Programming Assignment #2

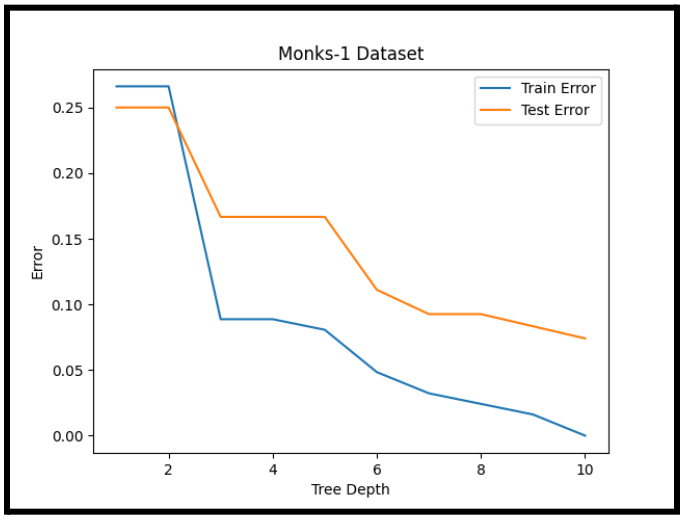
Decision Trees

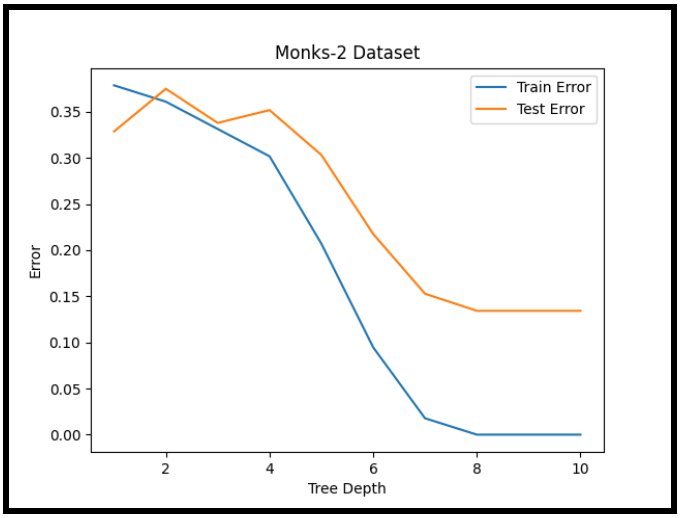
February 26th, 2023

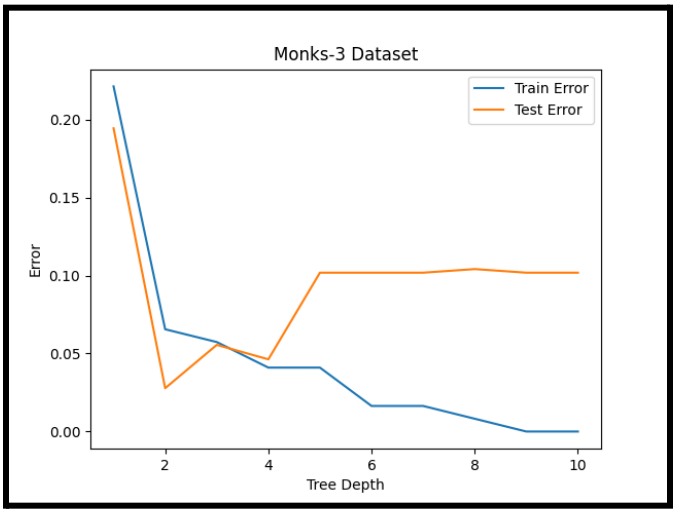
Gautam Sapre (*gss170001*)

Sanjeev Penupala (*sxp170022*)

a) Learning Curves







b) Weak Learners

MONKS-1 *DEPTH=1*		Classifier Prediction	
		Positive	Negative
Actual Value	Positive	108	108
	Negative	0	216

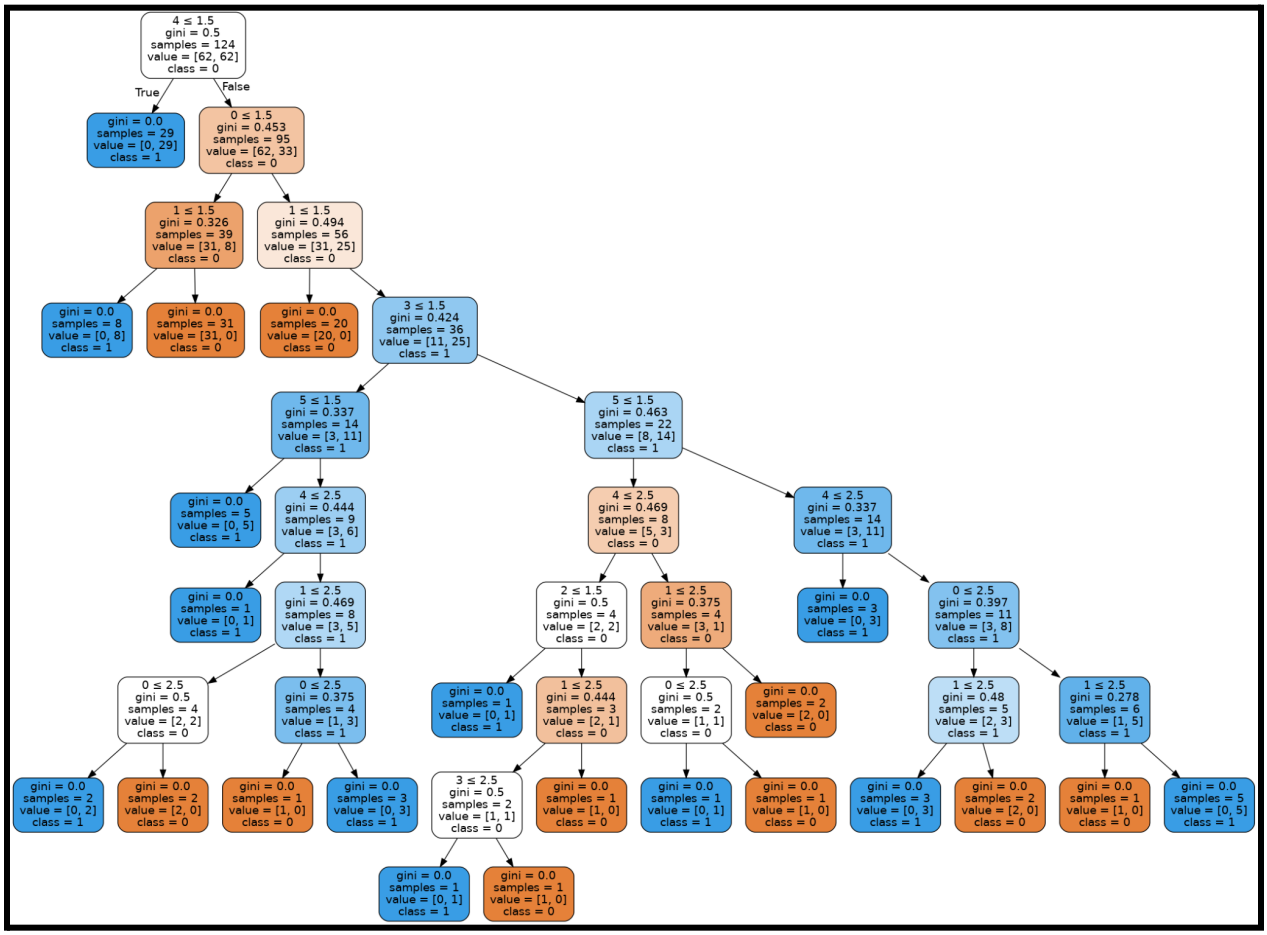
```
TREE (DEPTH=1)
+-- [SPLIT: x4 = 1 | False]
|       +-- [LABEL = 0]
+-- [SPLIT: x4 = 1 | True]
|       +-- [LABEL = 1]
```

MONKS-1 *DEPTH=2*		Classifier Prediction	
		Positive	Negative
Actual Value	Positive	108	108
	Negative	0	216

```
TREE (DEPTH=2)
+-- [SPLIT: x4 = 1 | False]
|       +-- [SPLIT: x0 = 1 | False]
|       |       +-- [LABEL = 0]
|       +-- [SPLIT: x0 = 1 | True]
|       |       +-- [LABEL = 0]
+-- [SPLIT: x4 = 1 | True]
|       +-- [LABEL = 1]
```

c) scikit-learn

MONKS-1 *Scikit-Learn*		Classifier Prediction	
		Positive	Negative
Actual Value	Positive	205	11
	Negative	29	187



d) Other Data Sets

We used the Breast Cancer dataset from Scikit-learn for our second dataset. We discretized all continuous features, according to the simple rule in the instructions page. Below are the confusion matrices for our algorithm and scikit's learning algorithm.

MONKS-1 *DEPTH=1*		**Classifier Prediction**	
		Positive	Negative
Actual Value	Positive	64	7
	Negative	6	37

MONKS-1 *DEPTH=2*		**Classifier Prediction**	
		Positive	Negative
Actual Value	Positive	64	7
	Negative	6	37

MONKS-1 *Scikit-Learn*		**Classifier Prediction**	
		Positive	Negative
Actual Value	Positive	68	3
	Negative	4	39