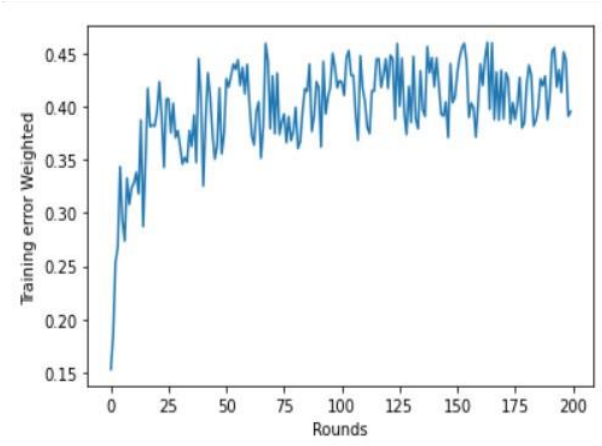
Assignment -3 Implementing Adaboost Algorithm

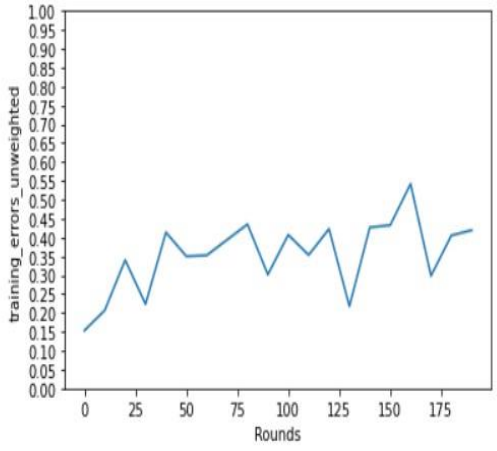
SRUJITHA REDDY MEKALA – G01326037 (miner username:ds18)

No. of Rounds vs Training Error:

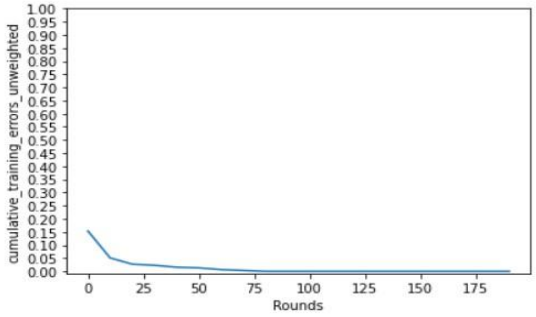
No. of Rounds	Training Error
10	0.15321252059308077
20	0.2067545304777595
30	0.34019769357495877
40	0.22322899505766058
50	0.413509060955519
60	0.3500823723228995
70	0.35255354200988465
80	0.3937397034596376
90	0.43492586490939045
100	0.30148270181219106
110	0.4069192751235585
120	0.3533772652388797
130	0.42257001647446457
140	0.21746293245469517
150	0.4266886326194399
160	0.4324546952224053
170	0.542009884678748
180	0.2981878088962109
190	0.4052718286655683
200	0.4192751235584844

Graphs for training data:



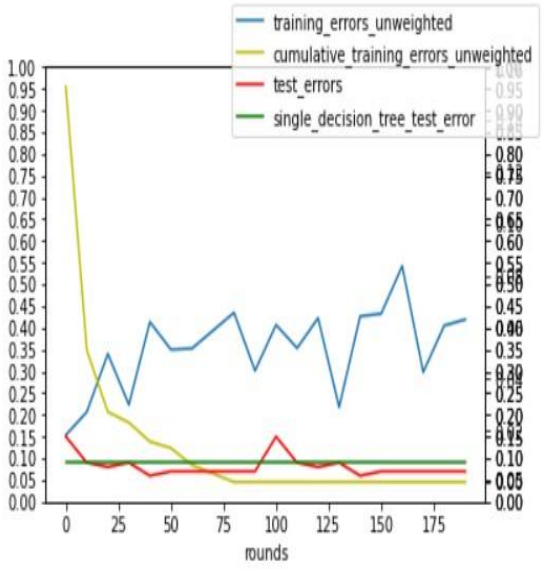


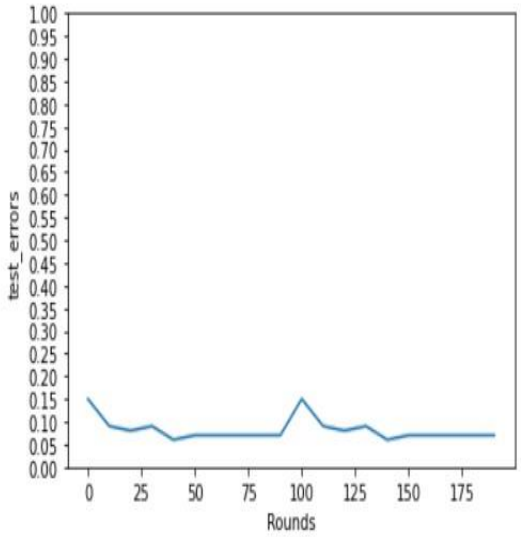
The training error is always less than 0.5 which is the main point of using adaboost algorithm. Here we are using individual weak classifiers whose output is not used again for the next round and hence the error value is increasing. But when all the weak classifiers are used together the cumulative error decreases.



No. of Rounds vs Test-set Error:

No. of Rounds	Accuracy	Test_set error
1	0.85	0.15
10	0.91	0.09
20	0.92	0.08
30	0.91	0.09
40	0.94	0.06
50	0.93	0.07
60	0.93	0.07
70	0.93	0.07
80	0.93	0.07
90	0.93	0.07
100	0.93	0.07
110	0.93	0.07
120	0.93	0.07
130	0.93	0.07
140	0.93	0.07
150	0.93	0.07
160	0.93	0.07
170	0.93	0.07
180	0.93	0.07
190	0.93	0.07
200	0.93	0.07





Here, we can observe that the test error decreases as the no. of rounds increases in the algorithm. At some point later, the test error might become zero. Also, when the no. of rounds is very large the model might undergo overfitting. Also, using adaboost algorithm (red line in the graph) decreases the test error compared to using only single decision tree (green line in graph). The training error decreases faster than the test error and at some point, becomes zero. The model undergoes overfitting at around round 75.