Lecture 10 Boosting

Adaboost

- Given training data $(x_i, y_i), i = 1, ..., n$, where $y_i \in Y = \{+1, -1\}$.
- Initialize weights $w_i = 1/n$
- For $t = 1, ..., T$
  1. Fit a weak classifier $\hat{f}_t(x)$ to the training sets with weigts $w_1, ..., w_n$.
  2. Compute the weighted misclassification error

$$\epsilon_t = \frac{\sum_{i=1}^{n} -w_i y_i \hat{f}_t(x_i)}{\sum_{i=1}^{n} w_i}$$

  3. Let $\alpha_t = \ln(\frac{1-\epsilon_t}{\epsilon_t})$. The intuition is that
     - $\epsilon_t = 0$ if $\hat{f}_t$ perfectly classifies all weighted data pts, $\alpha_t = \infty$
     - $\epsilon_t = 1$ if $\hat{f}_t$ perfectly wrongly classifies all weighted data pts, $\alpha_t = -\infty$
     - $\epsilon_t = 0.5$ if ... perform as random guess, $\alpha_t = 0$
  4. Update weights as

$$w_i \leftarrow w_i \times \exp[-\alpha_t y_i \hat{f}_t(x_i)]$$

     The rational is that wee increase the weight if wrong on pt i, i.e.,
     $y_i \hat{f}_t(x_i) = -1 < 0$

Note that the normalization of ϵ_t in step 2) can also be performed in the weights update step 4).