

Correction TD6: AdaBoost on binary classification

Exercise 10: $(x_i, y_i)_{1 \leq i \leq n} \in (X \times \{-1, 1\})^n$

$$\mathcal{H} = \{h_1, \dots, h_m\} \quad h_i: X \rightarrow \{-1, 1\}$$

$$h \in \mathcal{H} \text{ iff } -h \in \mathcal{H} \quad \exists i, j, h(x_i) = y_i, h(x_j) \neq y_j$$

$$\mathcal{F} = \left\{ \sum_{j=1}^n \theta_j h_j : \theta \in \mathbb{R}^n \right\}$$

$$\hat{f}_0 = 0$$

$$\hat{f}_m = \hat{f}_{m-1} + \beta_m h_{j_m}, \quad (\beta_m, h_{j_m}) \underset{h \in \mathcal{H}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^n e^{-y_i (\hat{f}_{m-1} + \beta h)(x_i)}$$

1. $\omega_i^m = \frac{1}{n} e^{-y_i \hat{f}_{m-1}(x_i)}$

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n e^{-y_i (\hat{f}_{m-1} + \beta h)(x_i)} &= \frac{1}{n} \sum_{i=1}^n \left[\underbrace{\mathbb{1}_{h(x_i) = y_i}}_{+ \mathbb{1}_{h(x_i) \neq y_i}} e^{-y_i (\hat{f}_{m-1} + \beta h)(x_i)} \right] \\ &= \frac{1}{n} \sum_{i=1}^n \left[\mathbb{1}_{h(x_i) = y_i} e^{-y_i (\hat{f}_{m-1} + \beta h)(x_i)} + \mathbb{1}_{h(x_i) \neq y_i} e^{-y_i (\hat{f}_{m-1} + \beta h)(x_i)} \right] \end{aligned}$$

$$= (e^\beta - e^{-\beta}) \frac{1}{n} \sum_{i=1}^n \omega_i^m \mathbb{1}_{h(x_i) \neq y_i} + e^{-\beta} \sum_{i=1}^n \omega_i^m$$

2. $h \in \mathcal{H}$ iff $\beta - h \in \mathcal{H}$ so we can only look at the $\beta \geq 0$.

Hence, $\forall \beta \geq 0, (e^\beta - e^{-\beta}) \geq 0$ and so, the function

$$\beta \mapsto (e^\beta - e^{-\beta}) \sum_{i=1}^n w_i^m \mathbb{1}_{h(x_i) \neq y_i} + e^{-\beta} \sum_{i=1}^n w_i^m$$

is minimal for

$$h_{\beta m} = \operatorname{argmin} \operatorname{err}_m(h)$$

which is indep't of β .

Then we minimise in β and we get the result.

3. Apply the rule decided above.