Report

Q 1,2,3,4,5) The Datasets are included in the submission
Q 6)

	Tree1	Tree2	Tree3	Tree4
RMSE_D1	3.16875842701	3.57731775128	3.71807099496	4.09050889632
RMSE_D2	4.45945091627	4.38047409524	4.45312966572	3.92355168332
RMSE_D3	4.09129612901	3.79154495363	3.92322659968	3.98000298929
RMSE_D4	4.19964366137	4.32563985929	4.04923033169	4.34730644635
MAE_D1	2.47304138688	2.92889434550	2.99637593457	3.24020511809
MAE_D2	3.58915488596	3.40471665670	3.44586658183	3.23288120238
MAE_D3	3.33737835568	3.07016855069	3.34724988159	3.34265877341
MAE_D4	3.43188330393	3.62021341094	3.45131228963	3.75330249346

Q 7)

(a) As the number of minimum nodes keeps on increasing from 5 to 40, both the RMSE(Root Mean Squared Error) and MAE(Mean absolute Error) increases for the case of data D1. This trend is observed because on increasing the minimum size of node the height of the tree decreases and separability of the data decreases while training. This leads to a high training error for smaller trees.

(b) Performance of D1 is better compared to that of D2 this is because the tree tends to overfits for training examples and as a result the test dataset has higher values of error. It can also be seen that the difference between values of RMSE and MAE for D1 and D2 decreases on increasing node size of tree. This is because effect of overfit decreases on reducing the tree size.

(c) (i)

In the case of Tree3 the order of both the types of error as follows
RMSE-> D1<D3<D4<D2
MAE-> D1<D3<D2<D4
The errors of D2 ,D3 and D4 are almost same.
Yes, in these cases there is bias variance trade-off, the tree overfitted on D1 decreasing

Error but the error on other testing data increases. Thus on decreasing bias(overfit) variance increases(test error) and vice versa

(ii)
To reduce both bias and variance, large number of overfitted trees can be trained and their average result can be taken to reduce the effect of overfit. Making an overfitted trees will have very low bias but high variance. This high variance can be taken care of by using the average result of all the trees. Thus reduction of both bias and variance can be achieved.

Q 8)

	Tree5	Tree6	Tree7
RMSE_D1	3.92066076202	3.82924364576	3.65381337328
RMSE_D2	3.87088461638	4.17649268994	4.49191948533
RMSE_D3	3.75824139492	3.69991051293	3.98590929328
RMSE_D4	3.84095772865	3.94624357762	4.1762999702
MAE_D1	3.12665271793	3.03862212927	2.98281750956
MAE_D2	3.07923072901	3.34714622614	3.64079413473
MAE_D3	3.14060669383	3.09298810182	3.47434432517
MAE_D4	3.19996998552	3.33430012248	3.36984374217

Q 9)
(a)
The error values are quite low, hence the performance of the trees are very good. The error of testing and training data is almost same, which means there is no overfit. In all the cases this phenomenon is seen because the aggregate predict values from multiple trees are taken tha just one single tree. Thus this simultaneously reduce error and overfit.
(b)
It can be seen that for all the cases the errors for D1(training data) has increased a bit, but the errors for test(D2) and validation data(D3 and D4) have reduced than the previous case.
(c)
When compared among themselves the results of Tree 5(Bagging) is the best followed by Tree 6(Random Forest) and then Tree 7(Adaboost). This is because the errors are nearly equal for

all datasets in bagging. The difference between test and training error is minimum. On going to left and right tree the training error decreases but the test errors increase. This is clear indication that Tree 3 handle the problem of overfitting very well. But the training error is greater for Tree 3 thus it has higher bias. The order is given below
Bias-> Tree 7 < Tree 6 < Tree 5
Variance-> Tree 5 < Tree 6 <Tree 7
(d)
The errors of training, test and validation data for Tree 5,6 and 7 are almost same. Thus the variance of these trees is less that previous ones. On the other hand for the previous trees there is difference between training and test error indicating overfit. But due to large training error the bias of these trees are higher compared to previous ones.
Bias-> Tree 1,2,3,4 < Tree 5,6,7
Variance-> Tree 5,6,7 < Tree 1,2,3,4

Q 10)
Taking the case of bagging tree
The trend is described below
On decreasing the value of minimum size of node the overfit increases and error decreases, and on increasing the number of trees the effect of overfit decrease. These two countering phenomenon can be used to obtain optimum value of the parameters. For the case the optimum values obtained are shown below->
Minimum size of leaf node=8
No of trees= 2000
The Results were as follows-
RMSE-> D1=3.68431857482 D2=3.74924065808
MAE-> D1=2.951125707 D2=2.994040529