\[ H_{\text{final}} = \text{sign} \left( \begin{array}{c} 0.42 \\ +0.65 \\ +0.92 \end{array} \right) \]
Training set: \( \{(x_1, y_1), ..., (x_m, y_m)\} \subseteq X \times \{-1, 1\} \)

Let \( D_1(i) = \frac{1}{m} \).

At each iteration \( t \):
1. Find weak learner \( h_t \) minimizing
   \[
   \varepsilon_t = \sum_{i=1}^{m} D_t(i) \cdot 1(h_t(x_i) \neq y_i).
   \]
2. Set \( D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)) \), where
   \[
   \alpha_t = \frac{1}{2} \ln \frac{1 - \varepsilon_t}{\varepsilon_t}
   \]
   and
   \[
   Z_t = \sum_{j=1}^{m} D_t(j) \exp \left( -\alpha_t y_j h_t(x_j) \right)
   \]
Adaboost uses this weighting mechanism to “force” the weak learner to focus on the problematic examples in the next iteration.

Formally,

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)), \quad \alpha_t = \frac{1}{2} \ln \frac{1-\varepsilon_t}{\varepsilon_t} \]

and

\[ Z_t = \sum_{j=1}^{m} D_t(j) \exp \left(-\alpha_t y_j h_t(x_j) \right) \]

Why?

Adaboost uses this weighting mechanism to focus on the problematic examples in the next iteration.

Formally,

\[ \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) = \frac{1}{2} \]
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)), \text{ where } \alpha_t = \frac{1}{2} \ln \frac{1-\varepsilon_t}{\varepsilon_t}. \]

What is \( D_{t+1}(i) \) when \( h_t(x_i) = y_i \)?

Your answer should only be in terms of \( \varepsilon_t, D_t(i), \) and \( Z_t \).
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)) \], where \( \alpha_t = \frac{1}{2} \ln \frac{1-\varepsilon_t}{\varepsilon_t} \).

What is \( D_{t+1}(i) \) when \( h_t(x_i) = y_i \)?

Your answer should only be in terms of \( \varepsilon_t, D_t(i), \) and \( Z_t \).

Answer.

\[ \frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)) = \frac{1}{Z_t} D_t(i) \exp(-\alpha_t) = \frac{1}{Z_t} D_t(i) \sqrt{\frac{\varepsilon_t}{1-\varepsilon_t}}. \]
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \sqrt{\frac{\varepsilon_t}{1-\varepsilon_t}} \text{ when } h_t(x_i) = y_i. \]

What is \( \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) = y_i) \)?

Your answer should only be in terms of \( Z_t \) and \( \varepsilon_t \).

\[ \varepsilon_t = \sum_{i=1}^{m} D_t(i) \cdot 1(h_t(x_i) \neq y_i) \]
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \sqrt{\frac{\varepsilon_t}{1-\varepsilon_t}} \] when \( h_t(x_i) = y_i \).

What is \( \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) = y_i) \)?

Your answer should only be in terms of \( Z_t \) and \( \varepsilon_t \).

\[ \varepsilon_t = \sum_{i=1}^{m} D_t(i) \cdot 1(h_t(x_i) \neq y_i) \]

Answer.

\[ \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) = y_i) = \frac{\sqrt{\varepsilon_t(1-\varepsilon_t)}}{Z_t} \]
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)), \text{ where } \alpha_t = \frac{1}{2} \ln \frac{1-\varepsilon_t}{\varepsilon_t}. \]

What is \( D_{t+1}(i) \) when \( h_t(x_i) \neq y_i \)?

Your answer should only be in terms of \( \varepsilon_t, D_t(i), \) and \( Z_t \).
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)), \text{ where } \alpha_t = \frac{1}{2} \ln \frac{1-\varepsilon_t}{\varepsilon_t}. \]

What is \( D_{t+1}(i) \) when \( h_t(x_i) \neq y_i \)?
Your answer should only be in terms of \( \varepsilon_t, D_t(i), \) and \( Z_t \).

Answer.

\[
\frac{1}{Z_t} D_t(i) \exp(-\alpha_t y_i h_t(x_i)) = \frac{1}{Z_t} D_t(i) \exp(\alpha_t) = \frac{1}{Z_t} D_t(i) \sqrt{\frac{1 - \varepsilon_t}{\varepsilon_t}}.
\]
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \sqrt{\frac{1-\varepsilon_t}{\varepsilon_t}} \] when \( h_t(x_i) \neq y_i \).

What is \( \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) \)?

Your answer should only be in terms of \( Z_t \) and \( \varepsilon_t \).

\[ \varepsilon_t = \sum_{i=1}^{m} D_t(i) \cdot 1(h_t(x_i) \neq y_i) \]
Question.

\[ D_{t+1}(i) = \frac{1}{Z_t} D_t(i) \sqrt{\frac{1-\varepsilon_t}{\varepsilon_t}} \text{ when } h_t(x_i) \neq y_i. \]

What is \( \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) \)?

Your answer should only be in terms of \( Z_t \) and \( \varepsilon_t \).

\[ \varepsilon_t = \sum_{i=1}^{m} D_t(i) \cdot 1(h_t(x_i) \neq y_i) \]

Answer.

\[ \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) = \frac{\sqrt{\varepsilon_t(1-\varepsilon_t)}}{Z_t} \]
Question.

We saw that

\[
\sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) = \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) = y_i)
\]

\[
= \frac{\sqrt{\varepsilon_t(1 - \varepsilon_t)}}{Z_t}
\]

Why does this mean that \(\sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) = \frac{1}{2}\)?
Answer.

\[
z_1 = \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) = \mathbb{P}_{(x,y) \sim D_{t+1}}[1(h_t(x) \neq y)]
\]

\[
z_2 = \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) = y_i) = \mathbb{P}_{(x,y) \sim D_{t+1}}[1(h_t(x) = y)]
\]

\[z_1 + z_2 = 1 \text{ and } z_1 = z_2.\]

Therefore,

\[
z_1 = \sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) = \frac{1}{2}.
\]
Adaboost uses this weighting mechanism to “force” the weak learner to focus on the problematic examples in the next iteration.

Therefore,

\[
\sum_{i=1}^{m} D_{t+1}(i) \cdot 1(h_t(x_i) \neq y_i) = \frac{1}{2}.
\]