Homework 2

1

given : $E_{agg}(x) = E[\{\frac{1}{M}\sum_{i=1}^{M}(h(x)-f(x))\}^2]$

$f(x) - h_i(x) = E(\varepsilon_i(x)^2)$

$E_{agg} = \frac{1}{M} E_{avg}$ prove

$$E_{agg} = E[\{\frac{1}{M}\sum_{i=1}^{M}(h_i(x)-f(x))\}^2]$$

$$= E[\frac{1}{M^2}\{\sum_{i=1}^{M}(h_i(x)-f(x))\}^2]$$

$$= \frac{1}{M^2}\sum_{i=1}^{M} E[h_i(x)-f(x)]^2$$

$$= \frac{1}{M^2}\sum_{i=1}^{M} E[-(-h_i(x)+f(x))]^2$$

$$= \frac{1}{M^2}\sum_{i=1}^{M} E[(f(x)-h_i(x))^2]$$

$$E_{agg} = \frac{1}{M^2}\sum_{i=1}^{M} E(\varepsilon_i(x)^2)$$

$$E_{agg} = \frac{1}{M}(\frac{1}{M}\sum_{i=1}^{M} E(\varepsilon_i(x)^2)$$

$$\boxed{E_{agg} = \frac{1}{M} E_{avg}}$$

HW2 2

$f(x)$ is convex $\because f(x)=x^2$

x_i are errors of individual models

λ_i are weights assigned to each inequality

$$f\left(\sum_{i=1}^{M} \lambda_i x_i\right) \leq \sum_{i=1}^{M} \lambda_i f(x_i)$$ since $f(x)=x^2$

$$\left(\sum_{i=1}^{M} \lambda_i x_i\right)^2 \leq \sum_{i=1}^{M} \lambda_i x_i^2$$

$$agg = \left(\frac{1}{M} \sum_{i=1}^{M} x_i\right)^2 \qquad avg = \frac{1}{M} \sum_{i=1}^{M} x_i^2$$

$$E[agg] = E\left[\left(\frac{1}{M} \sum_{i=1}^{M} x_i\right)^2\right]$$

$$E[avg] = E\left[\frac{1}{M} \sum_{i=1}^{M} x_i^2\right]$$

$$E\left[\left(\frac{1}{M} \sum_{i=1}^{M} x_i\right)^2\right] \leq E\left[\frac{1}{M} \sum_{i=1}^{M} x_i^2\right] \therefore E[agg] \leq E[avg]$$

3) $\varepsilon_t = \frac{1}{2} - \gamma_t$: Defining the training error after each iteration

the error of final hyp. $H(x)$ at end of T iterations:

$$\varepsilon(H) = \sum_{i=1}^{N} D_T(i) \cdot \mathbb{1}[H(x_i) \neq y_i]$$

$\mathbb{1}[H(x_i) \neq y_i]$ is indicator func. that =s 1 if condition is true + 0 if otherwise

Subbing it into Hoeffding's inequality

$$X_i = D_T(i) \cdot \mathbb{1}[H(x_i) \neq y_i]$$

$$P(|\varepsilon(H) - \mu| > \epsilon) \leq 2\exp\left(-\frac{2N^{\cancel{2}}\epsilon^2}{\cancel{N}}\right)$$

$$P(|\varepsilon(H) - \mu| > \epsilon) \leq 2\exp(-2N\epsilon^2)$$

setting $\epsilon = \sqrt{\frac{\ln(2)}{2N}}$

$$P\left[|\varepsilon(H) - \mu| > \sqrt{\frac{\ln(2)}{2N}}\right] \leq \frac{1}{2}$$

$$\exp(-2N\epsilon^2) = \exp\left(-2N\left(\frac{\ln 2}{2N}\right)\right) = \frac{1}{2}$$

$$P\left[\varepsilon(H) > \mu + \sqrt{\frac{\ln 2}{2N}}\right] < \frac{1}{2}$$

$$P[\varepsilon(H) > \mu] < \frac{1}{2}$$

training error $\varepsilon(H)$ is less than μ w/ prob. atleast $\frac{1}{2}$

overall training error of the hypothesis $H \leq \mu$

$$\mu = \frac{1}{2} - \gamma T$$

$\therefore$ training error bounded by

$$\exp\left(-2\sum_{t=1}^{T}\gamma_t^2\right)$$