

Variable Importance

- Measure the importance of a variable by the improvement of RSS contributed by this variable.
- At each split, attribute the improvement of RSS to the corresponding splitting variable.
- For each variable, accumulate its improvement of RSS across the tree and then averaged over all the trees in the forest.

- Another measure is computed from **permuting** OOB samples: For each tree T_b in the forest, calculate the prediction error (MSE for regression) based on OOB samples. Then the same is done after permuting the j th predictor in the OOB samples. The difference between the two (**before and after** permutation) is then averaged over all trees, and further normalized by the corresponding standard deviation^a.

^aIf the standard deviation of the differences is equal to 0 for a variable, then the division is not applied.