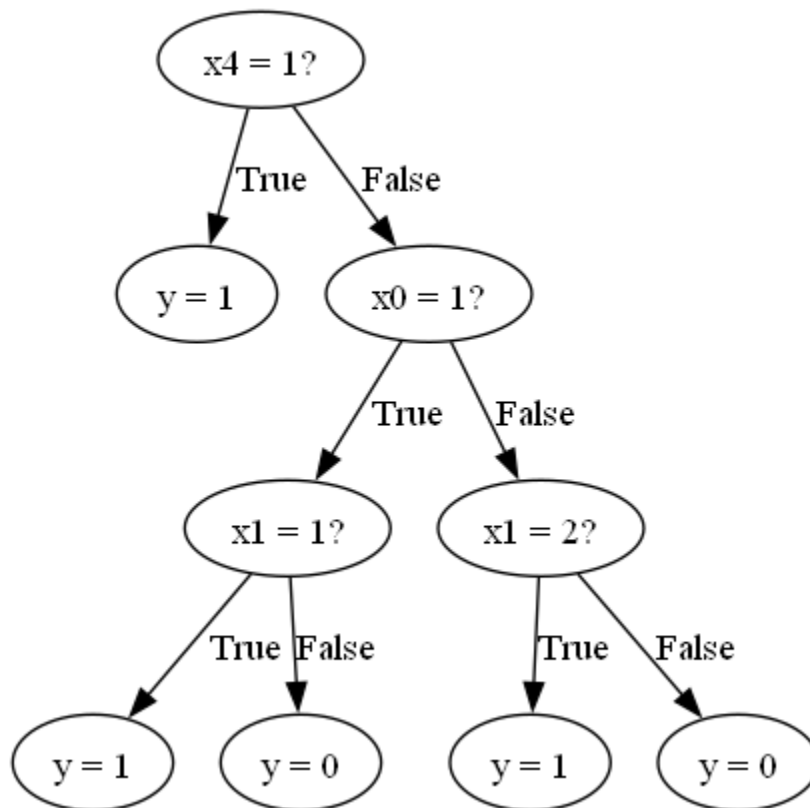


## ASSIGNMENT 2: Fixed-Length Decision Tree

### PART-A



**Fig.1. The decision tree for depth 3 of Monk-1 data**

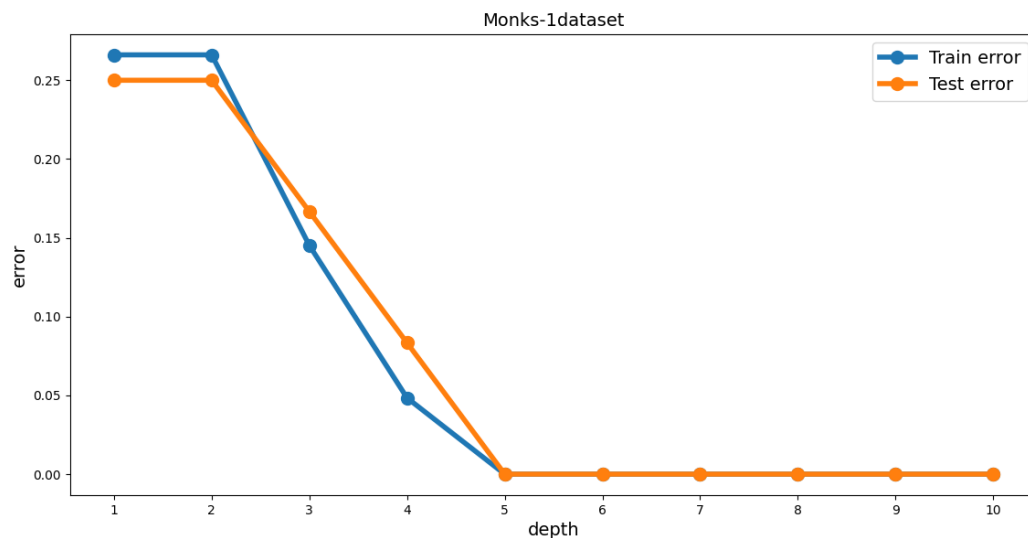
```

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%.

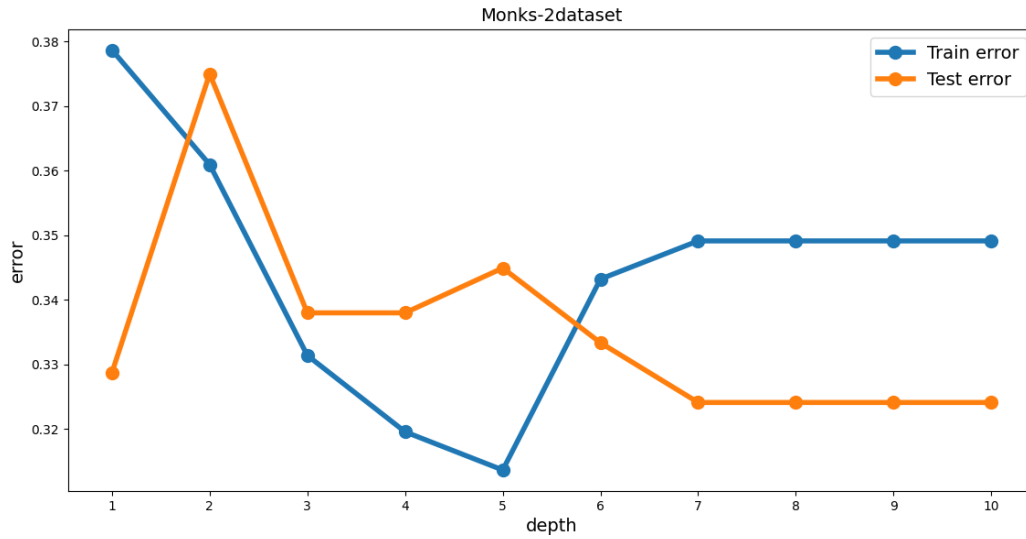
```

**Fig.2. Pretty print tree and Test Error for depth 3 of Monk-1 data**

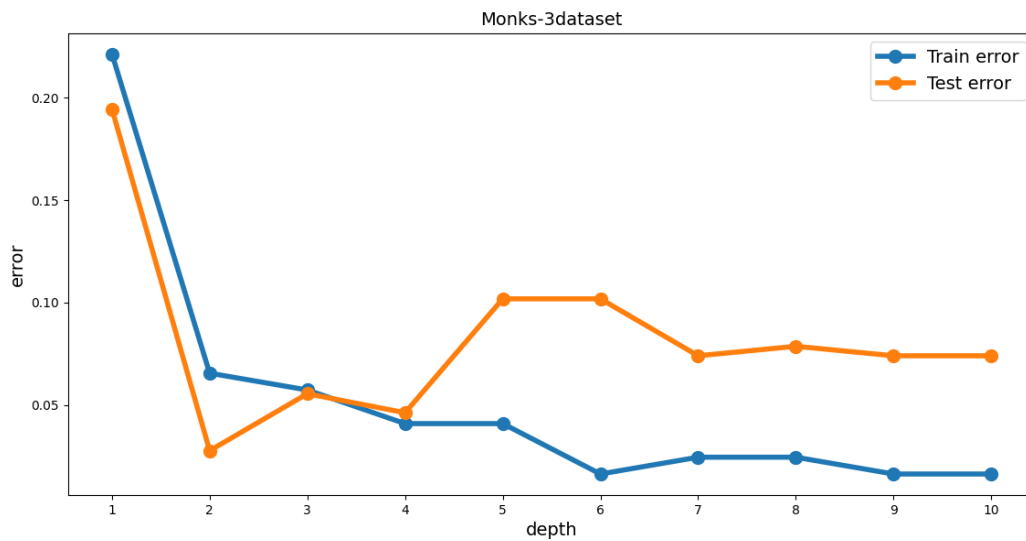
## PART-B



**Fig.3. The training and testing error plots for depth = 1, 2, 4, 5, 6, 7, 8, 9, 10 for Monks-1 dataset**



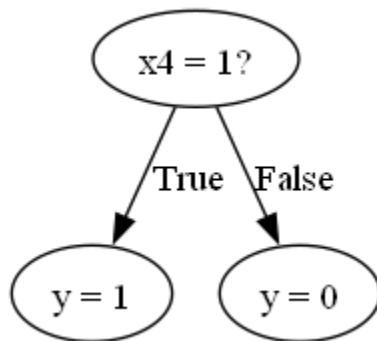
**Fig.4. The training and testing error plots for depth = 1, 2, 4, 5, 6, 7, 8, 9, 10 for Monks-2 dataset**



**Fig.5. The training and testing error plots for depth = 1, 2, 4, 5, 6, 7, 8, 9, 10 for Monks-3 dataset**

## PART-C

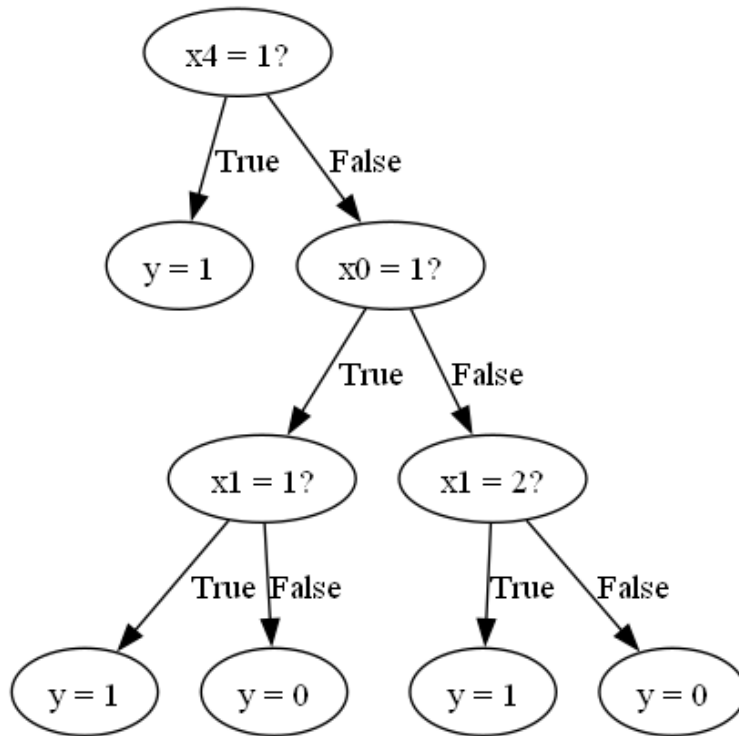
Below is the learned tree, test error and Confusion Matrix using id3 for Monks-1 dataset for Depth=1



```
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]]
```

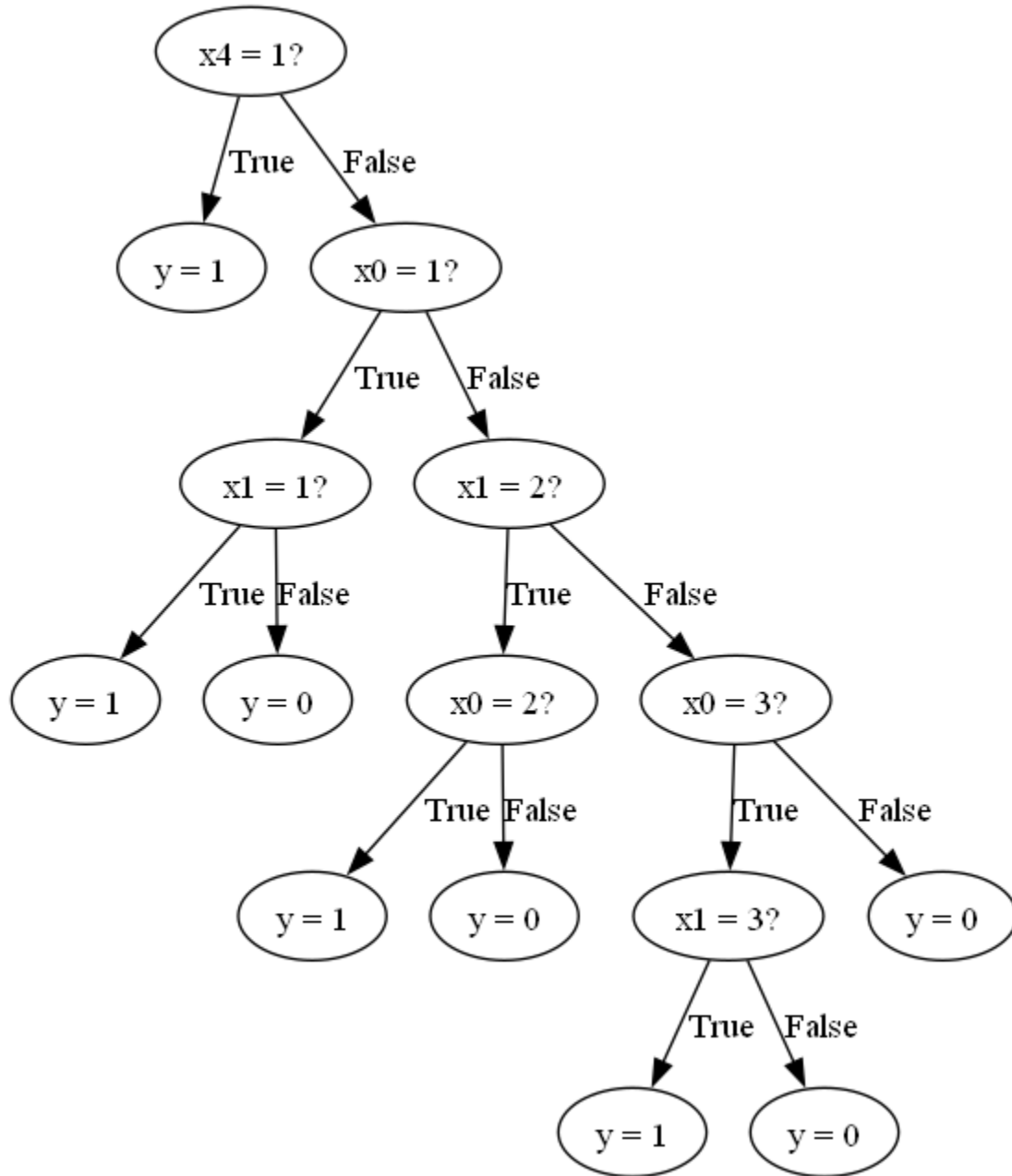
Below is the learned tree, test error and Confusion Matrix using id3 for Monks-1 dataset for Depth=3



```
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]]
```

Below is the learned tree, test error and Confusion Matrix using id3 for Monks-1 dataset for 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

Below is the Pretty Print tree, test error and Confusion Matrix using scikit for Monks-1 dataset for 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]]
```

Below are the Pretty Print tree, test error and Confusion Matrix using scikit for Monks-1 dataset for 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]]
```



**Below are the Pretty Print tree, test error and Confusion Matrix using scikit for Monks-1 dataset for 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]
```

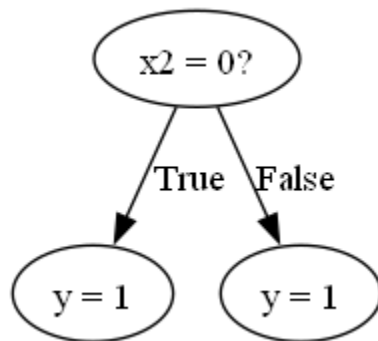
Test error = 0.00%.

```
Confusion Matrix for depth 5
[[168  48]
 [ 24 192]]
```

## PART-E

Below are the learned trees, test errors and confusion matrices for depth 1 of our own data Haberman's survival (from UCI repository) .

- These are generated using id3 and scikitlearn's decision tree classifier



```
decision tree from id3 for depth 1
TREE
+-- [SPLIT: x2 = 0 True]
|   +-- [LABEL = 1]
+-- [SPLIT: x2 = 0 False]
|   +-- [LABEL = 1]
Test Error = 29.03%.
```

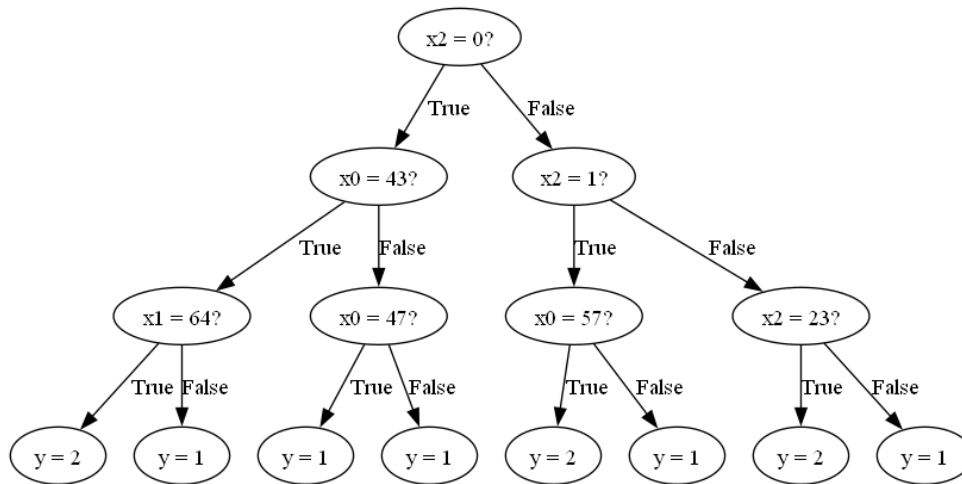
```
Confusion Matrix for the depth 1for id3 algorithm .
[[44  0]
 [18  0]]
```

```
decision tree from scikit for depth 1
|--- feature_2 <= 2.50
|   |--- class: 1
|--- feature_2 > 2.50
|   |--- class: 1
Test Error = 29.03%.
```

```
Confusion Matrix for the depth 1for decision tree using scikit learn.
[[44  0]
 [18  0]]
```

Below are the learned trees, test errors and confusion matrices for depth 3 of our own data Haberman's survival (from UCI repository) .

- These are generated using id3 and scikitlearn's decision tree classifier



```
decision tree from id3 for depth 3
TREE
+-- [SPLIT: x2 = 0 True]
|   +-- [SPLIT: x0 = 43 True]
|   |   +-- [SPLIT: x1 = 64 True]
|   |   |   +-- [LABEL = 2]
|   |   |   +-- [SPLIT: x1 = 64 False]
|   |   |   |   +-- [LABEL = 1]
|   |   +-- [SPLIT: x0 = 43 False]
|   |   |   +-- [SPLIT: x0 = 47 True]
|   |   |   |   +-- [LABEL = 1]
|   |   |   +-- [SPLIT: x0 = 47 False]
|   |   |   |   +-- [LABEL = 1]
|   +-- [SPLIT: x2 = 0 False]
|   |   +-- [SPLIT: x2 = 1 True]
|   |   |   +-- [SPLIT: x0 = 57 True]
|   |   |   |   +-- [LABEL = 2]
|   |   |   +-- [SPLIT: x0 = 57 False]
|   |   |   |   +-- [LABEL = 1]
|   |   +-- [SPLIT: x2 = 1 False]
|   |   |   +-- [SPLIT: x2 = 23 True]
|   |   |   |   +-- [LABEL = 2]
|   |   |   +-- [SPLIT: x2 = 23 False]
|   |   |   |   +-- [LABEL = 1]
Test Error = 29.03%.
```

```
Confusion Matrix for the depth 3for id3 algorithm .
[[44  0]
 [18  0]]
```

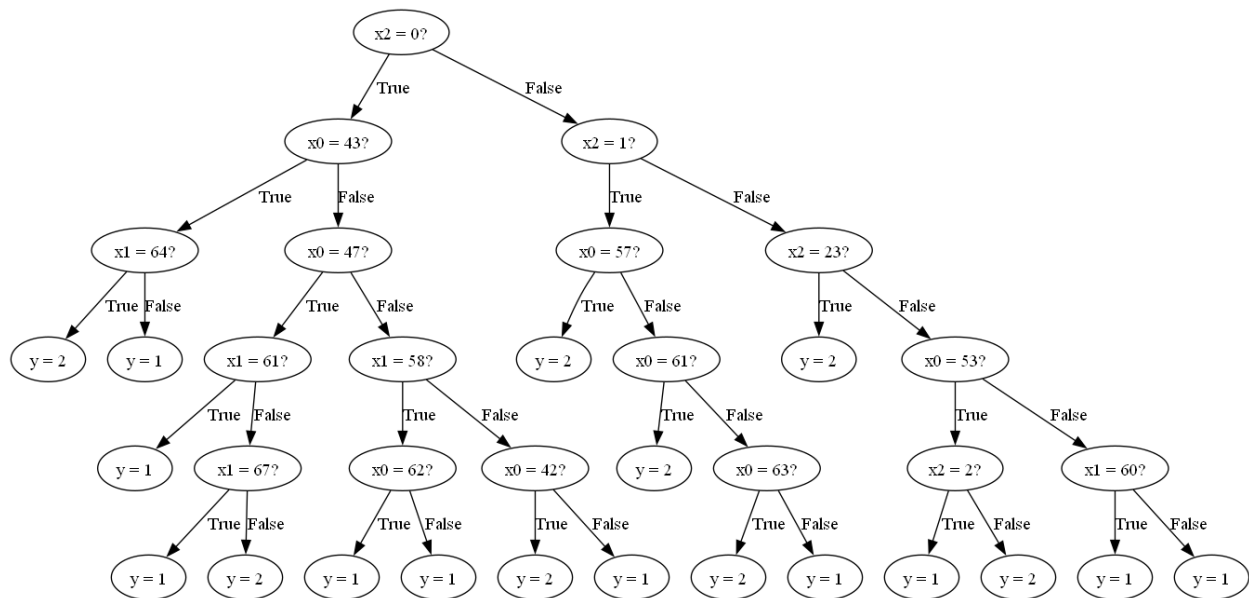
```
|--- feature_2 <= 2.50  
|   |--- feature_0 <= 77.00  
|   |   |--- feature_0 <= 38.50  
|   |   |   |--- class: 1  
|   |   |--- feature_0 > 38.50  
|   |   |   |--- class: 1  
|   |--- feature_0 > 77.00  
|   |   |--- class: 2  
|--- feature_2 > 2.50  
|   |--- feature_0 <= 43.50  
|   |   |--- feature_2 <= 21.50  
|   |   |   |--- class: 1  
|   |   |--- feature_2 > 21.50  
|   |   |   |--- class: 2  
|   |--- feature_0 > 43.50  
|   |   |--- feature_2 <= 8.50  
|   |   |   |--- class: 1  
|   |   |--- feature_2 > 8.50  
|   |   |   |--- class: 2
```

Test Error = 30.65%.

Confusion Matrix for the depth 3for decision tree using scikit learn.  
[[39 5]  
 [14 4]]

Below are the learned trees, test errors and confusion matrices for depth 5 of our own data Haberman's survival (from UCI repository) .

- These are generated using id3 and scikitlearn's decision tree classifier





```
--- feature_2 <= 2.50
|--- feature_0 <= 77.00
|   |--- feature_0 <= 38.50
|   |   |--- class: 1
|   |--- feature_0 > 38.50
|   |   |--- feature_0 <= 47.50
|   |   |   |--- feature_1 <= 68.50
|   |   |   |   |--- class: 1
|   |   |   |--- feature_1 > 68.50
|   |   |   |   |--- class: 2
|   |   |--- feature_0 > 47.50
|   |   |   |--- feature_0 <= 60.50
|   |   |   |   |--- class: 1
|   |   |   |--- feature_0 > 60.50
|   |   |   |   |--- class: 1
|   |--- feature_0 > 77.00
|   |--- class: 2
|--- feature_2 > 2.50
|   |--- feature_0 <= 43.50
|   |   |--- feature_2 <= 21.50
|   |   |   |--- feature_1 <= 65.50
|   |   |   |   |--- class: 1
|   |   |   |--- feature_1 > 65.50
|   |   |   |   |--- feature_2 <= 8.50
|   |   |   |   |   |--- class: 1
|   |   |   |   |--- feature_2 > 8.50
|   |   |   |   |   |--- class: 1
|   |   |   |--- feature_2 > 21.50
|   |   |   |   |--- class: 2
|   |--- feature_0 > 43.50
|   |   |--- feature_2 <= 8.50
|   |   |   |--- feature_0 <= 50.50
|   |   |   |   |--- feature_0 <= 46.50
|   |   |   |   |   |--- class: 2
|   |   |   |   |--- feature_0 > 46.50
|   |   |   |   |   |--- class: 1
|   |   |   |--- feature_0 > 50.50
|   |   |   |   |--- feature_1 <= 60.50
|   |   |   |   |   |--- class: 1
|   |   |   |   |--- feature_1 > 60.50
|   |   |   |   |   |--- class: 2
|   |   |--- feature_2 > 8.50
|   |   |   |--- feature_1 <= 65.50
|   |   |   |   |--- feature_2 <= 24.50
|   |   |   |   |   |--- class: 2
|   |   |   |   |--- feature_2 > 24.50
|   |   |   |   |   |--- class: 1
|   |   |   |--- feature_1 > 65.50
|   |   |   |   |--- feature_2 <= 15.50
|   |   |   |   |   |--- class: 2
|   |   |   |   |--- feature_2 > 15.50
|   |   |   |   |   |--- class: 1
Test Error = 32.26%.
```

```
Confusion Matrix for the depth 5for decision tree using scikit learn.
[[36  8]
 [12  6]]
```

[1] <https://archive.ics.uci.edu/dataset/43/haberman+s+survival>