ASSIGNMENT - 4

PART I:

1) Given,

$$\mathcal{E}_{avg} = \frac{1}{M}\sum_{i=1}^{m} \epsilon(\mathcal{E}_i(x)^2)$$

$$\epsilon_{agg}(x) = \epsilon\left[\{1/m \sum_{i=1}^{m} \mathcal{E}_i(x)\}^2\right]$$

$$h_{agg}(x) = \epsilon\left[\{1/m \sum_{i=1}^{m} \mathcal{E}_i(x)\}^2\right]$$

Now,

$$\epsilon_{agg}(x) = \epsilon\left[\{1/m \sum_{i=1}^{m} \epsilon_i(x)\}^2\right]$$

$$\left[1/m \sum_{i=1}^{m} \epsilon_i(x)\right]^2 = 1/m^2\left[\{\sum_{i=1}^{m}\epsilon_i(x)\}\{\sum_{j=1}^{m}\epsilon_j(x)\}\right]$$

$$= 1/m^2 \sum_{i=1}^{m} \epsilon_i(x)\left[\epsilon_j(x) + \sum_{i=j}^{M} \epsilon_j(x)\right]$$

$$= 1/m^2\left[\sum_{i=1}^{m} \epsilon_j(x^2) + \sum_{i\neq j} \epsilon_i(x)\epsilon_j(x)\right] \quad -1$$

<u>From the assumption:</u> $\sum_{i=j} \epsilon_i(x)\epsilon_j(x) = 0$

$$\left[1/m^2 \sum_{i=1}^{m} \epsilon_j(x)\right]^2 = 1/m^2 \sum_{j=1}^{m} \epsilon_j(x)^2 \quad -2$$

$$\therefore 1/m^2\left[\sum_{i=1}^{m} \epsilon_i(x)^2\right] + 1/m^2\left[\sum_{i\neq j} \epsilon_i(x)\epsilon_j(x)\right]$$

$$= 1/m^2\left[\sum_{i=1}^{m} \epsilon_i(x)^2\right]$$

2

$$\epsilon_{agg}(n) = \frac{1}{m^2} \epsilon\left[\sum_{i=1}^{m} \epsilon_i(n)^2\right]$$

$$= 1/m^2 \sum_{i=1}^{m} \epsilon\left[\epsilon_i(n)^2\right] = 1/m^2 \sum_{i=1}^{M} \epsilon\left[\epsilon_i(n)^2\right]$$

$\epsilon_{agg}(n) = 1/m\, \epsilon_{avg}(n)$ $\Rightarrow$ Hence, proved

2) <u>From Jensen's Inequality</u>:

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i) \quad -(1)$$

From question (1)

$$\epsilon_{avg} = 1/m \sum_{i=1}^{m} \epsilon(\epsilon_i(n)^2)$$

$$\epsilon_{avg} = \epsilon\left[\left\{1/m \sum_{i=1}^{m} \epsilon_i(n)\right\}^2\right]$$

$$\epsilon\left[\sum_{i=1}^{m} 1/m(\epsilon_i(n))^2\right] \leq \sum_{i=1}^{m} 1/m\, \epsilon(\epsilon_i(n))^2 \quad -(2)$$

Comparing (1) and (2)

$$f = \epsilon, \quad \lambda = \frac{1}{m}, \quad x_i = \epsilon_i(n)^2$$

In eq (2) $\therefore$ $\left[\frac{\epsilon_1(n) + \epsilon_2(n) + \ldots \epsilon_m(x)}{m}\right]^2 \leq \epsilon_1(n)^2 + \ldots + \epsilon_m(n)^2$

$$\therefore \epsilon\left[\left\{1/m \sum_{i=1}^{m} \epsilon_i(n)\right\}^2\right] \leq \epsilon\left[\sum_{i=1}^{m} 1/m(\epsilon_i(n))^2\right]$$

$\therefore$ $\epsilon_{agg} \leq \epsilon_{avg}$ $\Rightarrow$ Hence, proved

3) At step 1 of algorithm

$$D_1(i) = 1/N$$

Using weight update rule,

$$D_2(i) = \frac{1}{N} \frac{e^{-\alpha_1 y_i h_1(x_i)}}{Z_1}$$

$$D_3(i) = \frac{1}{N} \frac{e^{-\alpha_1 y_i h_1(x_i)}}{Z_1} \cdot \frac{e^{-\alpha_2 y_i h_2(x_i)}}{Z_2}$$

$$\therefore D_{T+1}(i) = \frac{1}{N} e^{-\alpha_1 y_i h_1(x_i)} \cdots \frac{e^{-\alpha_T y_i h_T(x_i)}}{Z_T}$$

$$= \frac{1}{N} \frac{e^{-y_i(\sum \alpha_t h_t(x_i))}}{\prod Z_t}$$

$$\therefore \boxed{D_{T+1}(i) = \frac{1}{N} \frac{e^{-y_i f(x_i)}}{\prod Z_t}} \quad — ①$$

Now, by definition of training error,

$$\text{training}(H) = \frac{1}{N} \sum_i \begin{cases} 1 & \text{if } y_i \neq H(x_i) \\ 0 & \text{if } y_i = H(x_i) \end{cases}$$

training error:

$$\therefore (H) = 1/N \sum_i \begin{cases} 1 & \text{if } y_i f(x_i) < 0 \\ 0 & \text{if } y_i f(x_i) > 0 \end{cases}$$

$$(H) \leq 1/N \sum e^{-y_i f(x_i)} = \sum D_{T+1}(i) \prod Z_t \text{ from } ①$$

$$= \prod Z_t .$$

$\therefore$ Training error is bounded by $\prod Z_t$.

Computing Z_t:

$$Z_t = \sum D_t(i)e^{-\alpha_t} + \sum D_t(i)e^{\alpha_t}$$

[for $h(x_i) = y_i$] [for $h(x_i) \neq y_i$]

$$\therefore Z_t = e^{-\alpha_t} \sum D_t(i) + e^{\alpha_t} \sum D_t(i)$$
$$= e^{-\alpha_t}(1-\epsilon_t) + e^{\alpha_t} e_t$$

$$= 2\sqrt{\epsilon_t(1-\epsilon_t)} \quad [\because \alpha_t = \tfrac{1}{2}\ln(1-\epsilon_t/\epsilon_t)]$$

$$= \sqrt{1-4\gamma_t^2} \quad [\because \epsilon_t = \tfrac{1}{2} - \gamma_t]$$
$$\leq e^{-2\gamma^2 t}$$

$\therefore$ Upper bound on training error $= e^{-2\gamma^2 t}$.