

Assignment 2 Report

Random states = [38, 53, 22, 69, 71]

Experiment 1

Training set = 25% , Testing set = 75%

RANDOM STATE	ACCURACY	NUMBER OF NODES
38	0.9689018464528668	27
53	0.9698736637512148	25
22	0.9747327502429544	27
69	0.9815354713313897	25
71	0.9620991253644315	23

Summary

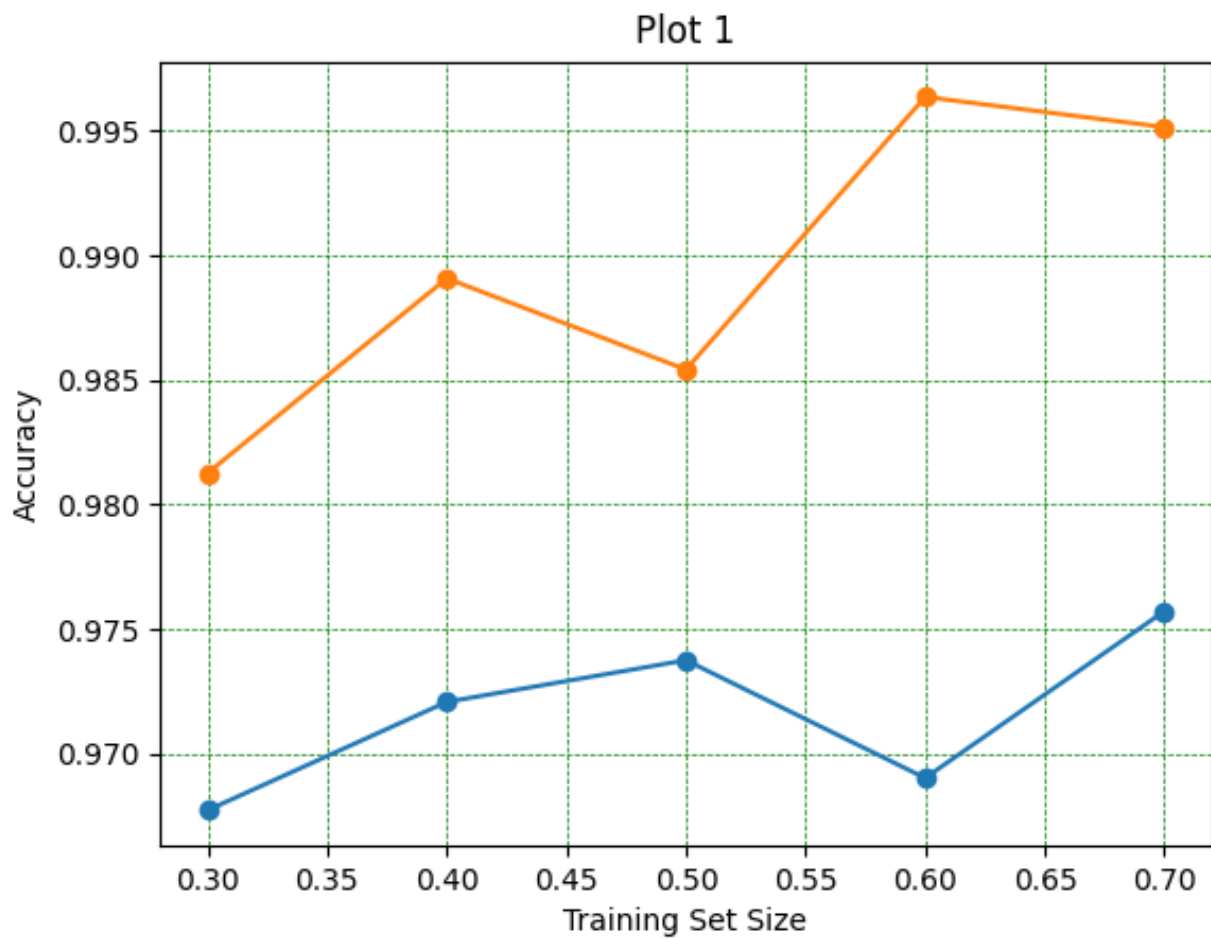
Dataset Split		Minimum Accuracy	Maximum Accuracy	Accuracy Mean	Minimum Number of nodes	Maximum Number of nodes	Number of nodes Mean
Fixed Range	Training 25%	0.9591836734693877	0.9825072886297376	0.9706511175898932	23	27	25.4
	Testing 75%						

Experiment 2

Dataset Split		Minimum Accuracy	Maximum Accuracy	Accuracy Mean	Minimum Number of nodes	Maximum Number of nodes	Number of nodes Mean
Range 1	Training 30%	0.967741935483871	0.9812695109261186	0.976066597294485	17	29	23.8
	Testing 70%						
Range 2	Training 40%	0.9684466019417476	0.9902912621359223	0.9810679611650486	17	29	25
	Testing 60%						
Range 3	Training 50%	0.9752186588921283	0.9854227405247813	0.9795918367346939	23	37	30.6
	Testing 50%						
Range 4	Training 60%	0.9690346083788707	0.9963570127504554	0.9836065573770492	31	45	36.2
	Testing 40%						
Range 5	Training 70%	0.9781553398058253	0.9927184466019418	0.9868932038834952	27	39	31.4
	Testing 30%						

Plots

1) shows accuracy against training set size



2) the number of nodes in the final tree against training set size

