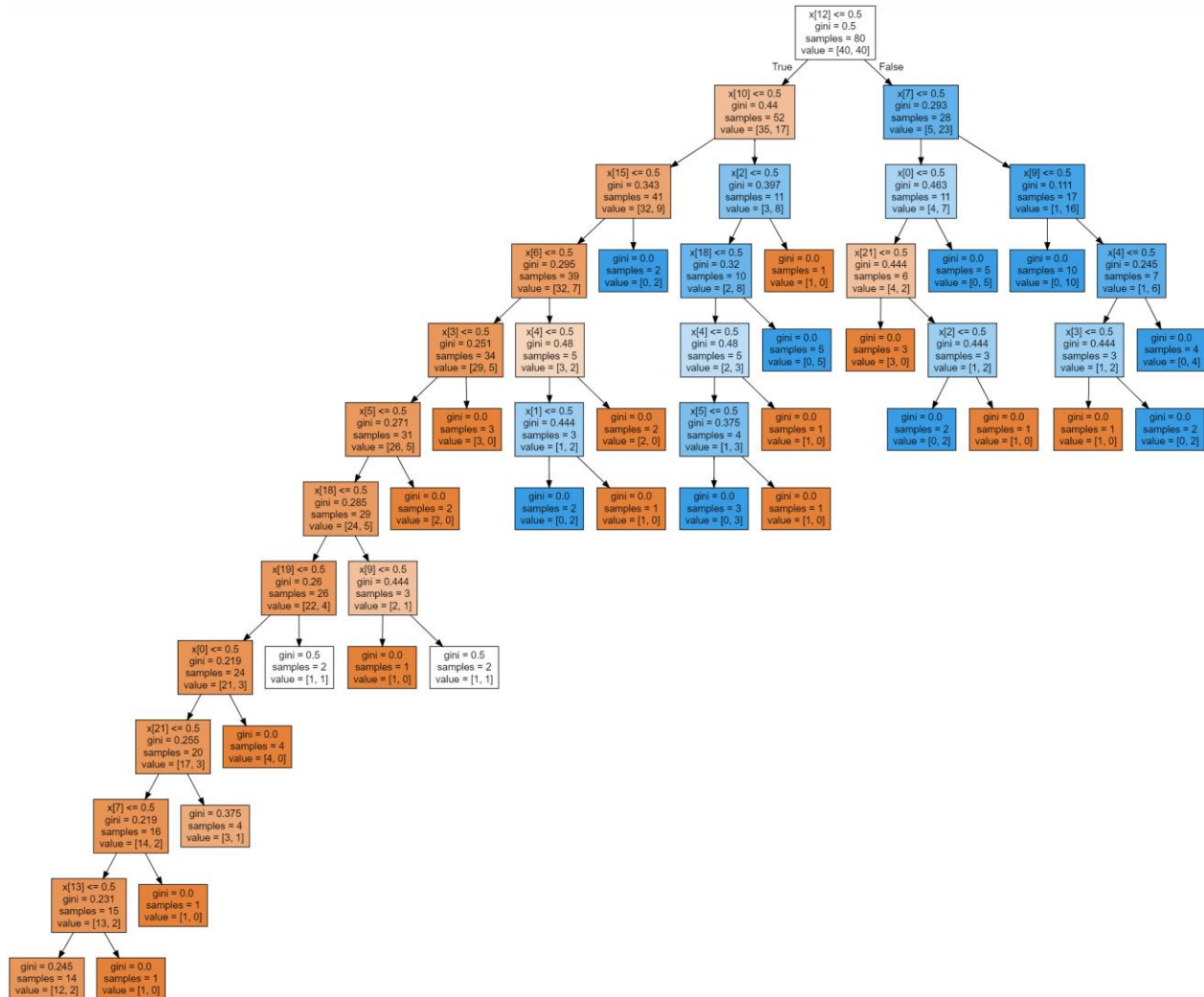
```
TREE
+-- [SPLIT: x0 = 0]
|   +-- [SPLIT: x1 = 0]
|   |   +-- [LABEL = 1]
|   |   +-- [SPLIT: x1 = 0]
|   |   |   +-- [LABEL = 1]
+-- [SPLIT: x0 = 0]
|   +-- [SPLIT: x1 = 0]
|   |   +-- [LABEL = 1]
|   +-- [SPLIT: x1 = 0]
|   |   +-- [LABEL = 0]
TP:  105 | FN:  67
-----
FP:   3 | TN:  12

TPR:  0.6104651162790697 % | FNR:  0.38953488372093026 %
-----
FPR:  0.2 % | TNR:  0.8 %
```



## Dataset 2: Part c)

## Scikit -learn Tree:



## Scikit -learn Confusion Matrix:

```
Test Error for SciKit Learn: 32.09%.  
TP: 115 | FN: 57  
-----  
FP: 3 | TN: 12  
  
TPR: 0.6686046511627907 % | FNR: 0.3313953488372093 %  
-----  
FPR: 0.2 % | TNR: 0.8 %
```

Dataset 3:

Numeric data to determine if somebody has donated blood in 2007

<https://archive.ics.uci.edu/ml/datasets/Blood+Transfusion+Service+Center>

## Dataset 3: Part B):

### Depth-1

```
Blood-Transfusion Dataset
TREE
+-- [SPLIT: x3 = 89]
|   +-- [LABEL = 0]
+-- [SPLIT: x3 = 89]
|   +-- [LABEL = 0]
TP:  0 | FN:  71
-----
FP:  0 | TN: 229

TPR:  0.0 % | FNR:  1.0 %
-----
FPR:  0.0 % | TNR:  1.0 %
```

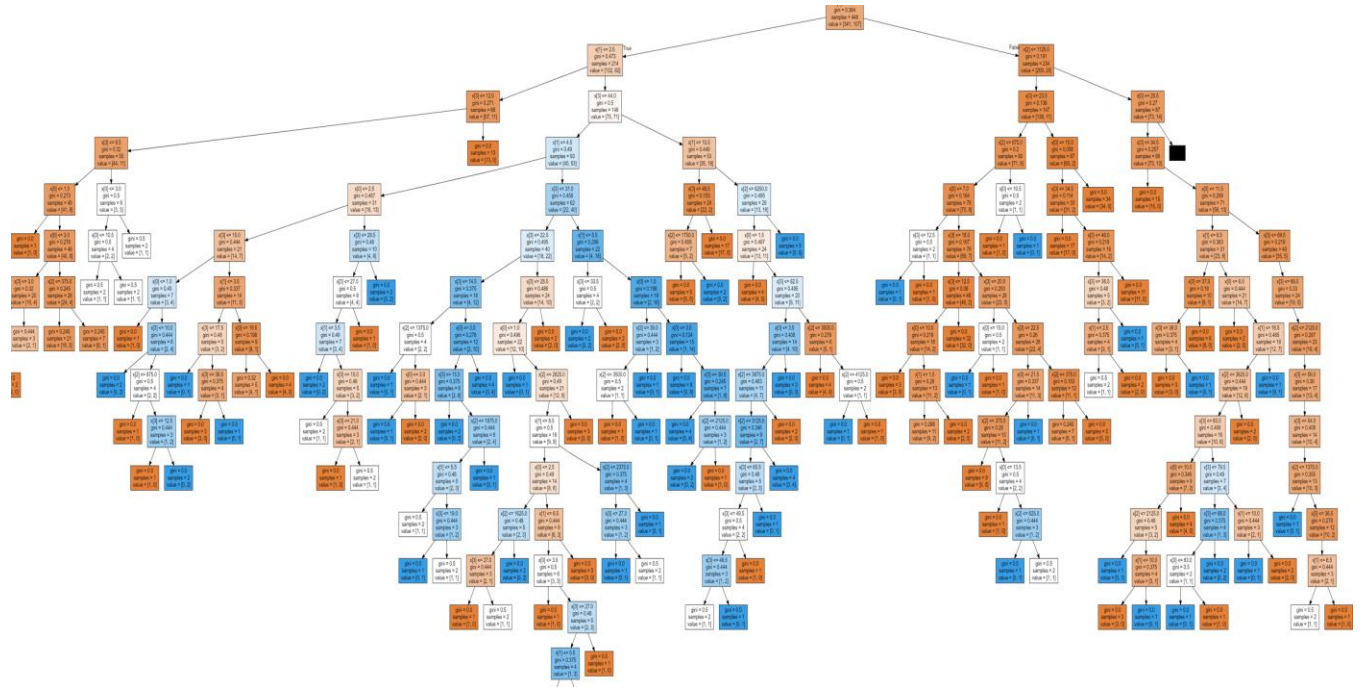
### Depth-2

```
TREE
+-- [SPLIT: x3 = 89]
|   +-- [SPLIT: x3 = 98]
|       |   +-- [LABEL = 0]
|       +-- [SPLIT: x3 = 98]
|           |   +-- [LABEL = 1]
+-- [SPLIT: x3 = 89]
|   +-- [SPLIT: x1 = 3]
|       |   +-- [LABEL = 0]
|       +-- [SPLIT: x1 = 3]
|           |   +-- [LABEL = 0]
TP:  2 | FN:  69
-----
FP:  4 | TN: 225

TPR:  0.028169014084507043 % | FNR:  0.971830985915493 %
-----
FPR:  0.017467248908296942 % | TNR:  0.982532751091703 %
```

## Dataset 3: Part c)

## Scikit-learn Tree:



## Scikit -learn Confusion Matrix:

Test Error for SciKit Learn: 26.67%.

TP: 27 | FN: 44

-----

FP: 36 | TN: 193

TPR: 0.38028169014084506 % | FNR: 0.6197183098591549 %

-----

FPR: 0.1572052401746725 % | TNR: 0.8427947598253275 %