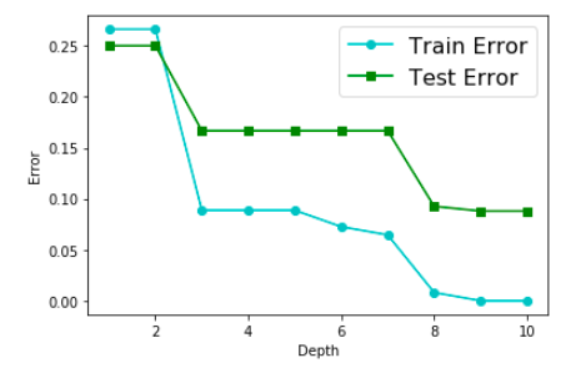
**Assignment 2 – Report**

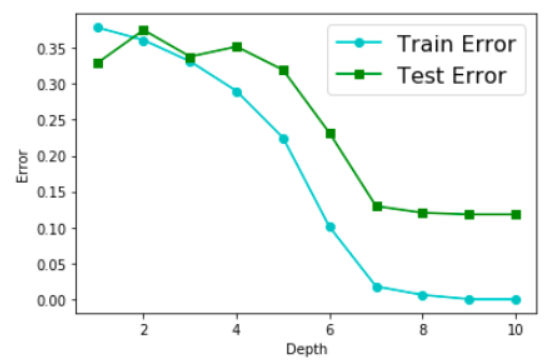
**Padma Priya Cheruvu – ppc180000**

**PART B.**

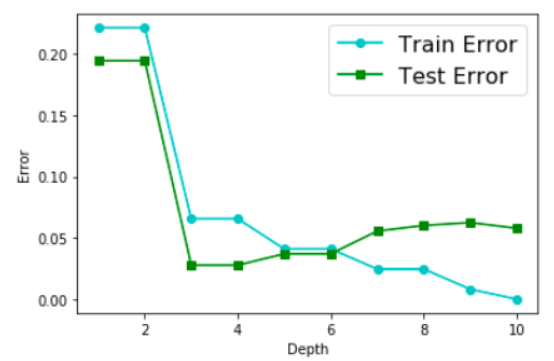
**Monks-1**



**Monks-2**



**Monks-3**



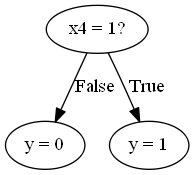
**PART C**

**Depth = 1:**

Predicted Positives, Predicted Negatives

True Positives 216 0

True Negatives 108 108

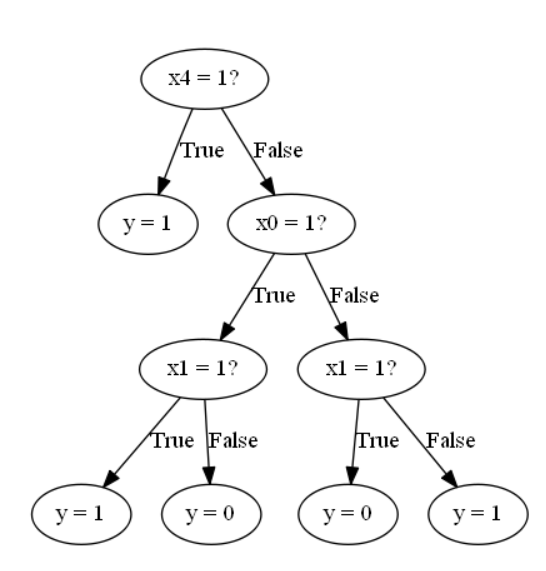


**Depth = 3:**

Predicted Positives Predicted Negatives

True Positives 144 72

True Negatives 0 216

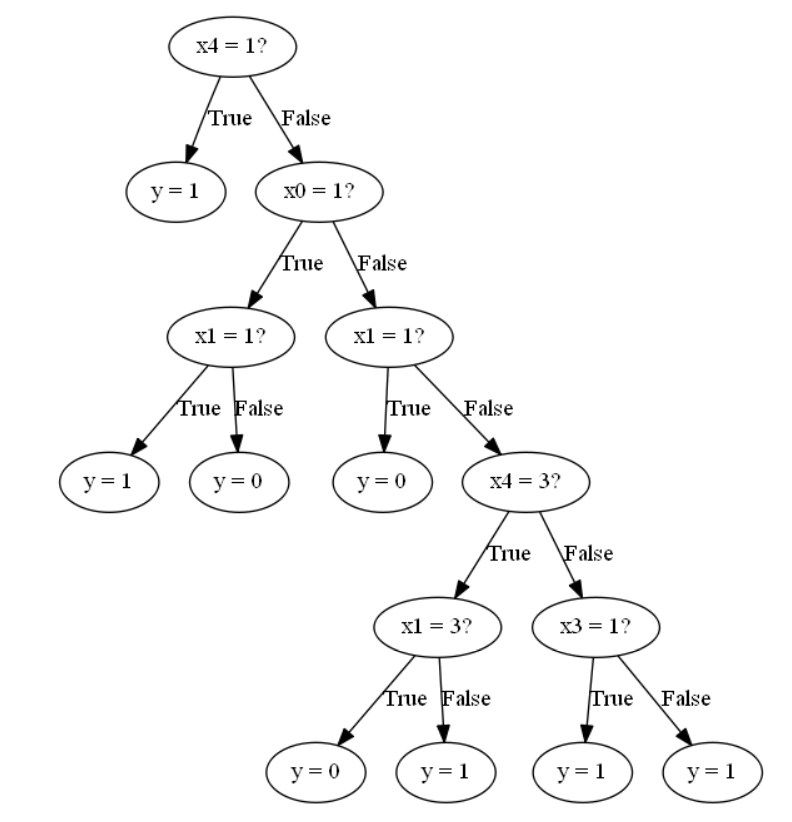


**Depth = 5:**

Predicted Positives Predicted Negatives

True Positives 156 60

True Negatives 12 204



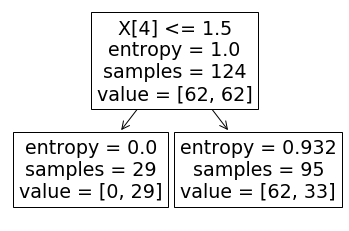
**PART D**

**Depth = 1:**

Predicted Positives, Predicted Negatives

True Positives 216 0

True Negatives 108 108

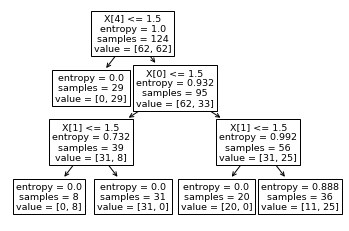


**Depth = 3:**

Predicted Positives, Predicted Negatives

True Positives 144 72

True Negatives 0 216

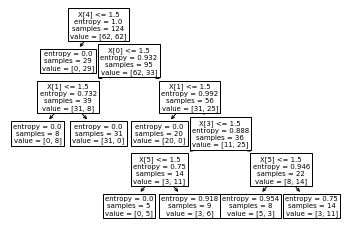


**Depth = 5:**

Predicted Positives, Predicted Negatives

True Positives 168 48

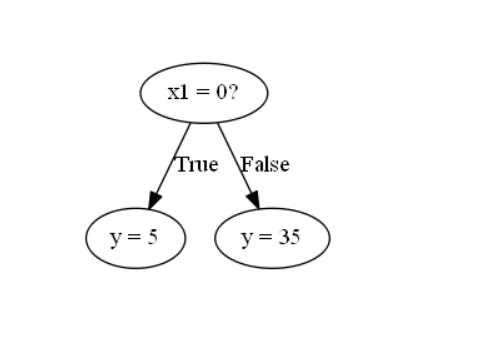
True Negatives 24 192



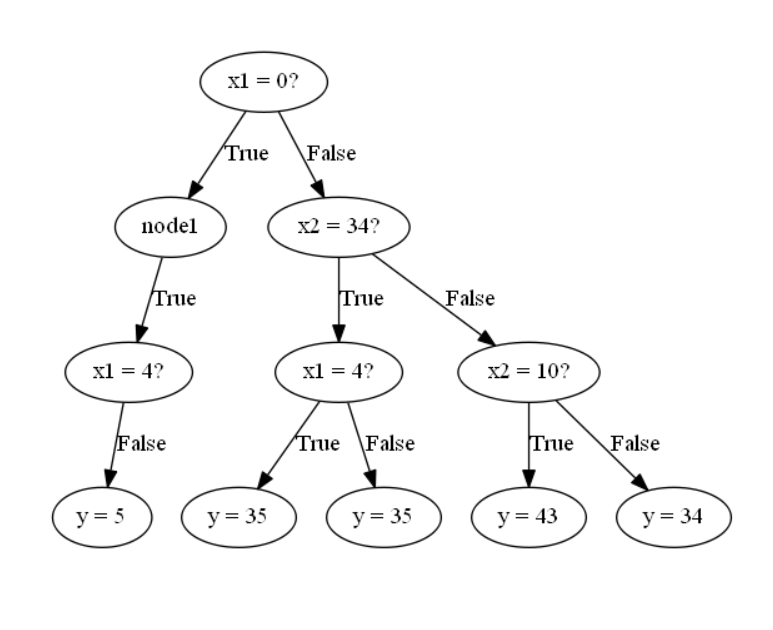
**PART E**

**For dataset Census :**

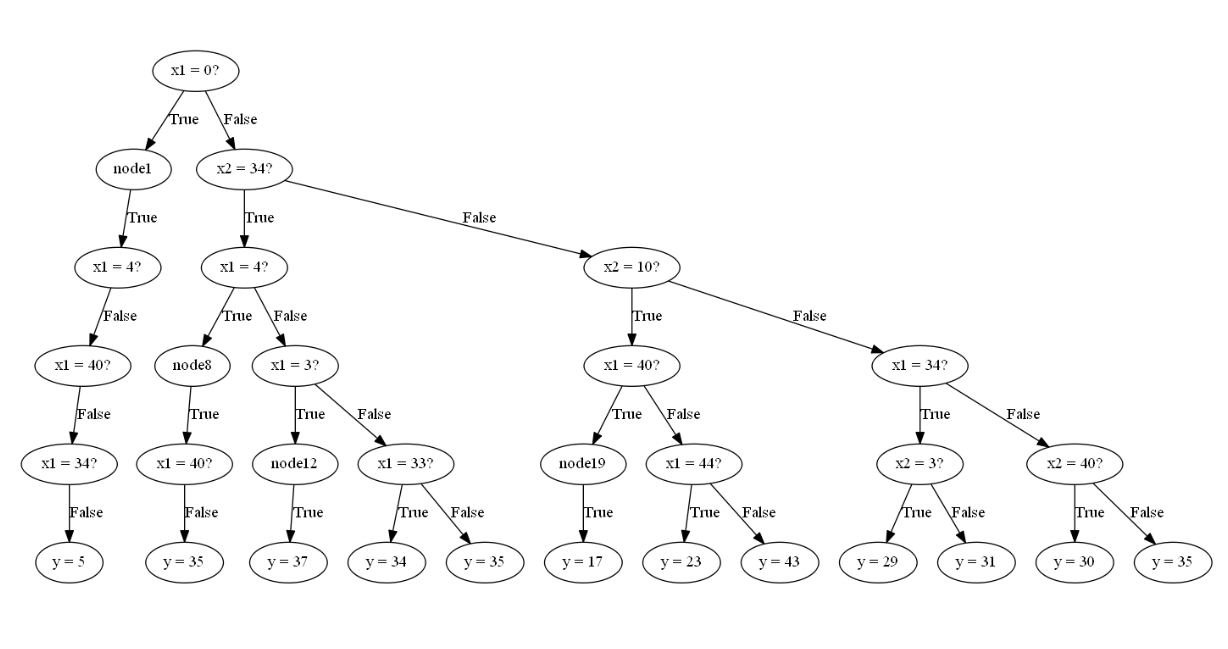
**Depth = 1:**



**Depth = 3:**

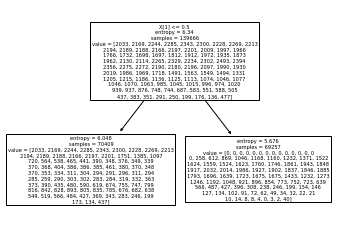


**Depth = 5:**

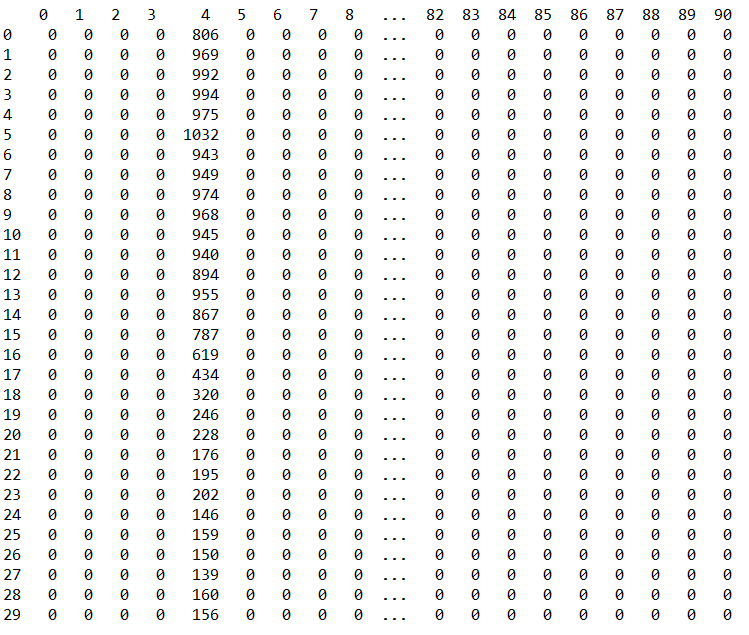


**Using Decision Tree Classifier:**

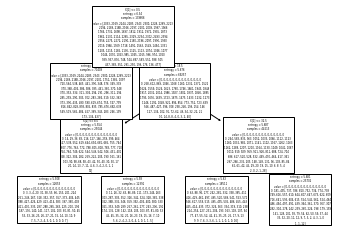
**Depth = 1:**

  
Confusion Matrix

90 x 90 confusion matrix

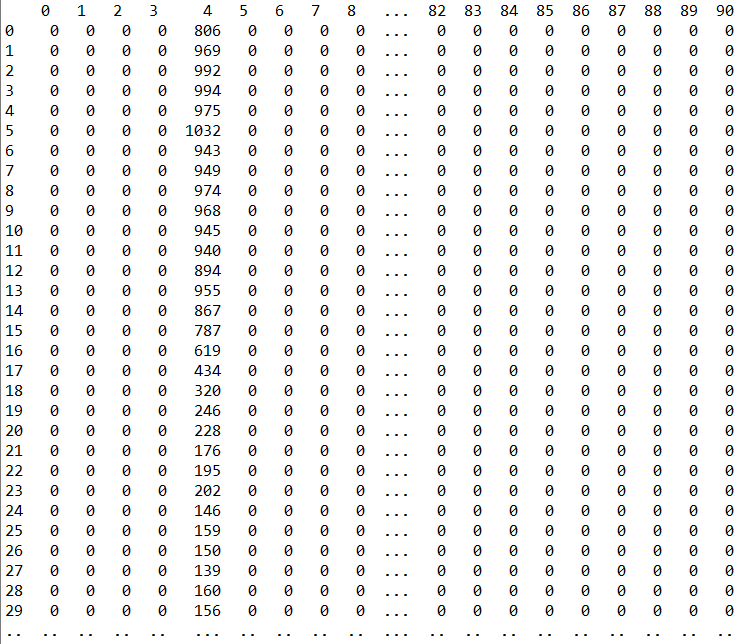


**Depth = 3:**

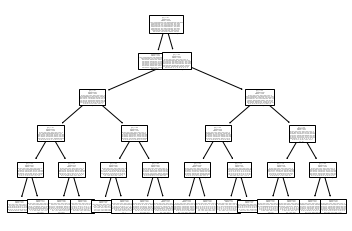


Confusion Matrix

90 x 90 confusion matrix

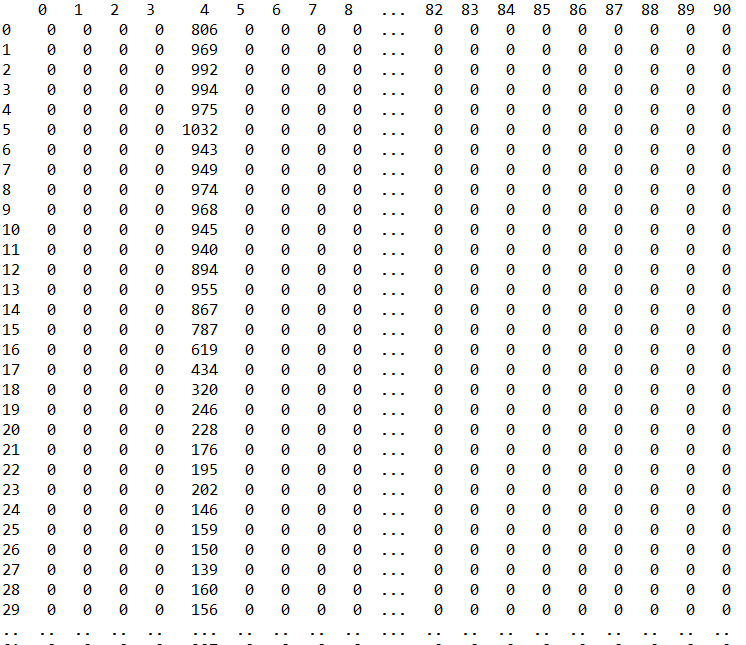


**Depth = 5:**



Confusion Matrix

90 x 90 confusion matrix



**Inference:**

* The confusion matrices and graphs have been plotted for all the datasets.
* In all the monk datasets, the training error is lesser than the test error.
* But, while plotting using small scale, we see that there is a significant difference between training and test error
* This shows that there is some amount of overfitting which has happened. This is one disadvantage of Decision Tree.
* Also, according to confusion matrices, our results are comparable to the one in scikit learn