Programming Assignment 1
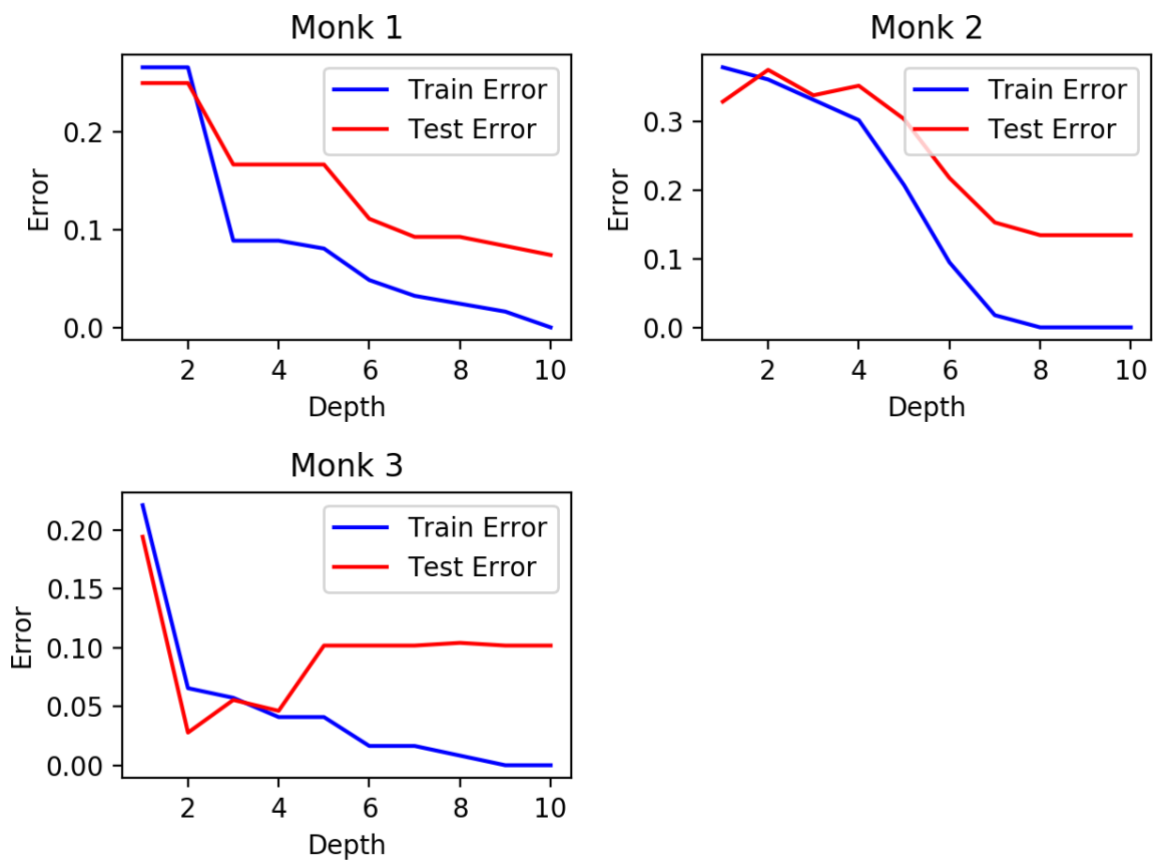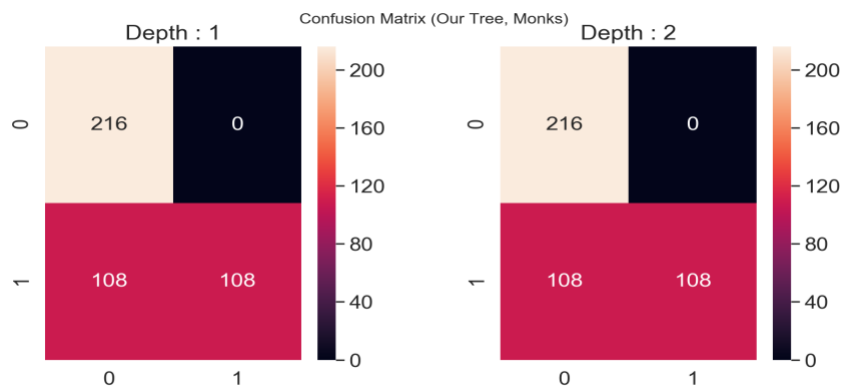Submitted by:
S K Aravind SXA190006
Paresh Dashore PXD190004

a) Learning Curves

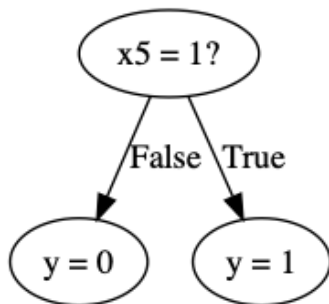

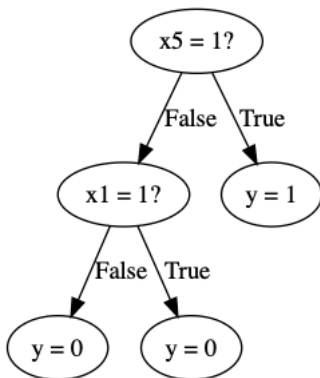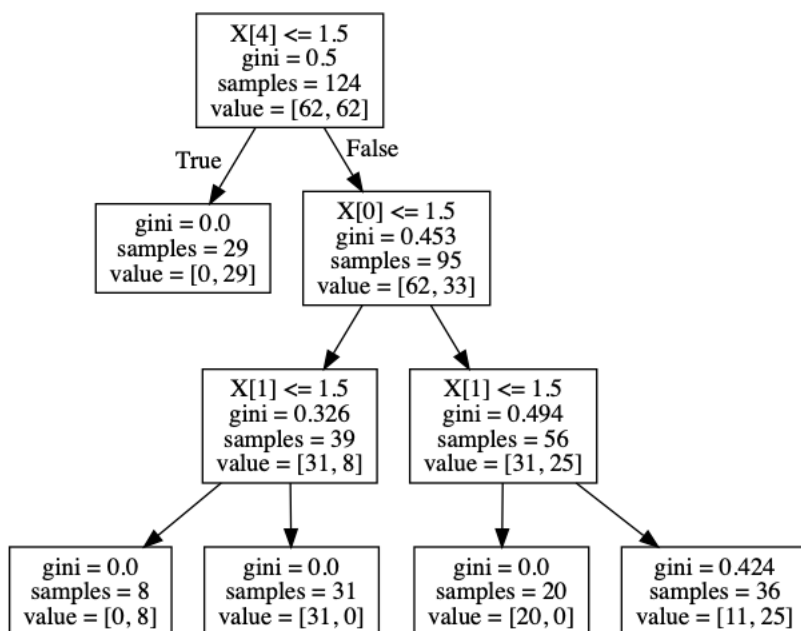## b) Weak Learners (Confusion matrix for Monks-1 dataset, depths=1,2)

Trees obtained:

Depth 1:

```
        ┌──────────┐
        │  x5 = 1? │
        └──────────┘
       False      True
       ↙            ↘
  ┌────────┐    ┌────────┐
  │ y = 0  │    │ y = 1  │
  └────────┘    └────────┘
```

Depth 2:

```
              ┌──────────┐
              │  x5 = 1? │
              └──────────┘
             False      True
             ↙            ↘
        ┌────────┐    ┌────────┐
        │ x1 = 1?│    │ y = 1  │
        └────────┘    └────────┘
       False    True
       ↙          ↘
  ┌────────┐  ┌────────┐
  │ y = 0  │  │ y = 0  │
  └────────┘  └────────┘
```

c) **Scikit-Learn Tree and Confusion matrix:**
        (Tree, depth = 3)

```
                     ┌────────────────────┐
                     │   X[4] <= 1.5       │
                     │   gini = 0.5        │
                     │   samples = 124     │
                     │   value = [62, 62]  │
                     └────────────────────┘
                  True  /        \ False
                       /          \
        ┌────────────────┐   ┌────────────────────┐
        │ gini = 0.0     │   │   X[0] <= 1.5       │
        │ samples = 29   │   │   gini = 0.453      │
        │ value = [0, 29]│   │   samples = 95      │
        └────────────────┘   │   value = [62, 33]  │
                             └────────────────────┘
                            /                      \
              ┌────────────────┐          ┌────────────────┐
              │ X[1] <= 1.5    │          │ X[1] <= 1.5    │
              │ gini = 0.326   │          │ gini = 0.494   │
              │ samples = 39   │          │ samples = 56   │
              │ value = [31, 8]│          │ value = [31, 25]│
              └────────────────┘          └────────────────┘
              /            \              /              \
  ┌────────────┐  ┌──────────────┐ ┌──────────────┐ ┌──────────────┐
  │ gini = 0.0 │  │ gini = 0.0   │ │ gini = 0.0   │ │ gini = 0.424 │
  │ samples = 8│  │ samples = 31 │ │ samples = 20 │ │ samples = 36 │
  │value=[0, 8]│  │value=[31, 0] │ │value=[20, 0] │ │value=[11, 25]│
  └────────────┘  └──────────────┘ └──────────────┘ └──────────────┘
```

Confusion Matrix:

Confusion Matrix (Sklearn, Monks)

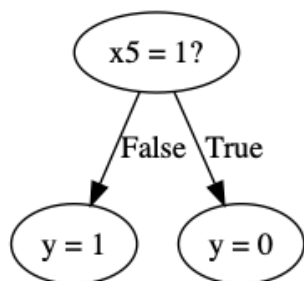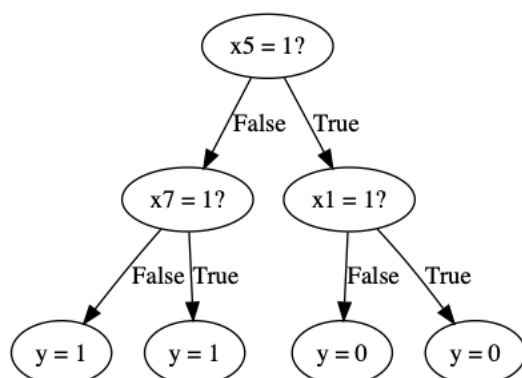|	0	1
0	144	72
1	0	216
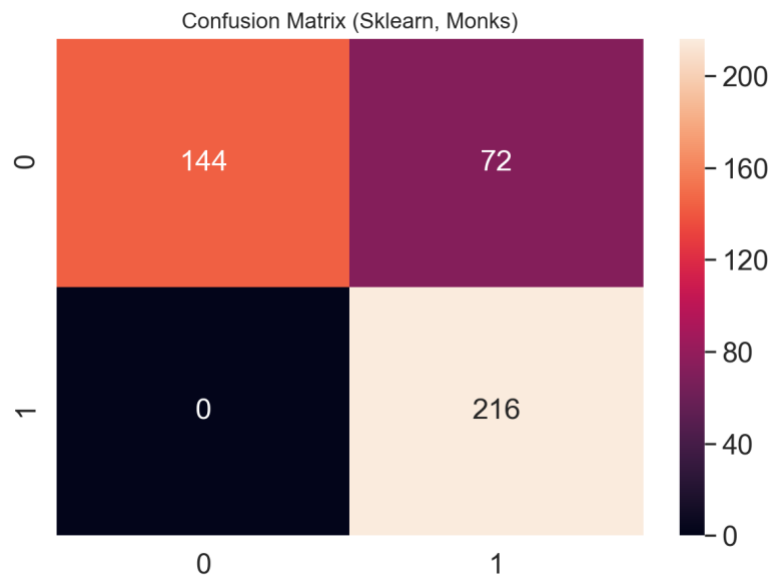
d) **Performance on our Dataset (Tic Tac Toe dataset)**
   No discretization was required for this dataset.

Our Learned Tree
Depth = 1

```
        x5 = 1?
      /        \
  False        True
    /            \
 y = 1         y = 0
```

Depth = 2

```
              x5 = 1?
            /         \
         False        True
          /             \
      x7 = 1?          x1 = 1?
      /     \          /      \
  False    True     False    True
    /        \        /         \
 y = 1     y = 1    y = 0      y = 0
```

Confusion Matrix

Confusion Matrix (Our Tree, Tic Tac Toe)

Depth : 1

	0	1
0	47	36
1	37	120

Depth : 2

	0	1
0	47	36
1	37	120
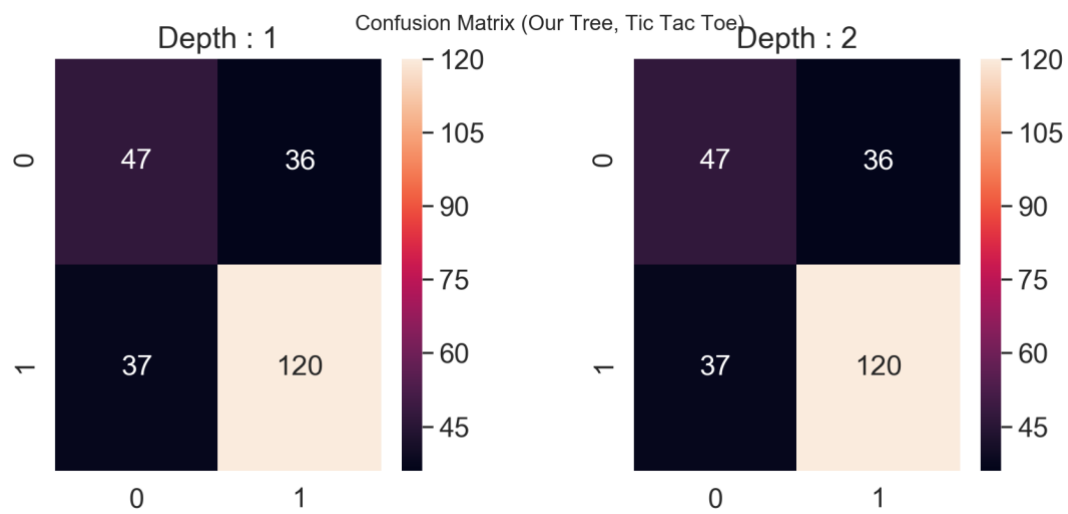
Scikit Learn's tree (depth = 3)


```
X[4] <= 1.5
gini = 0.453
samples = 718
value = [249, 469]
```

True / False

```
X[4] <= 0.5
gini = 0.499
samples = 376
value = [181, 195]
```

```
X[6] <= 0.5
gini = 0.319
samples = 342
value = [68, 274]
```

```
X[8] <= 1.5
gini = 0.42
samples = 120
value = [36, 84]
```

```
X[2] <= 1.5
gini = 0.491
samples = 256
value = [145, 111]
```

```
X[2] <= 0.5
gini = 0.202
samples = 79
value = [9, 70]
```

```
X[6] <= 1.5
gini = 0.348
samples = 263
value = [59, 204]
```

```
gini = 0.49
samples = 63
value = [27, 36]
```

```
gini = 0.266
samples = 57
value = [9, 48]
```

```
gini = 0.453
samples = 130
value = [85, 45]
```

```
gini = 0.499
samples = 126
value = [60, 66]
```

```
gini = 0.0
samples = 23
value = [0, 23]
```

```
gini = 0.27
samples = 56
value = [9, 47]
```

```
gini = 0.413
samples = 134
value = [39, 95]
```

```
gini = 0.262
samples = 129
value = [20, 109]
```

Confusion matrix

Confusion Matrix (Sklearn, Tic Tac Toe)

	0	1
0	26	57
1	16	141