Boosting I

Introduction to Boosting; AdaBoost

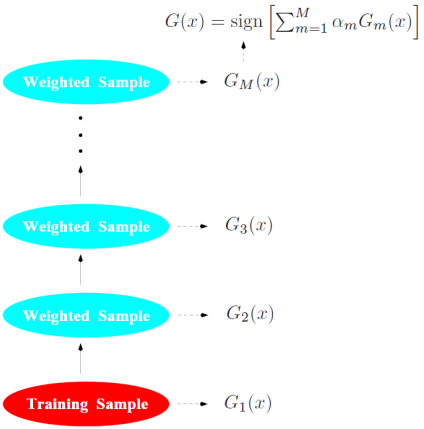
Introduction

Boosting

- Class of ensemble methods which combine **sequential** prediction models
- Adaptive approach with focus on "difficult observations"
- Different flavors exist
  - AdaBoost
  - Gradient Boosting Machines (GBM)
  - ...
- Can be applied to different (weak) base learners
  - Boosting trees
  - ...

AdaBoost

Figure: Process of AdaBoost algorithm[1]



[1] Source: Hastie et al. 2009

AdaBoost

AdaBoost

- Algorithm for classification problems ($Y \in \{-1, 1\}$)
- Estimate a sequence of classifiers using reweighted data
- AdaBoost process
  1. Fit classifier $G_m(x)$ to weighted data (intitial weights $w_i = \frac{1}{n}$)
  2. Compute the misclassification rate

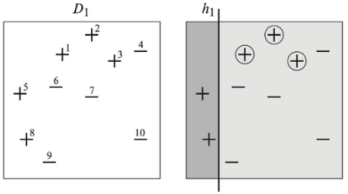
$$\text{err}_m = \frac{\sum_{i=1}^{n} w_i I(y_i \neq G_m(x_i))}{\sum_{i=1}^{n} w_i}$$

  3. Compute the classifier weight $\alpha_m = \log((1 - \text{err}_m)/\text{err}_m)$
  4. Recalculate weights $w_i = w_i \exp(\alpha_m I(y_i \neq G_m(x_i)))$
- Majority vote classification: $G(x) = \text{sign}\left[\sum_{m=1}^{M} \alpha_m G_m(x)\right]$

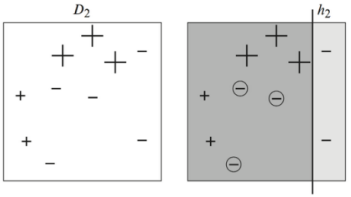
Boosting Stumps

Figure: (Ada)Boosting stumps (example)[2]

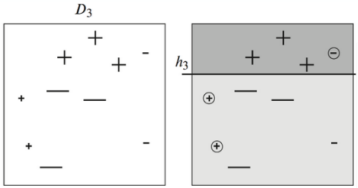
(a) Step 1: $\alpha_1 = 0.42$



(b) Step 2: $\alpha_2 = 0.65$



(c) Step 3: $\alpha_3 = 0.92$



[2] Source: Shapire & Freund 2012

Boosting Stumps

Figure: Step 4: Combine models