Homework 1

September 12, 2018

1 Homework 1A

We define:

$$Gain(X,Y) = H(Y) - H(Y|X)$$

$$H(Y) = -\sum_j P(I_j)log(P(I_j))$$

$$H(Y|X) = \sum_i H(Y|x_i)P(x_i)$$

Then the theorem can be formally defined as:

$$Gain(X,Y) = H(Y) - H(Y|X) >= 0$$

Theorem 1 $Gain(X,Y) = H(Y) - H(Y|X) >= 0$

Proof 1 *From the defination, we gain:*

$$Gain(X,Y) = H(Y) - H(Y|X)$$

$$= -\sum_j P(I_j)log(P(I_j)) - \sum_i H(Y|x_i)P(x_i)$$

$$= -\sum_j P(I_j)log(P(I_j)) + \sum_i (\sum_j P(I_j|x_i)log(P(I_j|x_i))))P(x_i)$$

$$= -\sum_j log(P(I_j)) \sum_i P(I_j|x_i)P(x_i) + \sum_i P(x_i) \sum_j P(I_j|x_i)log(P(I_j|x_i))$$

$$= \sum_{i,j} P(I_j|x_i)P(x_i)log(\frac{P(I_j|x_i)}{P(I_j)})$$

$$= \sum_{i,j} P(I_jx_i)log(\frac{P(I_jx_i)}{P(I_j)P(x_i)})$$

$$= \sum_{i,j} P(I_jx_i) - log(\frac{P(I_j)P(x_i)}{P(I_jx_i)})$$

-log is a convex function, then we can use Jensen unequation.
From Jensen unequation:

$$\sum_{i,j} P(I_jx_i) - log(\frac{P(I_j)P(x_i)}{P(I_jx_i)})$$

$$>= \sum_{i,j} -log(P(I_j x_i) \frac{P(I_j)P(x_i)}{P(I_j x_i)})$$

$$= \sum_{i,j} -log(P(I_j)P(x_i))$$

$$= -log(1) = 0$$

Therefore:

$$Gain(X, Y) = H(Y) - H(Y|X) >= 0$$

Also, we may import KL divergence, and prove that KL divergence is non-negative;
ref: https://blog.csdn.net/MathThinker/article/details/48375523