AdaBoost Algorithm
(Yoav Freund and Robert Schapire)

Input: \((x_1, y_1), \ldots, (x_m, y_m)\), where \(x_i \in X, y_i \in \{-1, +1\}\).
Output: A classifier of increased accuracy compared to our weak base classifier.

Initialize: \(D_1(i) = \frac{1}{m}\) for \(i \in \{1, \ldots, m\}\).

For \(t = 1, \ldots, T\):
- Train weak learner using distribution \(D_t\).
- Get weak hypothesis \(h_t : X \rightarrow \{-1, +1\}\).
- Aim: select \(h_t\) to minimize the weighted error:
  \[
  \varepsilon_t = \Pr_{i \sim D_t}(h_t(x_i) \neq y_i).
  \]
- Choose \(\alpha_t = \frac{1}{2} \ln \left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)\).
- Update, for \(i = 1, \ldots, m\):
  \[
  D_{t+1}(i) = \frac{D_t(i)}{Z_t} \cdot \begin{cases} 
  e^{-\alpha_t} & \text{if } h_t(x_i) = y_i \\
  e^{\alpha_t} & \text{if } h_t(x_i) \neq y_i
  \end{cases}
  = \frac{D_t(i) \cdot e^{-\alpha_t \cdot y_i \cdot h_t(x_i)}}{Z_t},
  \]
  where \(Z_t\) is a normalization factor (chosen so that \(D_{t+1}\) is a distribution).

Output the final hypothesis:
\[
H(x) = \text{sgn} \left(\sum_{t=1}^{T} \alpha_t h_t(x)\right),
\]
where \(\text{sgn}(z) = \begin{cases} 
+1 & \text{if } z > 0 \\
-1 & \text{otherwise}
\end{cases}\).