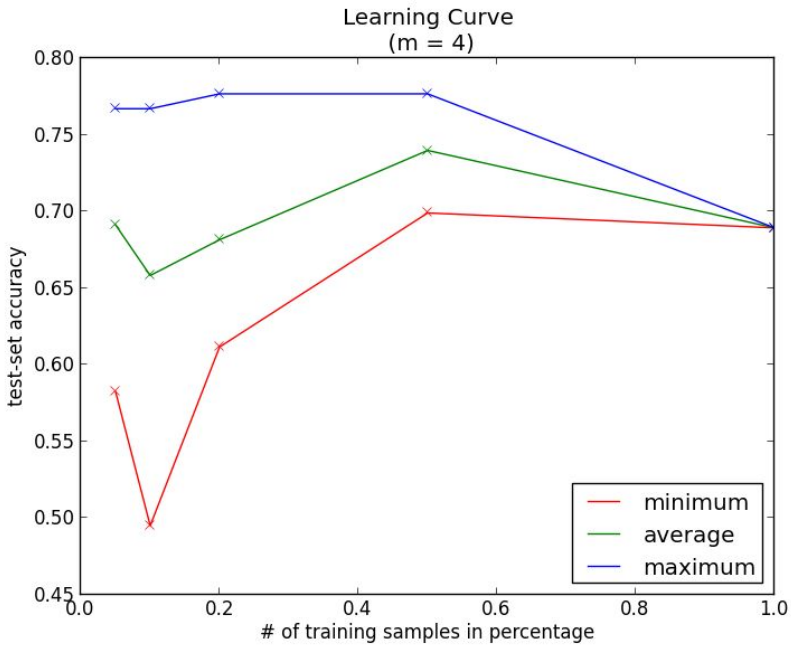
CS 760 Homework 1
Zhenyuan Shen

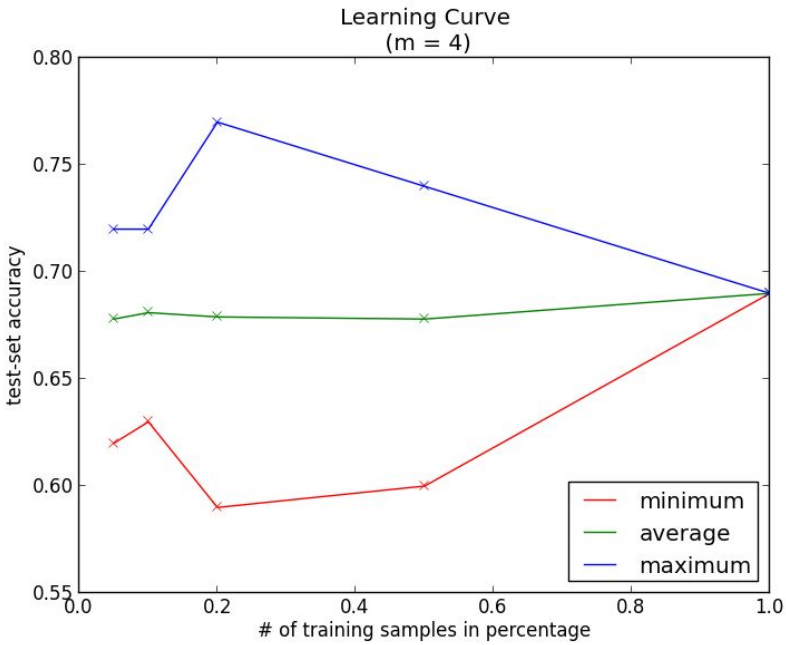
P1.

Usage: python dt-learn.py <train-set-file> <test-set-file> m

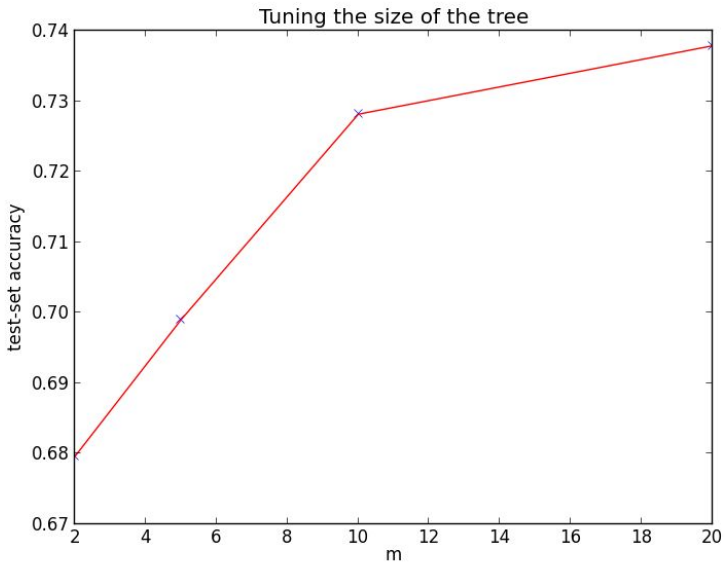
P2.
For the heart disease part, keeping m = 4, I plot the learning curve as follows:



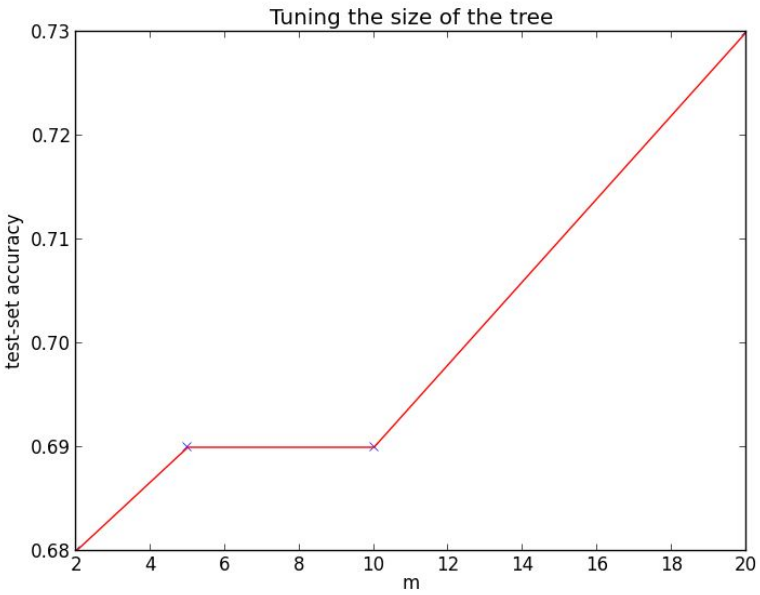
For the diabetes part, keeping m = 4, I plot the learning curve as follows:



P3.
For the heart disease part,



For the diabetes part,


Tuning the size of the tree
test-set accuracy
m
0.73
0.72
0.71
0.70
0.69
0.68
2
4
6
8
10
12
14
16
18
20