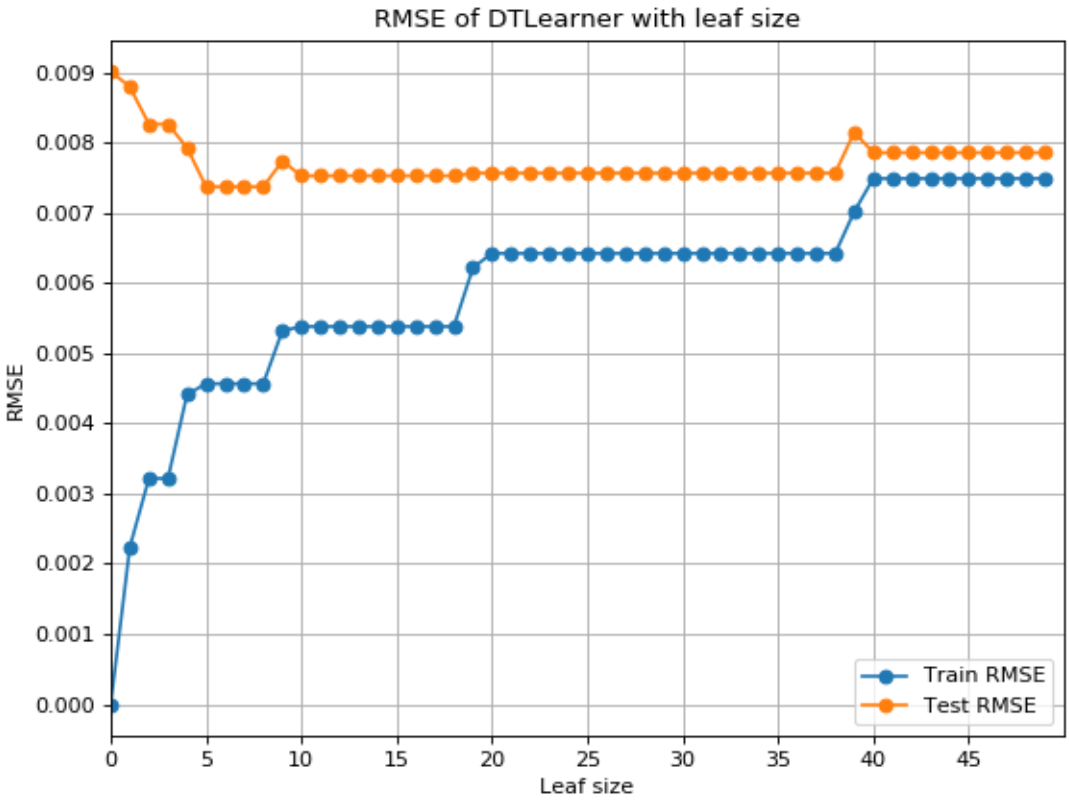


Figure 1: Variation of error with leaf size in DT Learner

1. (a) Training error increases with increasing leaf size, since it allows for more generalization It is 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. At leaf size 10, test error stabilizes till leaf size around 37, while the train error keeps on increasing. This is so because the model is allowing for more generalization.

   (c) Overfitting occurs on training data with lower leaf size.

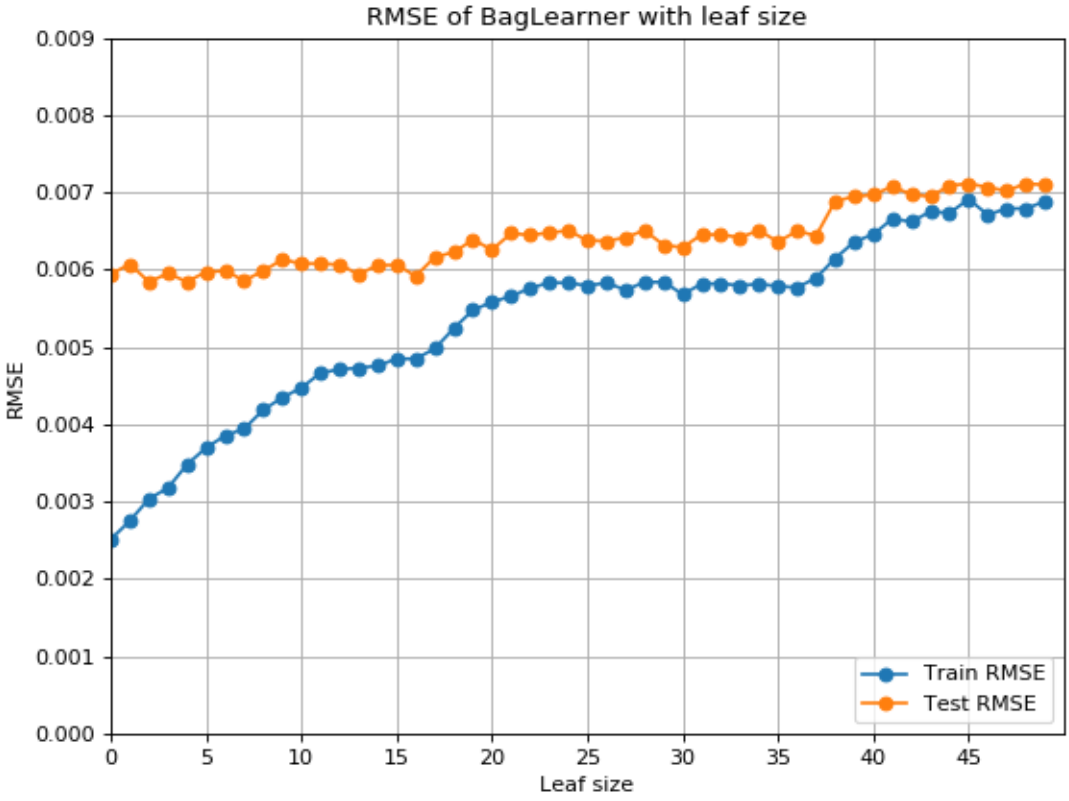


Figure 2: Variation of error with leaf size in Bag Learner

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 .007 for all cases, while for Bag Learner the maximum error is around .007.

   (b) Bag learner does not overfit for lower leaf size on training data. The test error of Bag Learner for leaf size 1 is around .0025 while for DT Learner it is .009.

   (c) The averaging of results from weak learners which were trained on sampled with replacement data reduces the variance in results and avoid overfitting.

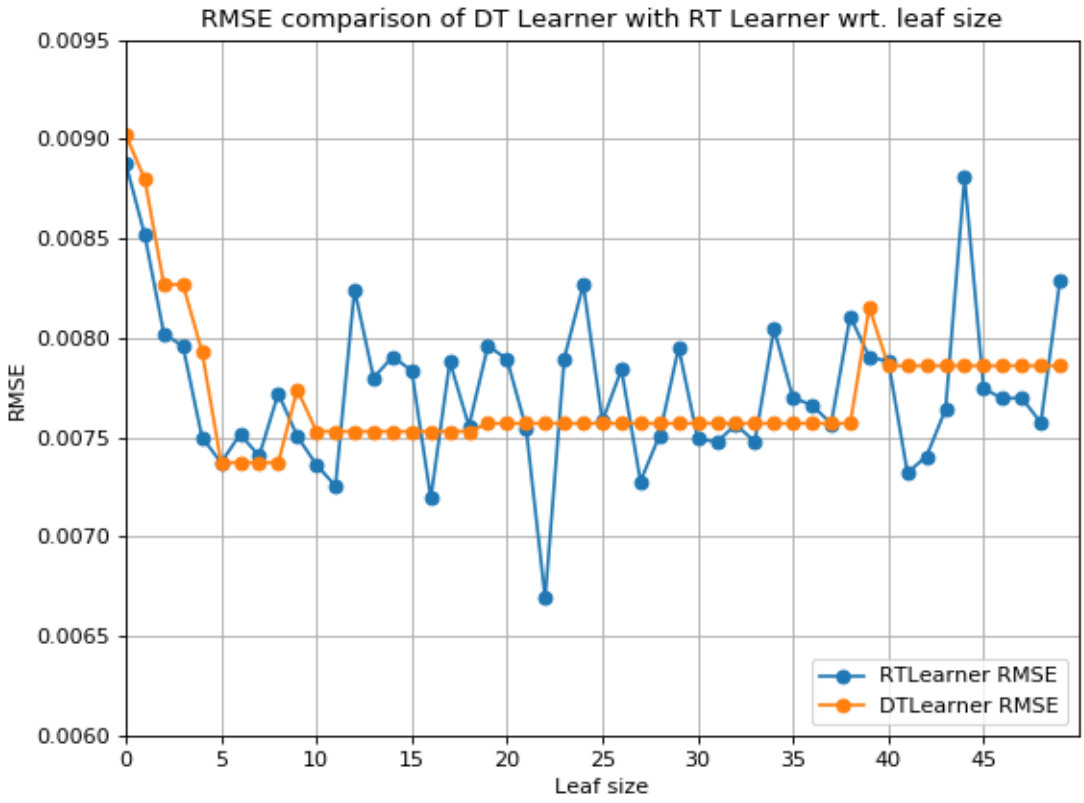


Figure 3: RMSE , RT vs DT Learner

3. The RT Learner's performance is quite erratic. At times, it performs better than DT Learner probably on happening to choose good features for splitting, and at times it performs worse than the RT Learner. However, it performs almost the same as DT Learner on leaf size upto 10, as both of them overfit for lower leaf sizes.

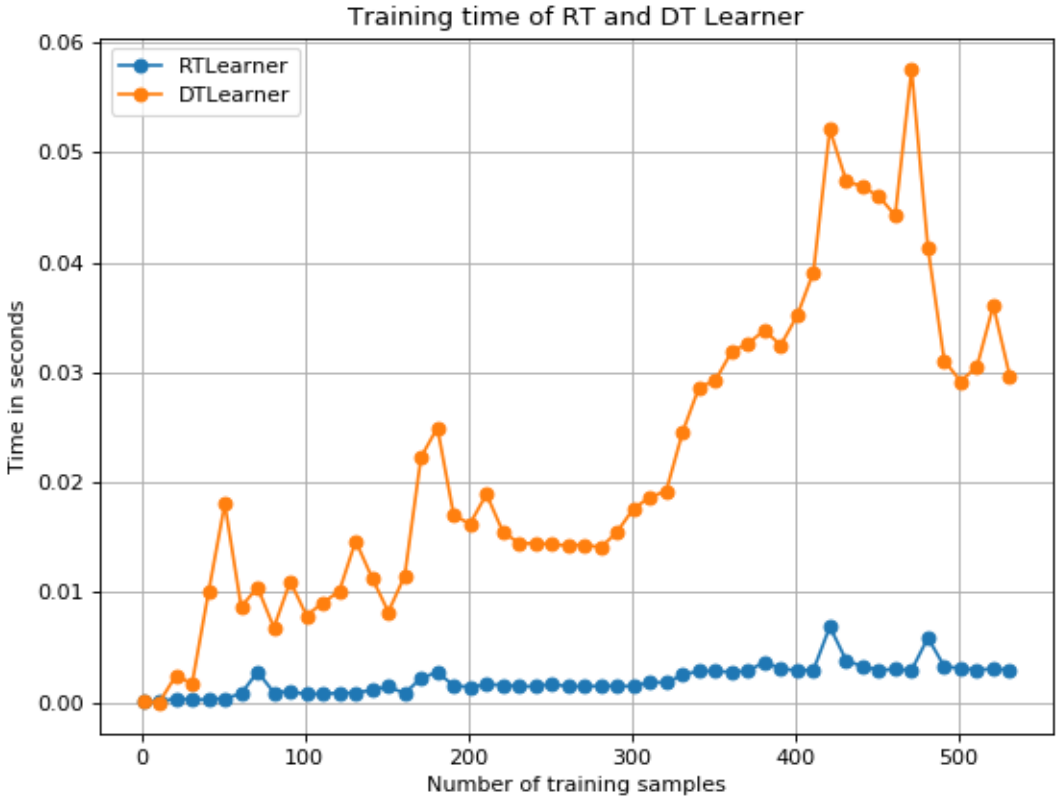


Figure 4: Training time, RT vs DT Learner

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.

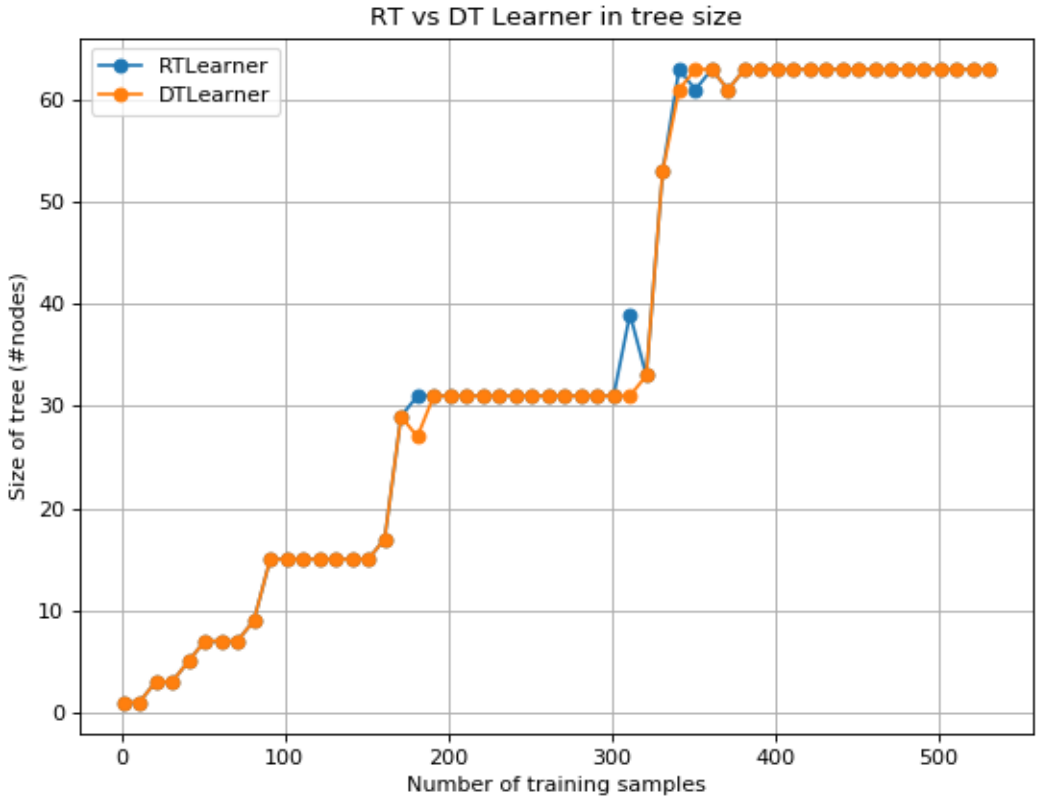


Figure 5: Size of tree, RT vs DT Learner

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.