

Classification vs Regression Learning

In this Project I calculated performance metrics for decision trees, random trees, and bagged learners(for effects on overfitting). The most interesting part of this project was being able to see how overfitting(one metric for assessing accuracy of a given model) changes with respect to leaf_size(a hyperparameter of a decision tree) and how deterministic models compare to non deterministic(random tree) in terms of accuracy using mean absolute error(MAE) and coefficient of determination(R squared). Below are a few results comparing metrics between learners and hyper variables such as leaf size, RMSE

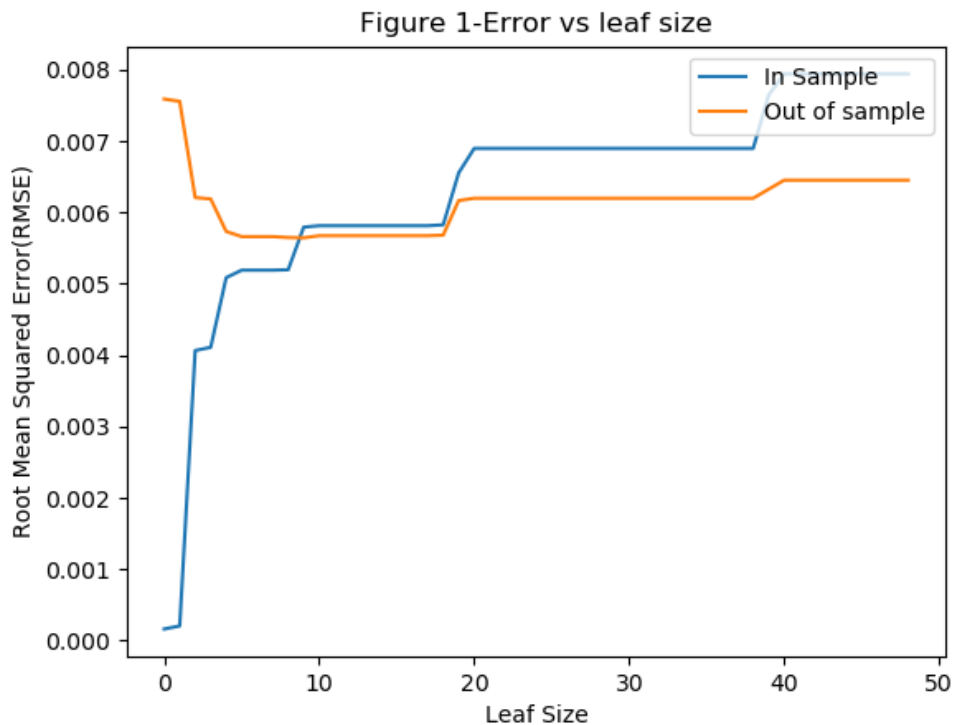


Figure 2-Error vs leaf size with Bagging(bags=20)

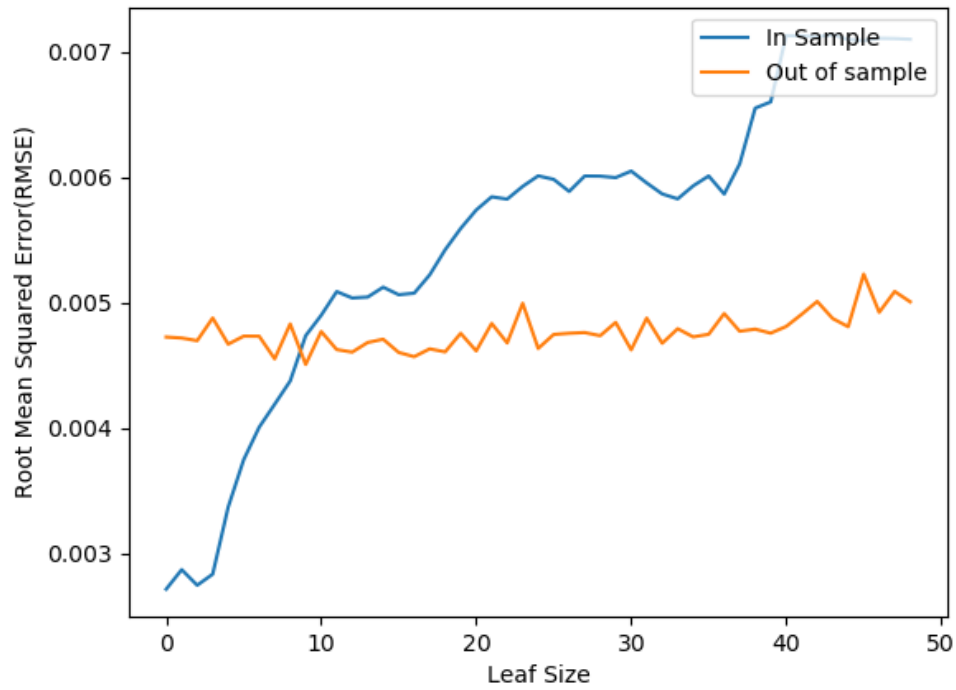


Figure 3- Mean Absolute Error DT vs RT

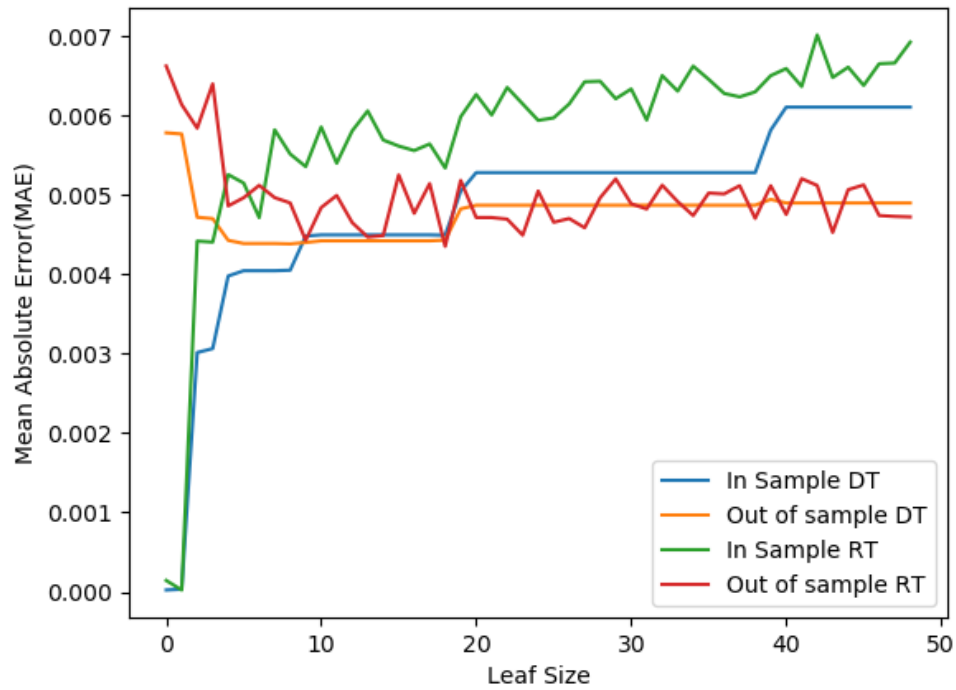


Figure 4- Coefficient of Determination DT vs RT

