Assignment 6

Shivu Gururaj
Sxg144730

Comparing Classifier Performances

Dataset	Number of total instances	Number of attributes	Percent split	Decision Tree Accuracy	Naive Bayesian Accuracy	Support Vector Machine	Neural Network Accuracy	Perceptron Accuracy
D1-1	351	34	90/10	80.55	83.33	83.33	86.11	80.55
D1-2	351	34	90/10	94.44	83.33	80.55	86.11	88.88
D2-1	150	4	90/10	93.33	100	93.33	93.33	80
D2-2	150	4	90/10	100	100	100	100	46.66
D3-1	5456	30	90/10	100	91.75	96.88	78.75	38.46
D3-2	5456	30	90/10	100	90.65	95.60	63.0	39.9
D4-1	47	35	90/10	100	100	100	100	60
D4-2	47	35	90/10	80	100	100	80	60
D5-1	90	8	90/10	66.66	66.66	66.66	66.66	66.66
D5-2	90	8	90/10	55.55	77.77	66.66	55.55	66.66

D1 - dataSet1 <- "http://archive.ics.uci.edu/ml/machine-learning-databases/ionosphere/ionosphere.data"
D2 - dataSet2 <- "http://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"
D3 - dataSet3 <- "http://archive.ics.uci.edu/ml/machine-learning-databases/00194/sensor_readings_2.data"
D4 - dataSet4 <- "http://archive.ics.uci.edu/ml/machine-learning-databases/soybean/soybean-small.data"

D5 - dataSet5 <- "http://archive.ics.uci.edu/ml/machine-learning-databases/postoperative-patient-data/post-operative.data"

Analysis:

Data set 1: Upon running the classifiers 2 times on this data set, it was observed that Decision tree and Neural networks performed better than other classifiers

Data set 2: Upon running the classifiers 2 times on this data set, it was observed that all the classifiers performed better except for perceptron.

Data set 3: Upon running the classifiers 2 times on this data set, it was observed that Decision tree, Naïve Bayesian and SVM performed better than other classifiers

Data set 4: Upon running the classifiers 2 times on this data set, it was observed that all the classifiers performed better except for perceptron.

Data set 5: Upon running the classifiers 2 times on this data set, it was observed that all the classifiers performed pretty much the same

Based on all above observations, I observed that Decision tree and SVM performed better on the given data sets and Naïve Bayesian and Neural networks were better. But Perceptron had the least accuracy. Also I observed that for more data, classifiers result in better accuracies.

Among Decision tree and SVM, I choose SVM because Decision Tree performs at a steady rate, since noise is also a factor and decision tree tries to learn everything.