Introduction to Machine Learning
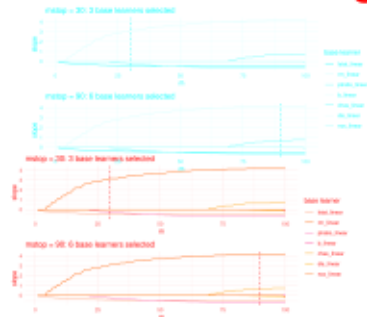
Gradient Boosting: CWB Basics 1

Learning goals

- Concept of CWB
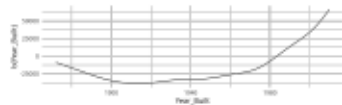- Which base learners do we use
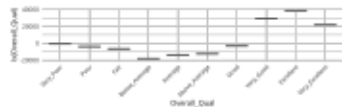- Built-in feature selection

BASE LEARNERS / 2

We restrict the base learners to additive model components, i.e.,
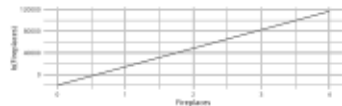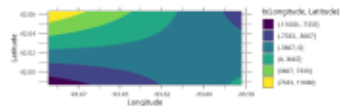

linear effect

non-linear (spline) effect

dummy encoded linear model of a cat. feature

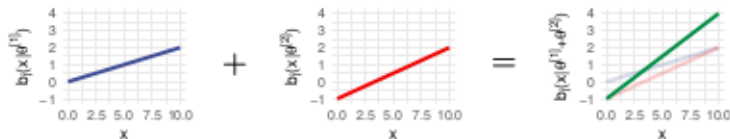tensor product spline for interaction modelling (e.g. spatial effects)

More advanced base learners could also be Markov random fields, random effects, or trees.

BASE LEARNERS / 3

Two BLs of the same type can simply be added by adding up their parameter vectors:

$$b_j(\mathbf{x}, \boldsymbol{\theta}^{[1]}) + b_j(\mathbf{x}, \boldsymbol{\theta}^{[2]}) = b_j(\mathbf{x}, \boldsymbol{\theta}^{[1]} + \boldsymbol{\theta}^{[2]}).$$

Thus, if $\{b_j(\mathbf{x}, \boldsymbol{\theta}^{[1]}), b_j(\mathbf{x}, \boldsymbol{\theta}^{[2]})\} \in \mathcal{B}_j$, then $b_j(\mathbf{x}, \boldsymbol{\theta}^{[1]} + \boldsymbol{\theta}^{[2]}) \in \mathcal{B}_j$.