

Optimal Efficiency

[Freund]

- for AdaBoost, saw: training error $\leq e^{-2\gamma^2 T}$
- is AdaBoost most efficient boosting algorithm?
no!
- given T rounds and γ -weak learning assumption, boost-by-majority (BBM) algorithm is provably exactly best possible:

$$\text{training error} \leq \sum_{j=0}^{\lfloor T/2 \rfloor} \binom{T}{j} \left(\frac{1}{2} + \gamma\right)^j \left(\frac{1}{2} - \gamma\right)^{T-j}$$

(probability of $\leq T/2$ heads in T coin flips if probability of heads $= \frac{1}{2} + \gamma$)

- AdaBoost's training error is like Chernoff approximation of BBM's