

## Boosting

Boosting is a way to iteratively increase model complexity by repeatedly reweighting the data to

1. Decide on a "weak" learning model  $b$
2. Initialize a baseline model:  $f(x_0) = 0$
3. For  $M$  iterations:
  - a) Calculate a "stage weight"  $\beta_m$  and model parameters  $\gamma_m$  by solving

$$(\beta_m, \gamma_m) = \arg \min_{\beta, \gamma} \sum_{i=1}^N L(y_i, f_{m-1}(x_i) + \beta b(x_i; \gamma))$$

- b) update the final model by

$$f_m(x) = f_{m-1}(x) + \beta_m b(x; \gamma_m)$$

If  $L$  is a squared error loss, this process is equivalent<sup>1</sup> to iteratively replacing the outcome  $y$  by the current residuals(!), this is because we then have

$$L(y_i, f_{m-1}(x_i) + \beta b(x_i, \gamma)) = (y_i - f_{m-1}(x_i) - \beta b(x_i, \gamma))^2 = (r_{im} - \beta b(x_i, \gamma))^2$$

---

<sup>1</sup>Assuming  $\beta = 1$