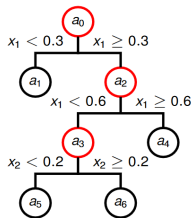# Interpretable Machine Learning

Rule-based Models

Learning goals

- Decision trees
- RuleFit
- Decision rules

DECISION TREES ▸ Breiman et al. (1984)

Idea of decision trees: Partition data into subsets based on cut-off values in features (found by minimizing a split criterion via greedy search) and predict constant mean c_m in leaf node $\mathcal{R}_m$:

$$\hat{f}(x) = \sum_{m=1}^{M} c_m \mathbb{1}_{\{x \in \mathcal{R}_m\}}$$

DECISION TREES ▸ Breiman et al. (1984)

Idea of decision trees: Partition data into subsets based on cut-off values in features (found by minimizing a split criterion via greedy search) and predict constant mean c_m in leaf node $\mathcal{R}_m$:

$$\hat{f}(x) = \sum_{m=1}^{M} c_m \mathbb{1}_{\{x \in \mathcal{R}_m\}}$$

- Applicable to regression and classification
- Able to model interactions and non-linear effects
- Able to handle mixed feature spaces and missing values
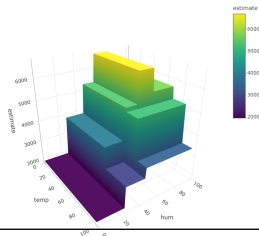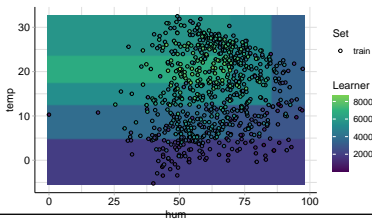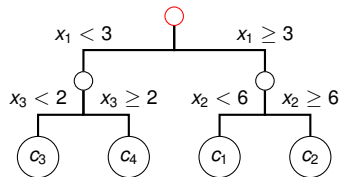
INTERPRETATION

- Directly by following the tree structure (i.e., sequence of decision rules)
- Importance of x_j: Aggregate "improvement in split criterion" over all splits where x_j was involved
  $\rightsquigarrow$ e.g., variance for regression or Gini index for classification

DECISION TREES - EXAMPLE

- Fit decision tree with tree depth of 3 on bike data
- E.g., mean prediction for the first 105 days since 2011 is 1798 (applies to $\hat{=}15\%$ of the data)
- `days_since_2011` shows highest feature importance (explains most of variance)
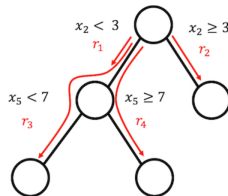
Feature	Importance
days_since_2011	79.53
temp	17.55
hum	2.92
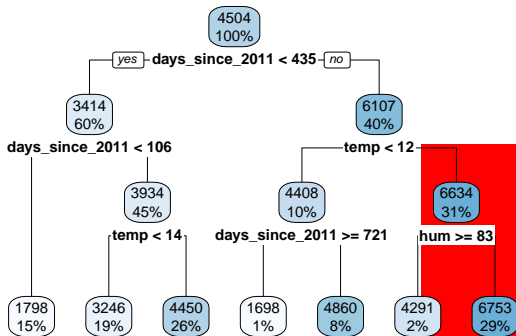
OTHER RULE-BASED MODELS
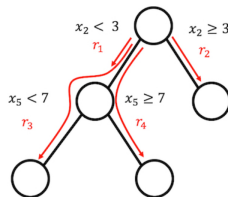
RuleFit ▸ Friedman and Popescu 2008

- Combination of linear models and decision trees
- Allows for feature interactions and non-linearities

OTHER RULE-BASED MODELS

RuleFit ▸ Friedman and Popescu 2008

- Combination of linear models and decision trees
- Allows for feature interactions and non-linearities

Decision Rules ▸ Holte 1993

- Simple "if – then" statements - very intuitive and easy-to-interpret
- Most methods work only for classification and categorical feat.

IF size=small THEN value=low
IF size=medium THEN value=medium
IF size=big THEN value=high

▸ Molnar 2022