RESULTS

1. Iris Dataset (Continuous, 4 attributes)

Fold	Majority Classifier	Unpruned Decision Tree		Pruned Decision Tree	
		Accuracy	No. of Nodes	Accuracy	No. of Nodes
1	0.27	0.87	9	0.93	5
5	0.13	0.93	13	0.93	5
3	0.40	0.93	13	0.93	9
4	0.53	0.93	13	0.93	9
5	0.47	1.0	13	1.0	9
6	0.33	1.0	13	1.0	5
7	0.53	1.0	13	1.0	5
8	0.13	0.93	13	1.0	5
9	0.20	0.80	13	0.93	5
10	0.33	0.87	13	0.87	9
Confidence Interval	±0.77	±0.12		±0.078	

2. Car Dataset (Discrete, 6 attributes)

Fold	Majority Classifier	Unpruned Decision Tree		Pruned Decision Tree	
		Accuracy	No. of Nodes	Accuracy	No. of Nodes
1	0.71	0.91	286	0.91	261
2	0.71	0.92	284	0.92	255
3	0.74	0.96	296	0.96	273
4	0.70	0.88	287	0.88	244
5	0.68	0.90	279	0.91	254
6	0.67	0.92	275	0.94	250
7	0.69	0.94	279	0.92	237
8	0.66	0.95	279	0.95	253
9	0.73	0.92	279	0.92	237
10	0.71	0.90	279	0.91	263
Confidence Interval	±0.32	±0.098		±0.095	

3. Breast Cancer Dataset (Continuous, 9 attributes)

Fold	Majority Classifier	Unpruned Decision Tree		Pruned Decision Tree	
		Accuracy	No. of Nodes	Accuracy	No. of Nodes
1	0.62	0.91	43	0.91	31
2	0.62	0.94	39	0.94	39
3	0.67	0.94	35	0.86	29
4	0.68	0.97	55	0.99	31
5	0.65	0.96	53	0.97	23
6	0.64	0.88	35	0.88	31
7	0.68	0.94	45	0.97	21
8	0.64	0.97	45	0.97	25
9	0.75	0.90	45	0.91	21
10	0.62	0.94	45	0.94	21
Confidence Interval	±0.37	±0.084		±0.078	

4. Mushroom Dataset (Discrete, 22 attributes)

Fold	Majority Classifier	Unpruned Decision Tree		Prune Decision Tree	
		Accuracy	No. of Nodes	Accuracy	No. of Nodes
1	0.55	1	29	1	29
2	0.51	1	29	1	29
3	0.54	1	29	1	29
4	0.50	1	29	1	29
5	0.49	1	29	1	29
6	0.51	1	29	1	29
7	0.53	1	29	1	29
8	0.51	1	29	1	29
9	0.53	1	29	1	29
10	0.52	1	29	1	29
Confidence Interval	±0.495	0		0	

5. Pima Indian Diabetes Dataset (Continuous, 8 attributes)

Fold	Majority Classifier	Unpruned Decision Tree		Pruned Decision Tree	
		Accuracy	No. of Nodes	Accuracy	No. of Nodes
1	0.67	0.64	111	0.71	23
2	0.59	0.70	121	0.74	31
3	0.58	0.63	99	0.62	19
4	0.75	0.74	107	0.72	35
5	0.72	0.68	111	0.68	61
6	0.55	0.70	101	0.75	7
7	0.64	0.62	125	0.78	19
8	0.64	0.70	125	0.87	21
9	0.62	0.71	125	0.75	21
10	0.74	0.75	125	0.82	21
Confidence Interval	±0.397	±0.3437		±0.305	

6. Phishing Dataset (Discrete, 30 attributes)

Note : Pruning takes a very long time to converge on this dataset

Fold	Majority Classifier	Unpruned Decision Tree		Pruned Decision Tree	
		Accuracy	No. of Nodes	Accuracy	No. of Nodes
1	0.55	0.96	773	0.96	495
2	0.55	0.97	769	0.97	589
3	0.56	0.96	774	0.95	440
4	0.57	0.96	786	0.97	558
5	0.54	0.97	778	0.97	568
6	0.57	0.96	782	0.96	523
7	0.58	0.95	738	0.96	612
8	0.55	0.95	738	0.95	626
9	0.55	0.97	738	0.96	561
10	0.55	0.97	738	0.97	644
Confidence Interval	±0.45	±0.041		±0.044	