

What AdaBoost Minimizes

- recall proof of training error bound:
 - training error(H_{final}) $\leq \prod_t Z_t$
 - $Z_t = \epsilon_t e^{\alpha_t} + (1 - \epsilon_t) e^{-\alpha_t} = 2\sqrt{\epsilon_t(1 - \epsilon_t)}$
- closer look:
 - α_t chosen to minimize Z_t
 - h_t chosen to minimize ϵ_t
 - same as minimizing Z_t
(since increasing in ϵ_t on $[0, 1/2]$)
- so: both AdaBoost and weak learner minimize Z_t on round t
 - equivalent to **greedily** minimizing $\prod_t Z_t$