Report for Assignment1

Yidan Sheng yxs173130

The report of accuracies on decision trees which are before pruning and post pruned has given below.

Dataset 1	Gain(Entropy)		Gain(Variance Impurity)	
(L,K)	Before pruning	After pruning	Before pruning	After pruning
(20,5)	0.7555	0.7555	0.762	0.7695
(30,20)	0.7555	0.763	0.762	0.767
(32,15)	0.7555	0.771	0.762	0.7675
(37,18)	0.7555	0.7675	0.762	0.766
(42,15)	0.7555	0.7665	0.762	0.769
(45,9)	0.7555	0.7755	0.762	0.7665
(44,18)	0.7555	0.761	0.762	0.767
(47,20)	0.7555	0.7555	0.762	0.777
(49,22)	0.7555	0.7555	0.762	0.7665
(50,12)	0.7555	0.7645	0.762	0.766

Dataset 2	Gain(Entropy)		Gain(Variance Impurity)	
(L,K)	Before pruning	After pruning	Before pruning	After pruning
(28,12)	0.7233	0.7267	0.725	0.7417
(30,15)	0.7233	0.7317	0.725	0.7567
(33,26)	0.7233	0.7317	0.725	0.7333
(36,28)	0.7233	0.7233	0.725	0.7483
(40,27)	0.7233	0.7283	0.725	0.7433
(44,19)	0.7233	0.73	0.725	0.745
(45,18)	0.7233	0.7383	0.725	0.7383
(47,40)	0.7233	0.7233	0.725	0.735
(49,22)	0.7233	0.7233	0.725	0.7267
(50,32)	0.7233	0.7416	0.725	0.725