All the experiments are on Istanbul dataset.

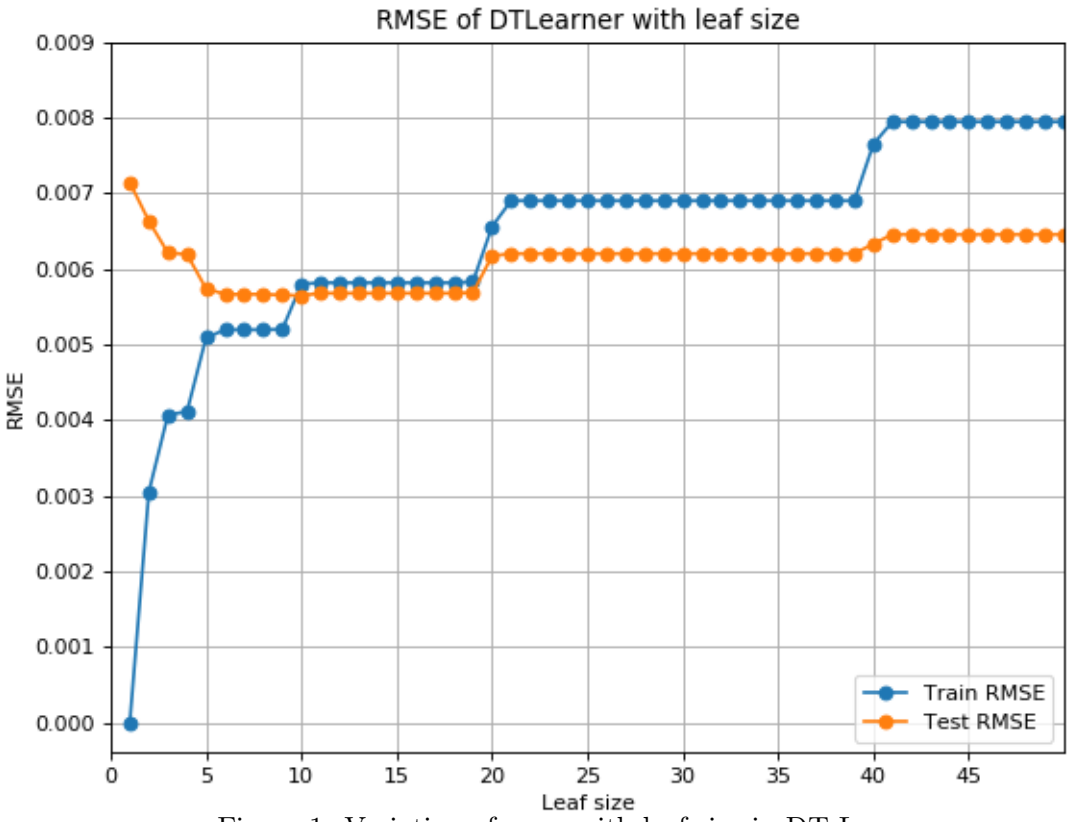

Figure 1: Variation of error with leaf size in DT Learner

1. (a) Training error increases with increasing leaf size, since it allows for more generalization It is around zero when leaf size is 1, since each leaf corresponds to a data sample.

   (b) Test error decreases with leaf size, since leaf size is a regularization parameter. Leaf size controls if the learner is overfitting on training data. Initially the difference between train and test error is large indicating overfitting. From leaf size 5 till around 19, the test error is lowest and stable, while train error keeps on increasing. At leaf size 10, training error becomes higher than test error. The errors start to increase after 19. Overfitting occurs on training data with lower leaf size.

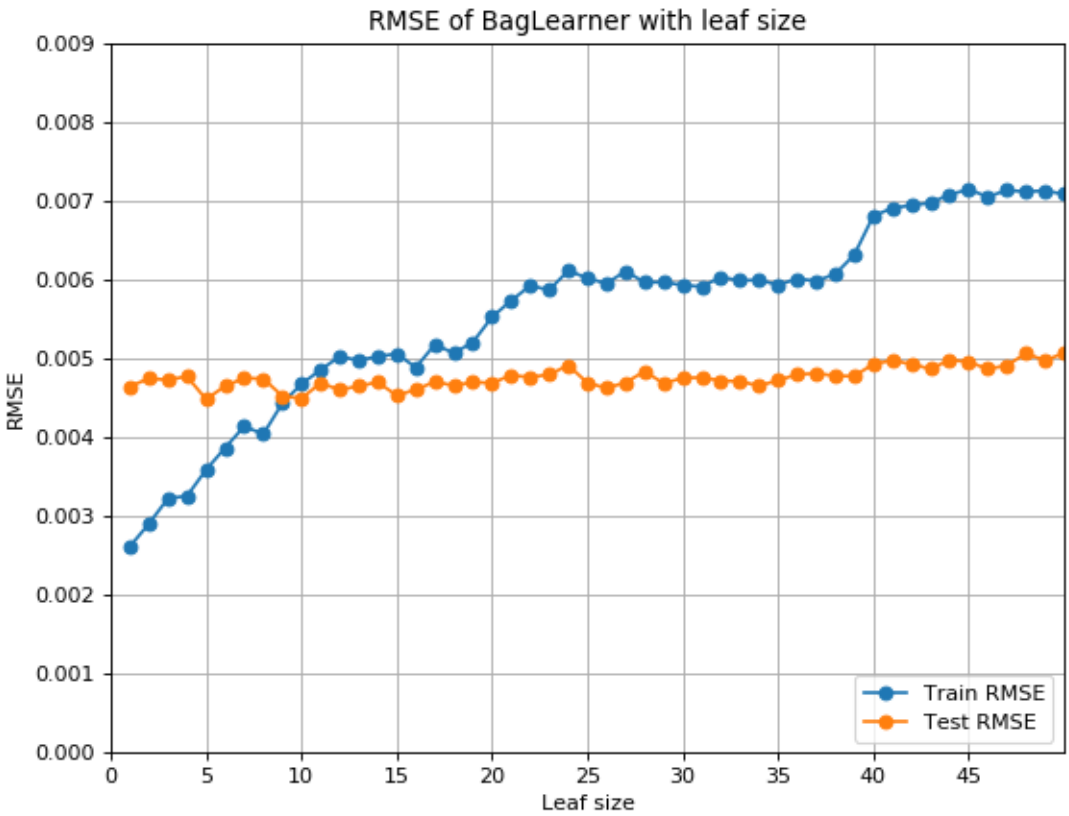

Figure 2: Variation of error with leaf size in Bag Learner

2. Number of bags - 30.

   (a) We can see the test error has reduced using Bag learner compared to DT Learner. The test error for DT Learner is above .005 for all cases, while for Bag Learner the maximum error is around .005. The bag learner gives better performance than DT learner.

   (b) The test errror for bag error does not vary much with leaf size. It does not overfit on lower leaf sizes. Bagging approach does not let the model overfit and test error remains relatively stable to DT Learner. The test error of Bag Learner for leaf size 1 is around .0045 while for DT Learner it is around .007.

   (c) The averaging of results from weak learners which were trained on sampled with replacement data reduces the variance on averaging the results, since the data was correlated.

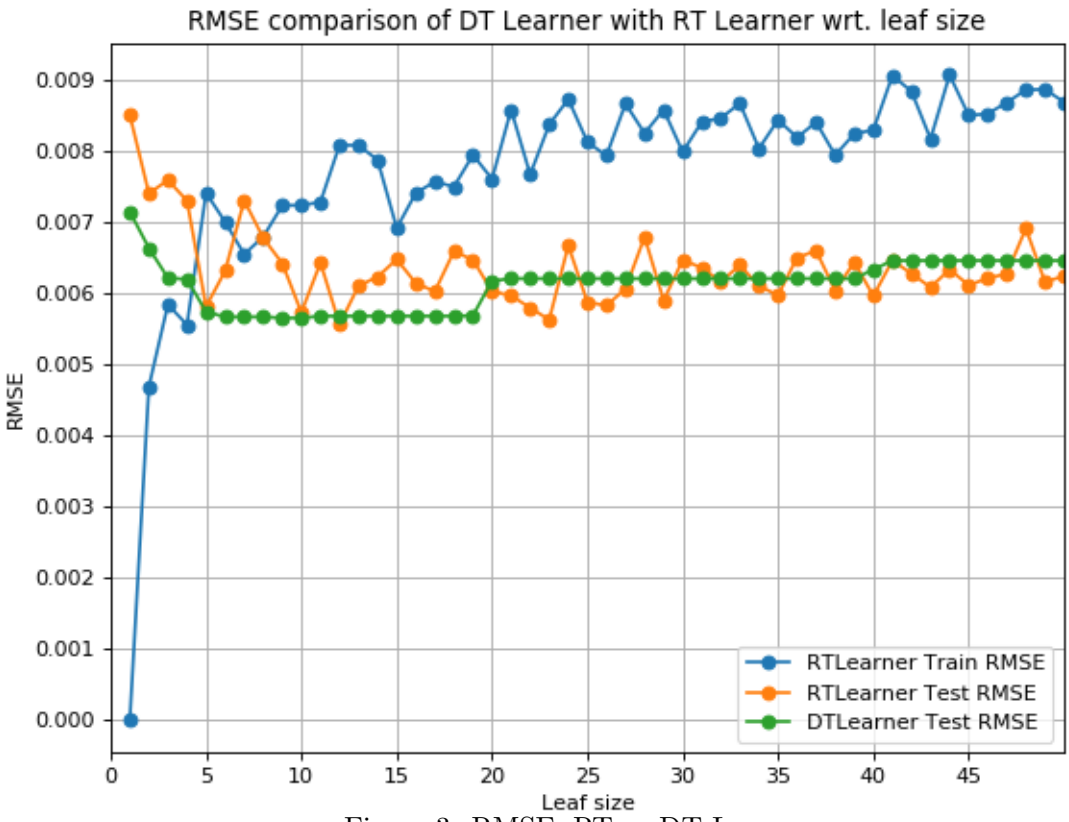

Figure 3: RMSE, RT vs DT Learner

3. In this figure, we compare the performance of RT and DT Learner on test data of Istanbul using RMSE. The RT Learner's test error is quite erratic. It also overfits for lower leaf sizes since there is a sharp drop in error at leaf size 2 and further reducing. In general, DT Learner ismore stable compared to RT, though RT performs better for certain leaf sizes.

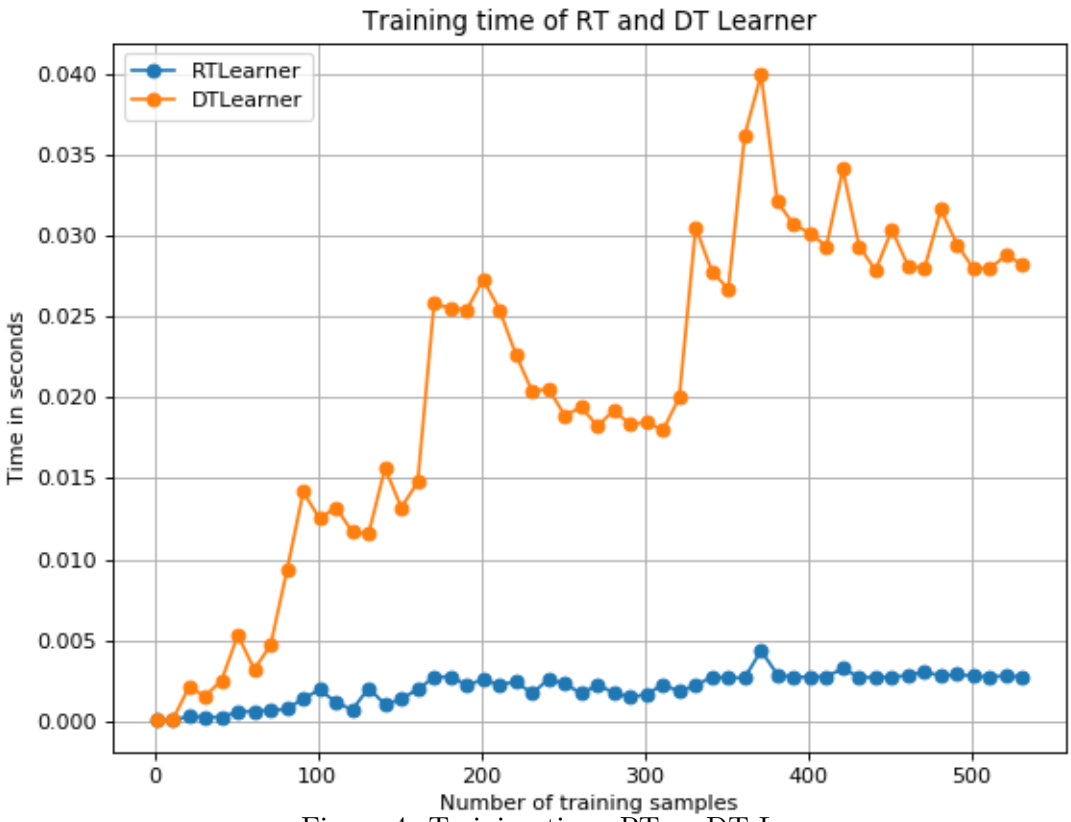

Figure 4: Training time, RT vs DT Learner

For this experiment, I chose randomly n training samples upto 531 from Istanbul data set for measuring the training time wrt number of samples.
For measuring train time, measured the time taken to by the method addEvidence.
The time taken to train DT Learner increases with increasing number of training samples. This is because DT Learner uses correlation for calculating the best feature to split the data while RT Learner chooses randomly without performing any calculations. The time taken to find correlation increases with increasing number of samples and hence training time also increases.

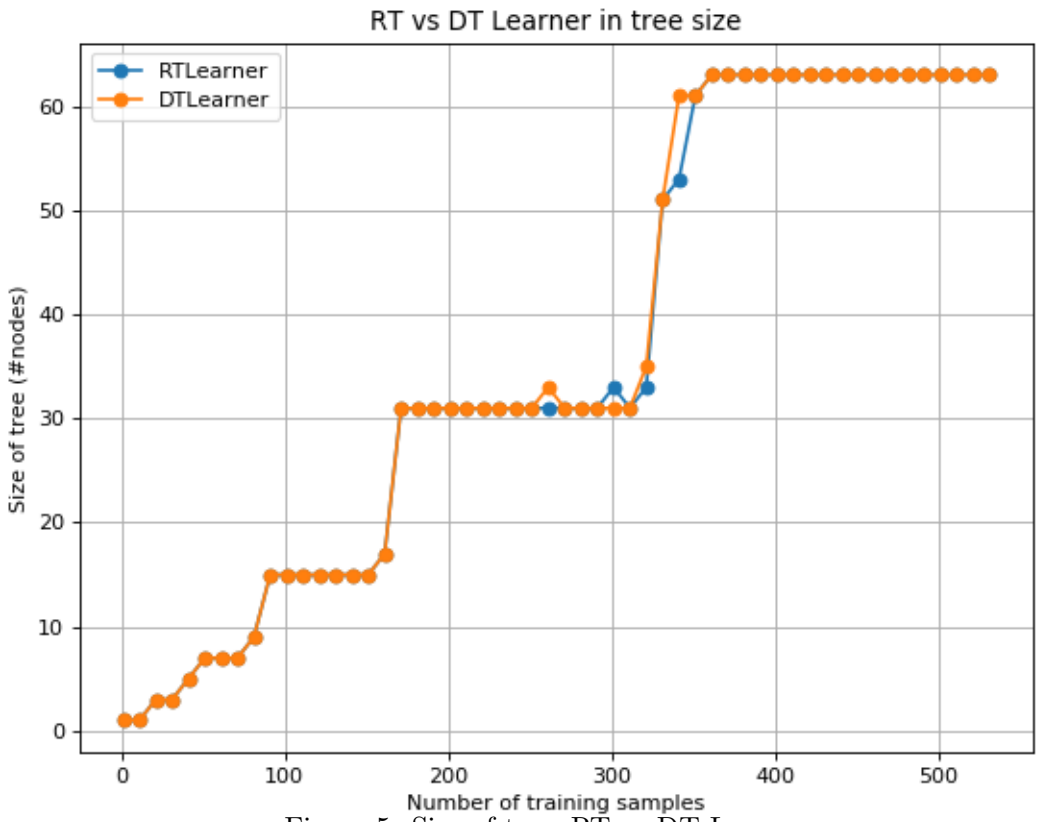

Figure 5: Size of tree, RT vs DT Learner

For this experiment, I chose randomly n training samples upto 531 from Istanbul data set for measuring the training time wrt number of samples.
The size of the tree increases with increasing number of samples for both DT and RT Learner. However, the RT Learner at times has larger size of tree, since it chooses to split on a random feature which can lead to increased number of branches, hence higher nodes.