Report Assignment 3 Asses learner

Rimon Mikhael
rmikhael3@gatech.edu

***Abstract*—**This report discusses the results of running different techniques to test decision knowledge based systems, and the accuracy of those different techniques. those techniques are bagging decision tree, random decision tree and a classic implementation of decision tree. we use coefficient correlation and root mean square error - RMSE - for both in sample and out of sample data to judge those techniques.

> The family's palindromic name emphasizes that its members carry out the Top-Down Induction of Decision Trees.(Quinlan, 1986)

1 INTRODUCTION

The Data was separated to ratios 60% to training the models and 40% to test them. the figures included demonstrate the relationship between over-fitting and leaf size among all the test techniques.

2 METHODS

- Over-Fitting tend to dissipate the higher the leaf count.
- Bagging does reduce over-fitting.
- the root square mean error reduces with more bags
- Random trees produces as accurate results even though the data is random because the repeated testing.

3 EXPERIMENT-1

This is an out-of-simple, in-sample for 1 to 80 leaf sizes DT Learners. Over-fitting happens with small leaf sizes 6 and 19 or less when the in-simple error starts decreasing drastically and the out-of-sample starts increasing. the in-sample is always decreasing when the leaf size is between 1 to 10. so outfitting occurrences are more prominent with leaf size 6 and less. leaf size has an effect on over-fitting.

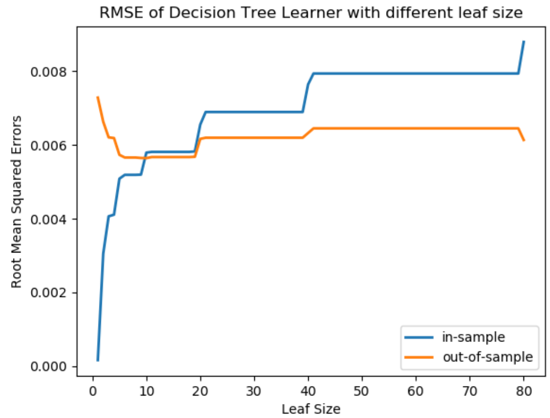
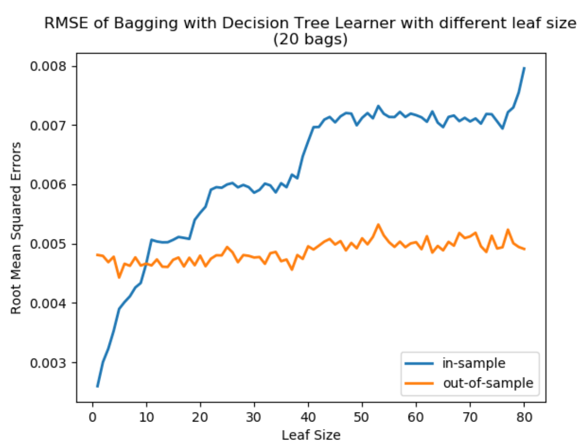


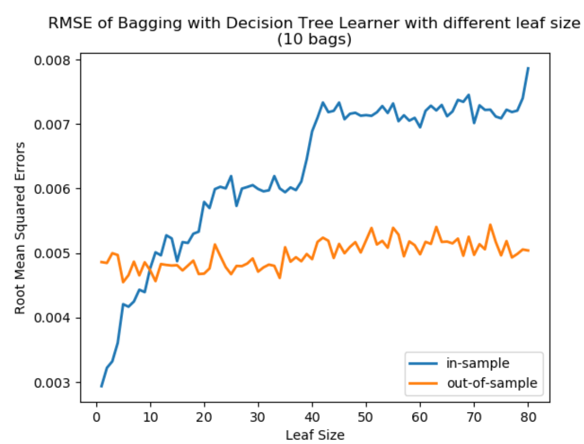
Figure 1—Experiment1 figure 1.

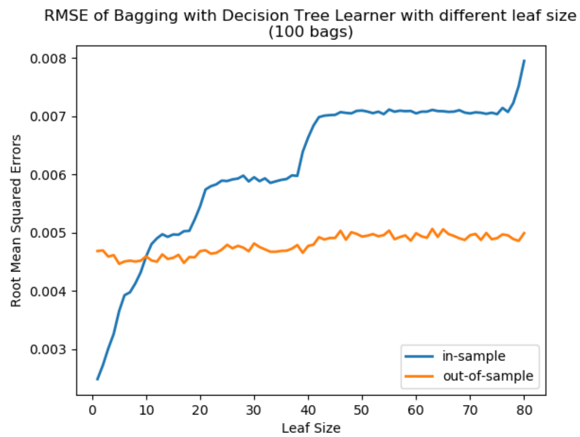
4 EXPERIMENT-2

4.1 Bagging Effect

bagging can reduce over-fitting and stabilize RMSE but can't eliminate it. as per the figures below the higher the count of bags the less over-fitting. The variance of RMSE is smoother the more bags we use it to train with it. so 10 bag has more variance than 20 and then more than 100.







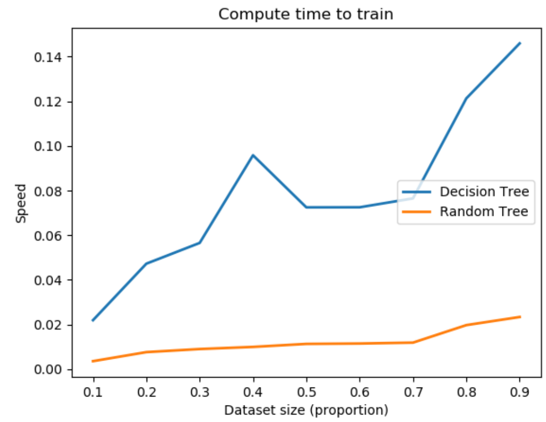
5 EXPERIMENT-3

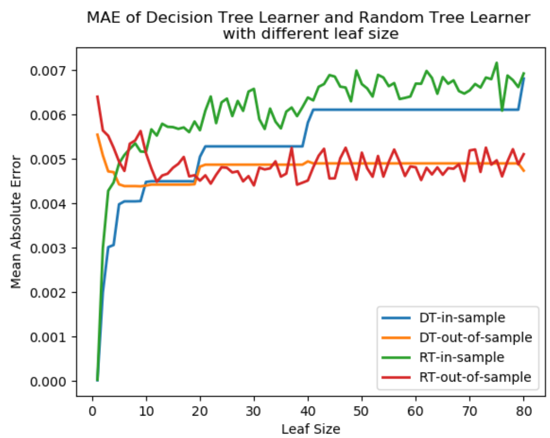
5.1 Time of query and Training

Comparison between DT Classic learner and RT Learner per time to train for the same data simple. In Figure 4a, the DT Learner runs faster than the RT Learner for training when the same data. also, as the data size increases, the DT Learner take more time to train while the training time of the RT Learner stays relatively the same and in figure 4b, the bigger the data set the more time to query though rt learner uses less time than DT learner because of the randomness of data.

5.2 MAE Comparison

Comparison for the MAE of the DT learner and RT learner. RT learner has more variance in MAE. so concisely that means that randomness of data create a lot of variance for the mean of errors between predicted result to the actual results. though MAE is less biased towards higher error count and values but it doesn't reflect performance adequately, while RMSE is a better count for performance.





6 REFERENCES

[1] Quinlan, J. Ross (1986). "Induction of decision trees". In: *Machine learning* 1.1, pp. 81–106.