

Machine Learning Assignment 2 (Decision Tree):

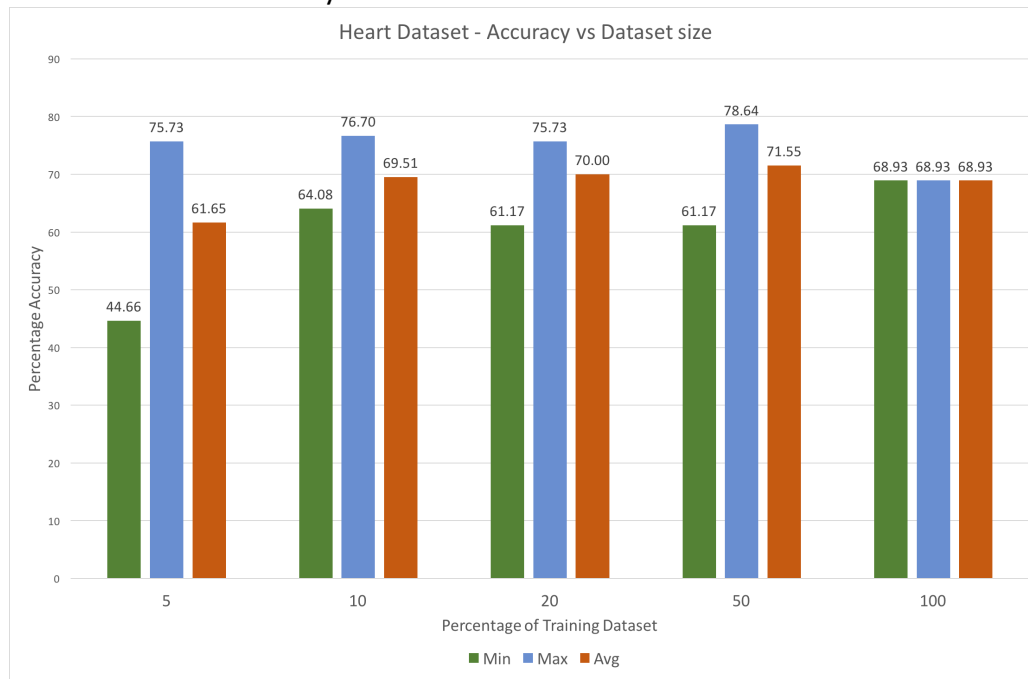
Name: Rohit Damkondwar

Email: damkondwar@wisc.edu

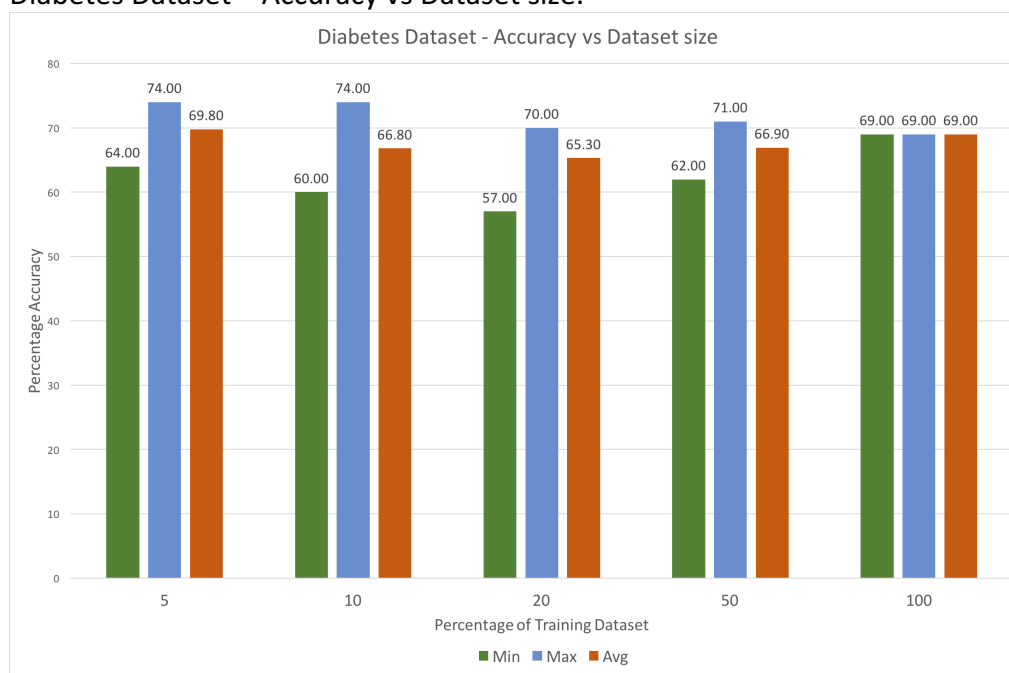
Part 2: Test-set Accuracy vs Dataset size

The DT Predictor was trained for 2 datasets- Heart dataset and Diabetes dataset for variable dataset sizes. The sizes include 5,10,20,50 and 100% of entire dataset. The dataset was chosen randomly from original dataset. Each dataset size was used for 10 iterations for measuring test-set accuracy.

Heart Dataset – Accuracy vs Dataset size.



Diabetes Dataset – Accuracy vs Dataset size.

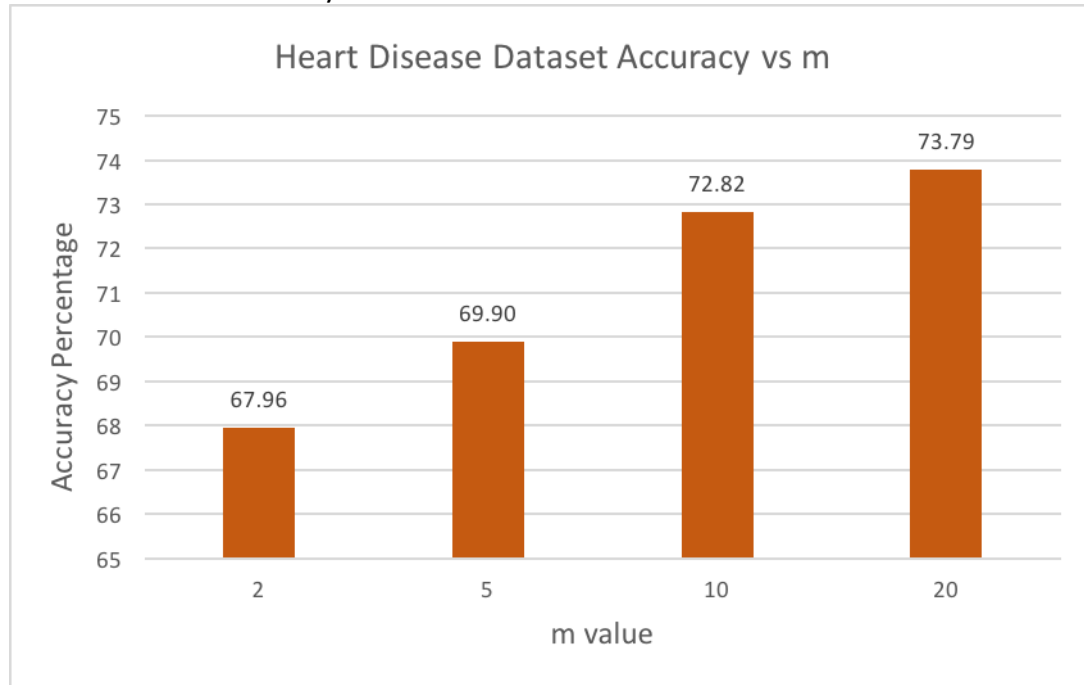


Part 3:

The DT Predictor was run for variable values of m – 2, 5, 10 and 20.

The graphs represent test-set accuracy for each value of m .

Heart Dataset – Accuracy vs m -value



Diabetes Dataset – Accuracy vs m -value

