Q3)

(a) $L(y:\hat{y}) = e^{-(y\,\hat{y})} = e^{-(y,\,H_t(x_i))}$

$$\boxed{\hat{y} = H_t(x_i)}$$

$$H_t(x) = \sum_{i=1}^{t} \alpha_i h_i(x)$$

$$= \sum_{i=1}^{t-1} \alpha_i h_i(x) + \alpha_t h_t(x)$$

$$\boxed{H_t(x) = H_{t-1}(x) + \alpha_t h_t(x)}$$

$$L(y,\hat{y}) = \frac{1}{n} \sum_{i=1}^{n} e^{-y_i H(x_i)}$$

$$= \frac{1}{n} \sum_{i=1}^{n} e^{-y_i (H_{t-1} + \alpha_t h_t)}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \underbrace{e^{-y_i H_{t-1}}}_{w_i^t} e^{-y_i \alpha_t h_t}$$

$$= \frac{1}{n} \sum_{i=1}^{n} w_i^t e^{-y_i \alpha_t h_t}$$

To minimize loss, equate derivative to '0'.

$$\frac{\partial L(y,\hat{y})}{\partial \alpha_t} = \frac{\partial}{\partial \alpha} \left(\frac{1}{n} \sum_{i=1}^{n} w_i^t e^{-y_i \alpha_t h_t} \right) = 0$$

$$0 \Rightarrow \frac{-1}{n} \sum_{i=1}^{n} w_i^t y_i h_t \, e^{-y_i \alpha_t h_t}$$

$\boxed{e^{x+y} = e^x e^y}$

$$\Rightarrow \frac{-1}{n} \left[\underbrace{\sum_{i: h(x_i) = y_i}^{n}}_{y_i * h_t = 1} w_i^t e^{-\alpha_t} - \underbrace{\sum_{i: h(x_i) \neq y_i}}_{y_i * h_t = -1} w_i^t e^{\alpha t} \right]$$

$$\Rightarrow -\frac{1}{n} \sum_{i: h(x_i) = y_i}^{n} w_i^t e^{-\alpha t} + \frac{1}{n} \sum_{i: h(x_i) \neq y_i}^{n} w_i^t e^{\alpha t}$$

$$\varepsilon_t = \sum_{i: h(x_i) \neq y_i}^{n} w_i^t \qquad 1 - \varepsilon_t = \sum_{i: h(x_i) = y_i}^{n} w_i^t$$

$$\Rightarrow 0 = (1 - \varepsilon_t) e^{-\alpha t} - \varepsilon_t e^{\alpha t}$$

$$\Rightarrow (1 - \varepsilon_t) e^{-\alpha t} = \varepsilon_t e^{\alpha t}$$

$$\Rightarrow \frac{1 - \varepsilon_t}{e^{\alpha t}} = \varepsilon_t e^{\alpha t}$$

$$\Rightarrow \frac{1 - \varepsilon_t}{\varepsilon_t} = e^{2\alpha t}$$

$$\Rightarrow \boxed{\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right)}$$

In this we proved that the α for a given iteration is $\frac{1}{2} \ln \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right)$ and we proved that this α_t minimizes the loss function.

So, an update, i.e. at the next step the empirical risk will indeed be reduced.