CSC 411 HW 2

Ruijie Sun 1003326046

Q1 (a) Prove that the entropy $H(x)$ is non-negative

$$H(X) = \sum_x P(x) \log_2 \frac{1}{P(x)}$$

Case 1. $P(x) = 0$

Since $P(x) = 0$ then $P(x) \log_2 \frac{1}{P(x)} = 0$

Case 2. $0 < P(x) \leq 1$

Since $0 < P(x) \leq 1$, then $\frac{1}{P(x)} \geq 1$

$\Rightarrow \log_2 \frac{1}{P(x)} \geq 0$

$\Rightarrow P(x) \log_2 \frac{1}{P(x)} \geq 0$

Since $0 \leq P(x) \leq 1$, for each case, $P(x) \log_2 \frac{1}{P(x)} \geq 0$

Thus, the summation, $H(x) \geq 0$.

$\square$

(b) Prove that $KL(p||q)$ is non-negative.

$$KL(p||q) = \sum_x p(x) \log_2 \frac{p(x)}{q(x)}$$
$$= E\left(\log_2 \frac{p(x)}{q(x)}\right)$$
$$= E\left(-\log_2 \frac{q(x)}{p(x)}\right)$$

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

$\Rightarrow \frac{q(x)}{p(x)}$ is positive real number. $\Rightarrow \log_2 \frac{q(x)}{p(x)}$ is concave

Since $\log_2 \frac{q(x)}{p(x)}$ is concave $\Rightarrow -\log_2 \frac{q(x)}{p(x)}$ is convex

Thus $KL(p||q) = E\left(-\log_2 \frac{q(x)}{p(x)}\right)$

$$\geq -\log_2 E\left(\frac{q(x)}{p(x)}\right)$$
$$= -\log_2\left(\sum_x p(x) \frac{q(x)}{p(x)}\right)$$
$$= -\log_2\left(\sum_x q(x)\right)$$
$$= -\log_2 1$$
$$= 0$$

Therefore, $KL(p||q) \geq 0$ $\square$

non-negative.

Scanned by CamScanner

(c) $I(Y;X) = H(Y) - H(Y|X)$

$= H(X,Y) - H(X|Y) - H(Y|X)$ (Chain Rule in Slide)

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

$- (-\sum_{x,y} p(x,y)\log_2 p(x|y))$ $\quad (H(Y|X) = -\sum_{x,y} p(x,y)\cdot \log_2 p(y|x)$

$- (-\sum_{x,y} p(x,y)\log_2 p(y|x))$ $\quad)$

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

$+ \sum_{x,y} p(x,y)\log_2 \frac{p(x,y)}{p(x)}$

$= \sum_{x,y} p(x,y)\cdot [\log_2 \frac{p(x,y)}{p(y)} + \log_2 \frac{p(x,y)}{p(x)} - \log_2 p(x,y)]$

$= \sum_{x,y} p(x,y)\cdot \log_2 \left(\frac{(p(x,y))^2}{p(y)\cdot p(x)\cdot p(x,y)} \right)$

$= \sum_{x,y} p(x,y)\cdot \log_2 \frac{p(x,y)}{p(x)\,p(y)}$

$= KL(p(x,y)\,||\,p(x)\,p(y))$

$\square$

Q2 $L(\bar{h}(x), t) = L(\frac{1}{m} \Sigma_{i=1}^{m} h_i(x), t)$

$= L(E(h(x)), t)$

$\leq E(L(h(x), t))$

Since $L(y,t) = \frac{1}{2}(y-t)^2$ is convex

Jensen's Inequality applies.

$= \frac{1}{m} \Sigma_{i=1}^{m} L(h_i(x), t)$

Q3 Define Set $E = \{ i \mid h_t(x^{(i)}) \neq t^{(i)} \}$

$$E^c = \{ i \mid h_t(x^{(i)}) = t^{(i)} \}$$

$$err'_t = \frac{\sum_{i=1}^{N} w_i' \mathbb{I}\{h_t(x^{(i)}) \neq t^{(i)}\}}{\sum_{i=1}^{N} w_i'}$$

$$= \frac{\sum_{i \in E} w_i'}{\sum_{i \in E} w_i' + \sum_{i \in E^c} w_i'} \quad \left(\text{Since } \forall i \in E^c, \mathbb{I}\{h_t(x^{(i)}) \neq t^{(i)}\} = 0 \quad \forall i \in E, \mathbb{I}\{h(x^{(i)}) \neq t^{(i)}\} = 1 \right)$$

$$= \frac{\sum_{i \in E} w_i \cdot \exp(-\alpha_t \cdot t^{(i)} \cdot h_t(x^{(i)}))}{\sum_{i \in E} w_i \cdot \exp(-\alpha_t \cdot t^{(i)} \cdot h_t(x^{(i)})) + \sum_{i \in E^c} w_i \cdot \exp(-\alpha_t \cdot t^{(i)} \cdot h_t(x^{(i)}))}$$

$$= \frac{\sum_{i \in E} w_i \cdot \exp(\alpha_t)}{\sum_{i \in E} w_i \cdot \exp(\alpha_t) + \sum_{i \in E^c} w_i \cdot \exp(-\alpha_t)}$$

$$\left(\text{Since } \forall i \in E, \ t^{(i)} h_t(x^{(i)}) = -1 \quad \forall i \in E^c, \ t^{(i)} h_t(x^{(i)}) = 1 \right)$$

$$= \frac{\sum_{i \in E} w_i}{\sum_{i \in E} w_i + \sum_{i \in E^c} w_i \cdot \exp(-2\alpha_t)}$$

$$= \frac{\sum_{i \in E} w_i}{\sum_{i \in E} w_i + \sum_{i \in E^c} w_i \cdot \exp\left(-\log \frac{1 - err_t}{err_t}\right)}$$

$$= \frac{\sum_{i \in E} w_i}{\sum_{i \in E} w_i + \sum_{i \in E^c} w_i \cdot \frac{err_t}{1 - err_t}}$$

where $err_t = \frac{\sum_{i \in E} w_i}{\sum_{i=1}^{N} w_i}$

$$= \frac{\sum_{i \in E} W_i}{\sum_{i \in E} W_i + \sum_{i \in E^c} W_i \cdot \frac{\frac{\sum_{i \in E} W_i}{\sum_{i=1}^{N} W_i}}{1 - \frac{\sum_{i \in E} W_i}{\sum_{i=1}^{N} W_i}}}$$

$$= \frac{\sum_{i \in E} W_i}{\sum_{i \in E} W_i + \sum_{i \in E^c} W_i \cdot \frac{\sum_{i \in E} W_i}{\sum_{i \in E^c} W_i}}$$

$$= \frac{\sum_{i \in E} W_i}{\sum_{i \in E} W_i + \sum_{i \in E} W_i}$$

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

The intepretation of this is that we need a new learner since the error is not good. The error should be less than $\frac{1}{2}$.