

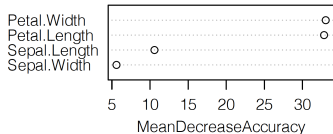
Introduction to Machine Learning

Random Forests: Feature Importance



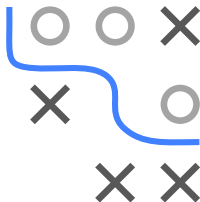
Learning goals

- Understand that the goal of defining variable importance is to enhance interpretability of the random forest
- Know definition of variable importance based on improvement in split criterion
- Know definition of variable importance based on permutations of OOB observations



VARIABLE IMPORTANCE

- Single trees are highly interpretable
- Random forests as ensembles of trees lose this feature
- Contributions of the different features to the model are difficult to evaluate
- Way out: variable importance measures
- Basic idea: by how much would the performance of the random forest decrease if a specific feature were removed or rendered useless?



VARIABLE IMPORTANCE / 2

Measure based on improvement in split criterion

for features $x_j, j = 1$ to p **do**

for tree base learners $\hat{b}^{[m]}, m = 1$ to M **do**

 Find all nodes \mathcal{N} in $\hat{b}^{[m]}$ that use x_j .

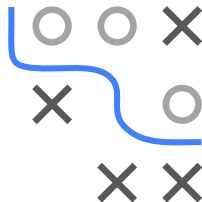
 Compute improvement in splitting criterion achieved by them.

 Add up these improvements.

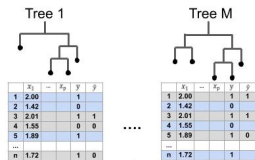
end for

 Add up improvements over all trees to get feature importance of x_j .

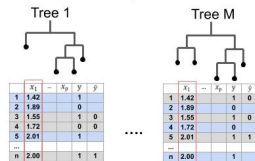
end for



VARIABLE IMPORTANCE / 3



Permutation of feature 1



In-bag observations
 Out-of-bag observations

Measure based on permutations of OOB obs.

Estimate OOB error $\widehat{\text{err}}_{\text{OOB}}$.

for features x_j , $j = 1$ to p **do**

 Perform permutation ψ_j on x_j to distort
 feature-target relation for x_j .

for distorted observations $(\mathbf{x}_{\psi_j}^{(i)}, y^{(i)})$, $i = 1$ to n **do**

 Compute OOB prediction $\hat{y}_{\text{OOB}, \psi_j}^{(i)}$.

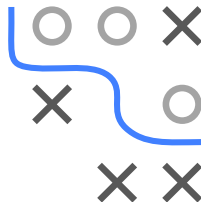
 Compute corresponding loss $L(y^{(i)}, \hat{y}_{\text{OOB}, \psi_j}^{(i)})$.

end for

 Estimate importance of j -th variable

$$\begin{aligned} \widehat{\text{VI}}_j &= \widehat{\text{err}}_{\text{OOB}, \psi_j} - \widehat{\text{err}}_{\text{OOB}} \\ &= \frac{1}{n} \sum_{i=1}^n L(y^{(i)}, \hat{y}_{\text{OOB}, \psi_j}^{(i)}) - \widehat{\text{err}}_{\text{OOB}}. \end{aligned}$$

end for



VARIABLE IMPORTANCE / 4

- Measure based on improvement in split criterion:
MeanDecreaseGini (average total decrease in node impurities, measured by the Gini index)
- Measure based on permutations of OOB observations:
MeanDecreaseAccuracy (average decrease in accuracy for predictions of OOB observations after permuting the j -th feature)

