Worked Example

AdaBoost with Weighted Gini (2 Rounds)

Dataset

We have 4 labeled 1D examples:

n	x_n	t_n
1	0.1	$+1$
2	0.3	$+1$
3	0.6	-1
4	0.8	-1

Let each base learner be a **decision stump** of the form:

$$h(x;\theta) = \begin{cases} +1 & \text{if } x < \theta \\ -1 & \text{otherwise} \end{cases}$$

Round 1: Initialize sample weights

Set initial weights:

$$w_n^{(1)} = \frac{1}{4}, \quad \text{for all } n$$

Evaluate candidate thresholds

Candidate thresholds between data points:
$\theta \in \{0.2, 0.45, 0.7\}$

We use **weighted Gini impurity** to choose the best stump.

Gini impurity of a region R Let w_+ be the total weight of positive examples in R, w_- for negative. Define:

- $W_R = w_+ + w_-$ (total weight in region)
- $p_+ = \frac{w_+}{W_R}$, $p_- = \frac{w_-}{W_R} = 1 - p_+$

Then:

$$G(R) = 1 - p_+^2 - p_-^2$$

Total cost of a split Let W_L, W_R be total weights left and right of split, then:

$$G_{\text{split}} = \frac{W_L}{W_L + W_R} G_L + \frac{W_R}{W_L + W_R} G_R$$

Try split at $\theta = 0.2$

- **Left**: point 1 ($t = +1$) $\rightarrow G_L = 0$
- **Right**: points 2, 3, 4 ($t = [+1, -1, -1]$)

Weights:

- $w_+ = 0.25$, $w_- = 0.5$
- $p_+ = \frac{1}{3}$, $p_- = \frac{2}{3}$

$$G_R = 1 - \left(\frac{1}{3}\right)^2 - \left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

Total weighted Gini:

$$G = \frac{0.25}{1.0} \cdot 0 + \frac{0.75}{1.0} \cdot \frac{4}{9} = \frac{1}{3}$$

We choose this stump (not perfect, demonstrates reweighting).

Train first stump $h_1(x)$

Let:

$$h_1(x) = \begin{cases} +1 & x < 0.2 \\ -1 & x \geq 0.2 \end{cases}$$

Predictions: $[+1, -1, -1, -1]$
Truth: $[+1, +1, -1, -1]$

Only point 2 is misclassified $\rightarrow$ weighted error:

We can rewrite ϵ_1 from the notes as

$$\epsilon_1 = \sum_{n=1}^{N} w_n^{(1)} \cdot \mathbb{I}\left(h_1(x_n) \neq t_n\right)$$

since weights are normalized.

Then

$$\epsilon_1 = w_2 = 0.25$$

Compute model weight β_1

$$\beta_1 = \frac{1}{2} \ln\left(\frac{1 - \epsilon_1}{\epsilon_1}\right) = \frac{1}{2} \ln(3) \approx 0.5493$$

Update weights

$$w_n^{(2)} = w_n^{(1)} \cdot \exp(-\beta_1 t_n h_1(x_n))$$

n	t_n	$h_1(x_n)$	$t_n h$	$\exp(-\beta_1 t_n h)$	$w_n^{(2)}$ (unnormalized)
1	$+1$	$+1$	$+1$	0.577	$0.25 \cdot 0.577 = 0.144$
2	$+1$	-1	-1	1.722	$0.25 \cdot 1.722 = 0.430$
3	-1	-1	$+1$	0.577	0.144
4	-1	-1	$+1$	0.577	0.144

Normalize:

$$Z = 0.862$$

Final weights:

- $w_1 = 0.167$, $w_2 = 0.499$, $w_3 = 0.167$, $w_4 = 0.167$

Round 2: Train next stump

Try threshold $\theta = 0.45$:

- Left: $x_1, x_2 \to +1$, $+1 \to$ pure $\to G_L = 0$
- Right: $x_3, x_4 \to -1$, $-1 \to$ pure $\to G_R = 0$

Perfect split (assume small $\epsilon_2 = 0.01$ for numerical stability)

Then:

$$\beta_2 = \frac{1}{2} \ln\left(\frac{0.99}{0.01}\right) \approx 2.30$$

Final Classifier

$$f(x) = \beta_1 h_1(x) + \beta_2 h_2(x)$$

Prediction:

$$y(x) = \text{sign}(f(x))$$

Summary

- **Weighted Gini** used to evaluate splits with current AdaBoost weights
- Round 1 made a mistake $\rightarrow$ weights increased for misclassified point
- Round 2 corrected mistake with a better stump
- Final model is:

$$f_M(x) = \sum_{m=1}^{M} \beta_m h_m(x)$$