Q1

a) $H(x) = \sum_x p(x) \log_2(\frac{1}{p(x)})$

$= \sum_x p(x)(-\log_2 p(x))$

Since $0 \le p(x) \le 1$, $\log_2 p(x) \le 0$

$-\log_2 p(x) \geq 0$

thus $p(x) \cdot (-\log_2 p(x)) \geq 0$,

$\sum_x p(x)(-\log_2 p(x)) \geq 0$.

$H(x) \geq 0$. (non-negative function)

b) $KL(p \| q) = \sum_x p(x) \log_2 \frac{p(x)}{q(x)}$

$-KL(p \| q) = \sum_x p(x) \log \frac{q(x)}{p(x)}$

since $\log_2(\frac{\sum x_i}{n}) \geq \frac{\sum \log_2 x_i}{n}$

by Jensen's Inequality

$\sum_x p(x) \log \frac{q(x)}{p(x)} \le \log_2 \sum_x p(x) \frac{q(x)}{p(x)}$

$\le \log_2 \sum_x q(x)$

$\le \log_2 1$

≤ 0.

so $-KL(p \| q) \le 0$

$KL(p \| q) \geq 0$

c) $I(Y; X) = KL(p(x,y) \| p(x)p(y))$

LHS: $I(Y;X) = H(Y) - H(Y|X)$

$$= \Sigma_y\, p(y) \log_2 \frac{1}{p(y)} - \Sigma\, p(x)\, H(Y \mid X = x)$$

$$= \Sigma_y\, p(y) \log_2 \frac{1}{p(y)} - \Sigma\, p(x,y) \log_2 \frac{p(x)}{p(x,y)}$$

RHS: $KL(p(x,y) \| p(x)p(y)) = \Sigma\, p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$

$$= -\Sigma\, p(x,y) \log_2 \frac{p(x)p(y)}{p(x,y)}$$

$$= -\Sigma\, p(x,y) \log_2 \frac{p(x)}{p(x,y)} - \Sigma\, p(x,y) \log p(y)$$

$$= -\Sigma\, p(x,y) \log_2 \frac{p(x)}{p(x,y)} + \Sigma\, p(x,y) \log \frac{1}{p(y)}$$

LHS = RHS

Q2

$$\bar{h}(x) = \frac{1}{m} \Sigma\, h_i(x)$$

$$L(\bar{h}(x), t) \leq \frac{1}{m} \Sigma\, L(h_i(x), t)$$

LHS $= L(\frac{1}{m} \Sigma\, h(x_i, t)) = L(Z(h(x_i), t))$

RHS $= E(h(x_i), t)$

$L(y,t) = \frac{1}{2}(y-t)^2$ since $p = 2 \geq 1$

thus.. $L(y,t)$ is convex, $L(\bar{h}(x),t) \leq \frac{1}{m}\sum L(h_i(x),t)$

Q3

$$err' = \frac{\sum^N w_i' \mathbb{I}\{h_t(x^i) \neq t^i\}}{\sum^N w_i'}$$

$$= \frac{\sum^E we\text{^}(+\alpha t^i h_t(x)) \overbrace{\mathbb{I}\{h_t(x) \neq t_i\}}^{=1} + \overbrace{\sum^{E^c} We\text{^}(-\alpha t^i h_t(x))}^{=0}}{\sum^E we\text{^}(2\alpha \underbrace{\mathbb{I}\{h_t(x) \neq t_i\}}_{=1} + \sum^{E^c} we\text{^}\underbrace{(2\alpha \mathbb{I}\{h_t(x) \neq t_i\}}_{=1}}$$

$$= \frac{\sum^E W_i e\text{^}(2\alpha \mathbb{I}\{h_t(x) \neq t_i\})}{\sum^E W_i e\text{^}(2\alpha) \quad \sum^{E^c} W_i}$$

$$= \frac{\frac{\sum^E W_i}{\sum W_i}}{\frac{\sum^E W_i}{\sum^N W_i} + \frac{\sum^{E^c} W \cdot e\text{^}(-2\alpha)}{\sum^N W_i}}$$

$$= \frac{\text{error}}{\text{error} + e\text{^}(-2\alpha)(1-\text{error})}$$

$$= \frac{\text{error}}{\text{error} + \frac{\text{error}}{1-\text{error}}(1-\text{error})}$$

$$= \frac{1}{2}$$