第五次作业

自硕21 崔昊菲 202121097 6

1、解:

(1) 证明:

$$E_{BAG} = E_x\{(y_{BAG}(x) - d(x))^2\} = E_x\{(\frac{1}{M}\sum_{i=1}^{M} y_m(x) - d(x))^2\}$$

$$= E_x\{\frac{1}{M^2}(\sum_{i=1}^{M} y_m(x) - d(x))^2\}$$

$$= \frac{1}{M^2} E_x\{(\sum_{m=1}^{M} \varepsilon_m(x))^2\}$$

$$= \frac{1}{M^2} E_x\{\sum_{m=1}^{M} \varepsilon_m(x)^2 + \sum_{m\neq l} \varepsilon_m(x)\varepsilon_l(x)\}$$

$$= \frac{1}{M^2}\left(\sum_{m=1}^{M} E_x\{\varepsilon_m(x)^2\} + \sum_{m\neq l} E_x\{\varepsilon_m(x)\varepsilon_l(x)\}\right)$$

$$= \frac{1}{M^2}\sum_{m=1}^{M} E_x\{\varepsilon_m(x)^2\} = \frac{1}{M} E_{AV}$$ 证毕

(2) 由于 $y = x^2$ 是下凸函数，因此对于 $\forall x$，恒有

$$\left(\frac{1}{M}\sum_{m=1}^{M} \varepsilon_m(x)\right)^2 \leq \frac{1}{M}\sum_{m=1}^{M} \varepsilon_m(x)^2$$

因此有 $$E_x\{(\frac{1}{M}\sum_{m=1}^{M}\varepsilon_m(x))^2\} \leq E_x\{\frac{1}{M}\sum_{m=1}^{M}\varepsilon_m(x)^2\}$$

即 $$E_{BAg} \leq \frac{1}{M} E_x\{\sum_{m=1}^{M}\varepsilon_m(x)^2\} = E_{AV}$$

证毕

2、解: 有8个正例, 9个反例, 故总体的信息熵为

$$S_0 = -\frac{8}{17}\log_2\left(\frac{8}{17}\right) - \frac{9}{17}\log_2\left(\frac{9}{17}\right) \approx 0.9975$$

若 Weight 作为根节点, 则

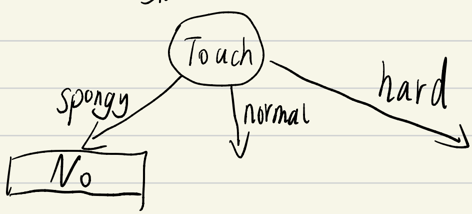
$S(light) \approx 0.9183$ $S(average) \approx 0.7219$ $S(heavy) \approx 1$

得 $Gain(Weight) = S_0 - \frac{5}{17}S(average) - \frac{6}{17}S(heavy) - \frac{6}{17}S(light) \approx 0.1081$

同理得

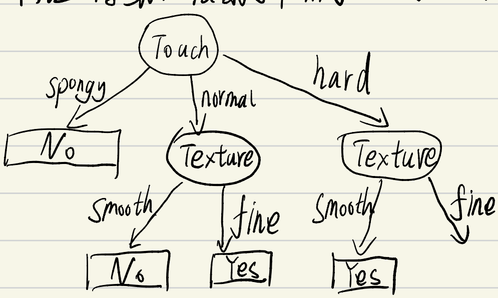
$Gain(Size) \approx 0.1427$ $Gain(Touch) \approx 0.3806$ $Gain(Texture) \approx 0.0060$

故根节点为 Touch. 第一层为



当 Touch = normal 时, $S(normal) \approx 0.7219$

同理得第二层的子节点为 Texture, 即



再之后 $Gain(Weight) = Gain(Size)$

故选择任意一个当第三层结点都可

得