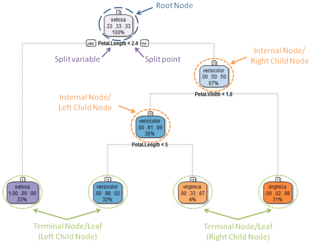
Introduction to Machine Learning

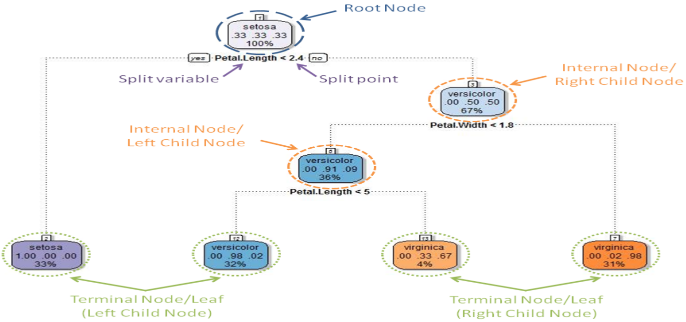
Classification and Regression Trees (CART): Basics



Learning goals

- Understand the basic structure of a tree model
- Understand that the basic idea of a tree model is the same for classification and regression
- Know how the label of a new observation is predicted via CART
- Know hypothesis space of CART

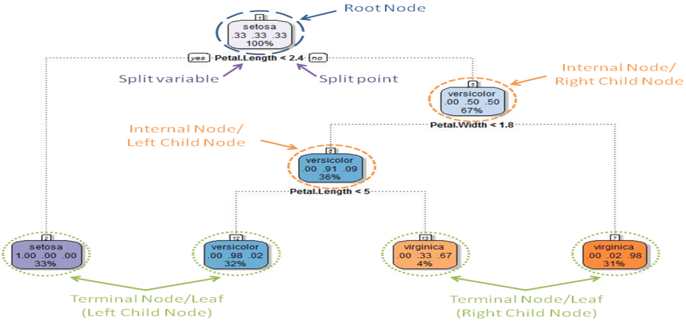
TREE MODEL AND PREDICTION



- Classification and Regression Trees, introduced by Breiman
- Binary splits are constructed top-down
- Constant prediction in each terminal node (leaf): either a numerical value, a class label, or a probability vector over class labels.

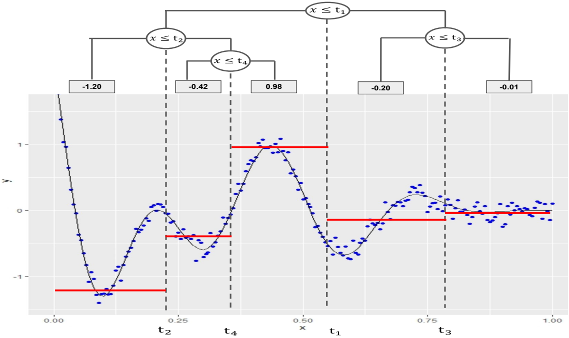
TREE MODEL AND PREDICTION

- For predictions, observations are passed down the tree, according to the splitting rules in each node
- An observation will end up in exactly one leaf node
- All observations in a leaf node are assigned the same prediction for the target



TREE MODEL AND PREDICTION

- For predictions, observations are passed down the tree, according to the splitting rules in each node
- An observation will end up in exactly one leaf node
- All observations in a leaf node are assigned the same prediction for the target

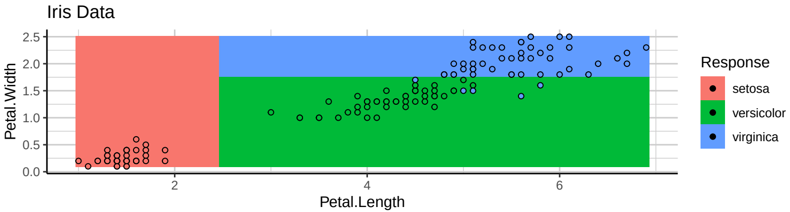


TREE AS AN ADDITIVE MODEL

Each point in $\mathcal{X}$ is assigned to exactly one leaf, and each leaf has a set of input points leading to it through axis-parallel splits.
Hence, trees divide the feature space $\mathcal{X}$ into **rectangular regions**:

$$f(\mathbf{x}) = \sum_{m=1}^{M} c_m \mathbb{I}(\mathbf{x} \in Q_m),$$

where a tree with M leaf nodes defines M “rectangles” Q_m.
c_m is the predicted numerical response, class label or class distribution in the respective leaf node.

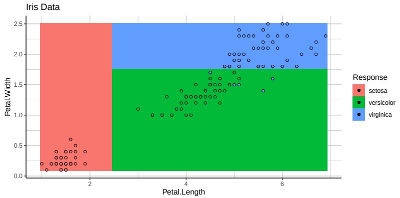


TREES

The hypothesis space of a CART is the set of all step functions over rectangular partitions of $\mathcal{X}$:

$$f(\mathbf{x}) = \sum_{m=1}^{M} c_m \mathbb{I}(\mathbf{x} \in Q_m)$$

Classification:



Regression: