AI-homework1

晏悦 2017K8009918013

August 30 2019

证明 Gain(X,Y) = H(Y) - H(Y,X)>0.

证明:

由于先获取 X 的信息熵，再获取基于 X 的 Y 的信息熵，等价于先获取 Y 的信息熵，再获取 X 的信息熵，所以:

$$H(X,Y) = H(X) + H(X|Y) = H(Y) + H(Y|X)$$

Gain(X,Y) = H(Y) - H(Y|X) = $\Sigma_y p(y) * log(p(y)) + \Sigma_x p(x) \Sigma_y p(y|x) log(p(y|x))$

=H(X) - H(X|Y) = $\Sigma_x p(x) * log(p(x)) + \Sigma_y p(y) \Sigma_x p(x|y) log(p(x|y))$

=H(X) + H(Y) - H(X,Y) = $\Sigma_x \Sigma_y p(x,y) * (-log(p(x) - log(p(y) + log(p(x,y)))$

=$\Sigma_x \Sigma_y p(x,y) * log(\frac{p(x,y)}{p(x)p(y)})$

=D(P(X,Y)||P(X)P(Y))

其中 D 是 KL 散度公式，那么只用证明 KL 散度大于零即可

D(P||Q) = $\Sigma_x p(x) log(\frac{p(x)}{q(x)})$

>=-log($\Sigma_x p(x) * \frac{q(x)}{p(x)}$)

(这一步由 Jensen 不等式得到)

=0

所以证明完毕，信息熵增量大于等于 0，其中为 0 时，当且仅当 X，Y 独立。