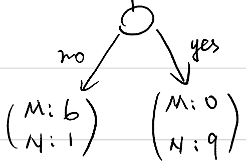
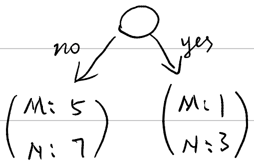
1. (a) $Gini(t) = 1 - \sum_{i=1}^{c} p(i|t)^2$

1 step: Attribute (Lay Eggs)



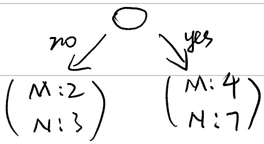
$$I = \frac{7}{16} \cdot Gini(no) + \frac{9}{16} \cdot Gini(yes)$$
$$= \frac{7}{16}\left(1-\left(\frac{6}{7}\right)^2-\left(\frac{1}{7}\right)^2\right) + \frac{9}{16}(1-1-0)$$
$$= \frac{3}{28}$$

Attribute: (Can Fly)



$$I = \frac{12}{16} \cdot \left(1-\left(\frac{5}{12}\right)^2-\left(\frac{7}{12}\right)^2\right) + \frac{4}{16}\left(1-\left(\frac{1}{4}\right)^2-\left(\frac{3}{4}\right)^2\right)$$
$$= \frac{13}{30}$$

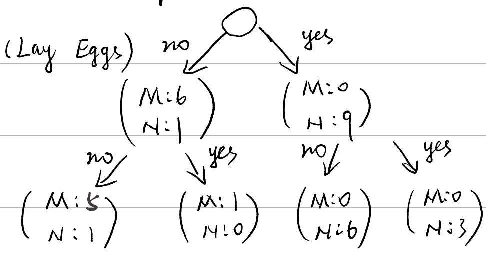
Attribute: (Have Legs)



$$I = \frac{5}{16} \cdot \left(1-\left(\frac{2}{5}\right)^2-\left(\frac{3}{5}\right)^2\right) + \frac{11}{16}\left(1-\left(\frac{4}{11}\right)^2-\left(\frac{7}{11}\right)^2\right)$$
$$= \frac{3}{20} + \frac{7}{22}$$

First $\boxed{\text{(Lay Eggs)}}$ attribute, $\boxed{I = \frac{3}{28} \approx 0.1071}$

(b) Step 2:

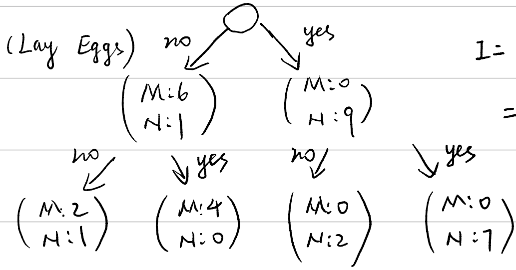


Attribute: (Can Fly)

$$I = \frac{6}{16} \cdot \left(1-\left(\frac{5}{6}\right)^2-\left(\frac{1}{6}\right)^2\right) + 0$$
$$= \frac{5}{48}$$

日期: /

Attribute: (Have Legs)



$$I = \frac{3}{16}\left(1-\left(\tfrac{2}{3}\right)^2-\left(\tfrac{1}{3}\right)^2\right)+0$$

$$= \frac{1}{12}$$

2nd step: $\boxed{\text{(Have Legs)}}$ attribute, $\boxed{I = \frac{1}{12} \approx 0.083}$

(c) $$g(l,c,h) = \log \frac{P(\text{Mammal} \mid l,c,h)}{P(\text{Non-Mammal} \mid l,c,h)}$$

$$= \log \frac{P(l|M)\cdot P(c|M)\cdot P(h|M)\cdot P(M)}{P(l|N)\cdot P(c|N)\cdot P(h|N)\cdot P(N)}$$

$$z_0 = \log \frac{P(M)}{P(N)} = \log \frac{6}{10}$$

$$g_1(l) = \log \frac{P(l|M)}{P(l|N)} = \begin{cases} \log \dfrac{(0+1)/(6+2)}{(9+1)/(10+2)} & l=1 \\ \log \left(\dfrac{(6+1)/(6+2)}{(1+1)/(10+2)}\right) & l=-1 \end{cases}$$

$$g_2(c) = \log \frac{P(c|M)}{P(c|N)} = \begin{cases} \log \dfrac{(1+1)/(6+2)}{(3+1)/(10+2)} & c=1 \\ \log \dfrac{(5+1)/(6+2)}{(7+1)/(10+2)} & c=-1 \end{cases}$$

$$g_3(h) = \log \frac{P(h|M)}{P(h|N)} = \begin{cases} \log \dfrac{(4+1)/(6+2)}{(7+1)/(10+2)} & c=1 \\ \log \dfrac{(2+1)/(6+2)}{(3+1)/(10+2)} & c=-1 \end{cases}$$

$z_0 = \log \frac{3}{5} \approx -0.5108$, $g_1(1) = \log \frac{3}{20} \approx -0.8239$, $g_1(-1) = \log \frac{21}{4} \approx 1.6582$, $g_2(1) = \log \frac{3}{4} \approx -0.1249$, $g_2(-1) = \log \frac{9}{8} \approx 0.0512$

$$\boxed{g_3(1) = \log\frac{15}{16} \approx -0.0280, \quad g_3(-1) = \log\frac{9}{8} \approx 0.0512}$$

2. (a)

		truth P	truth N
predicted	P	4 (TP)	2 (FP)
	N	1 (FN)	7 (TN)
		AP = 5	AN = 9

positive spam y, negative legitimate x, $\theta = 20$

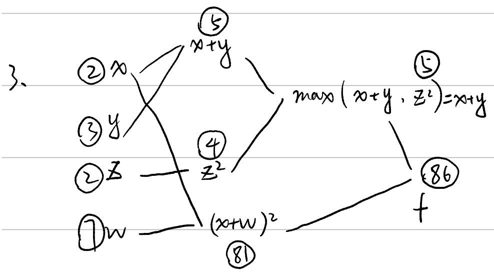
$$\boxed{\begin{aligned} TPR &= \frac{TP}{AP} = \frac{4}{5} = 0.8 \\ FPR &= \frac{FP}{AN} = \frac{2}{9} = 0.2222 \end{aligned}}$$

(b)
$$AUC = \frac{\sum_{x \in AN} \sum_{y \in AP} 1_{f(y) > f(x)}}{\# AP \cdot \# AN}$$

$$= \frac{9+9+9+8+4}{5 \times 9}$$

$$\boxed{= \frac{39}{45} = 0.8667}$$

3.



$$\frac{\partial f}{\partial(x+y)} = \frac{\partial f}{\partial(x+y)} \cdot \frac{\partial(x+y)}{\partial(x+y)} = 1$$

$$\frac{\partial f}{\partial z^2} = \frac{\partial f}{\partial(x+y)} \cdot \frac{\partial(x+y)}{\partial(z^2)} = 0$$

$$\frac{\partial f}{\partial(x+w)^2} = 1$$

日期:

$$\frac{\partial(x+y)}{\partial x}=1,\quad \frac{\partial(x+y)}{\partial y}=1,\quad \frac{\partial z^2}{\partial z}=2z=4$$

$$\frac{\partial(x+w)^2}{\partial x}=2(x+w)=18,\quad \frac{\partial(x+w)^2}{\partial w}=18$$

$$\frac{\partial f}{\partial x}=1\cdot1+1\cdot18=19,\quad \frac{\partial f}{\partial y}=1\cdot1=1$$

$$\frac{\partial f}{\partial z}=0,\quad \frac{\partial f}{\partial w}=1\cdot18=18,\quad \boxed{\nabla f=(19,1,0,18)}$$

4. (a) consider a3,

yes / no

$\begin{pmatrix} d2\ \text{yes}: 9 \\ d2\ \text{no}: 4 \end{pmatrix}$ $\begin{pmatrix} d2\ \text{yes}: 1 \\ d2\ \text{no}: 6 \end{pmatrix}$

$$\text{Gini}=\frac{13}{20}\left(1-\left(\frac{9}{13}\right)^2-\left(\frac{4}{13}\right)^2\right)+\frac{7}{20}\left(1-\left(\frac{1}{7}\right)^2-\left(\frac{6}{7}\right)^2\right)$$

$$=\frac{33}{91}=0.3626$$

consider a5 :

yes / no

$\begin{pmatrix} d2\ \text{yes}: 1 \\ d2\ \text{no}: 3 \end{pmatrix}$ $\begin{pmatrix} d2\ \text{yes}: 9 \\ d2\ \text{no}: 7 \end{pmatrix}$

$$\text{Gini}=\frac{4}{20}\left(1-\left(\frac{1}{4}\right)^2-\left(\frac{3}{4}\right)^2\right)+\frac{16}{20}\left(1-\left(\frac{9}{16}\right)^2-\left(\frac{7}{16}\right)^2\right)$$

$$=\frac{15}{32}=0.4687$$

$\boxed{a3}$ better.

(b) a1 → d1

$$\text{AUC}=\frac{\sum_{x\in AN}\sum_{y\in AP} 1_{f(y)>f(x)}}{\#AP\cdot\#AN}=\frac{1+2\times4+3\times2+5+8}{9\times11}=\boxed{\frac{28}{99}=0.2828}$$

日期: /

(c) $\overset{a_1}{\theta_1} = 37.95, \; a_2, a_3 \rightarrow d_2$

$$\frac{p(d_2 = yes \mid a_1 > \theta_1, a_2 = no, a_3 = no)}{p(d_2 = no \mid a_1 > \theta_1, a_2 = no, a_3 = no)}$$

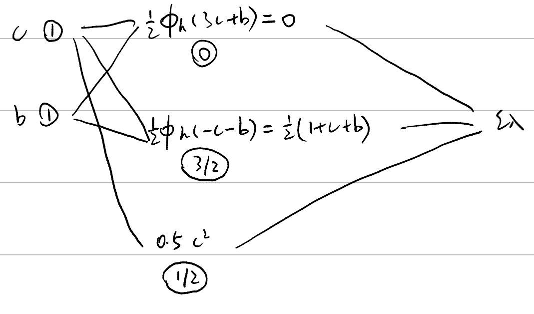
$$\approx \frac{p(d_2 = yes) \cdot p(a_1 > \theta_1 \mid yes) \cdot p(a_2 = no \mid yes) \cdot p(a_3 = no \mid yes)}{p(d_2 = no) \cdot p(a_1 > \theta_1 \mid no) \cdot p(a_2 = no \mid no) \cdot p(a_3 = no \mid no)}$$

$$= \frac{10}{10} \cdot \frac{9/10}{1/10} \cdot \frac{10/10}{10/10} \cdot \frac{1/10}{6/10}$$

$$= \frac{3}{2} > 1$$

$$\arg\max_{d_2} \sim = \boxed{yes}$$

5. (a) $\mathcal{E}_\lambda(c=1, b=1) = \frac{1}{2}\phi_h(3c+b) + \frac{1}{2}\phi_h(-c-b) + 0.5c^2$



$$\frac{\partial \mathcal{E}_\lambda}{\partial c} = \frac{\partial \frac{1}{2}(1+c+b)}{\partial c} + \frac{\partial(0.5c^2)}{\partial c} = \frac{1}{2} + c = 1.5$$

$$\frac{\partial \mathcal{E}_\lambda}{\partial b} = \frac{\partial \frac{1}{2}(1+c+b)}{\partial b} + 0 = 0.5$$

$$\boxed{\nabla \mathcal{E}_\lambda = \left(\frac{\partial \mathcal{E}_\lambda}{\partial c}, \frac{\partial \mathcal{E}_\lambda}{\partial b}\right) = (1.5, 0.5)}$$

日期: /

(b) $\mathcal{E}_\lambda = \frac{1}{2}\max\{0, 1-3c-b\} + \frac{1}{2}\max\{0, 1+c+b\} + \frac{1}{2}c^2$.

① if $1-3c-b>0$, $1+c+b>0$, $\mathcal{E}_\lambda = 1-c+\frac{1}{2}c^2$

$\nabla\mathcal{E}_\lambda = 0 \Rightarrow c=1$, $b<-2$, $b>-2$, contradictory.

② if $1-3c-b>0$, $1+c+b<0$, $\mathcal{E}_\lambda = \frac{1}{2}(1-3c-b) + \frac{1}{2}c^2$

$\nabla\mathcal{E}_\lambda \neq 0$

③ if $1-3c-b<0$, $1+c+b>0$, $\mathcal{E}_\lambda = \frac{1}{2}(1+c+b) + \frac{1}{2}c^2$

$\nabla\mathcal{E}_\lambda \neq 0$

④ if $1-3c-b<0$, $1+c+b<0$, $\mathcal{E}_\lambda = \frac{1}{2}c^2$.

$\nabla\mathcal{E}_\lambda = 0 \Rightarrow c=0$, $b\geq 1$, $b\leq -1$, contradiction.

Since $\mathcal{E}_\lambda$ is convex, so that if $\mathcal{E}_\lambda$ is differentiable,

$\nabla\mathcal{E}_\lambda = 0$.

if $\mathcal{E}_\lambda$ is not differentiable,

⑤ if $1-3c-b=0$, $1+c+b>0$, $\nabla\mathcal{E}_\lambda \neq 0$

⑥ if $1-3c-b=0$, $1+c+b<0$. then $c=0$, contradictory

⑦ if $1-3c-b=0$, $1+c+b=0$,

then $c=1$, $b=-2$, $\mathcal{E}_\lambda = 0.5$

$c^* = 1$

日期:

6. Fix $0<q<1$, $\gamma>0$,

$\rho^2 \leq (x-\rho)^2 + \gamma \cdot |x|^q$, $\forall x$, holds

find range of ρ.

assume that $f(x) \triangleq x^2 - 2x\rho + \gamma |x|^q \geq 0$

① if $\forall x \geq 0$, $f(x) = x^2 - 2x\rho + \gamma \cdot x^q$, $\frac{df}{dx} = 2(x^* - \rho) + \gamma q \cdot x^{*q-1} = 0$

$\frac{d^2 f}{dx^2} = 2 + \gamma q(q-1) x^{q-2}$, we know that f is concave $\to$ convex.

$\triangle$ we consider if there $\exists\, x^* \geq 0$, s.t. $f'(x^*) = 0$.

so that if $f(x^*) \geq 0$, the $f(x) \geq 0$, $\forall x \geq 0$.

$\Rightarrow \rho = x^* + \frac{1}{2}\gamma q x^{*q-1}$

$\Rightarrow f(x^*) = -x^{*2} + \gamma(1-q) x^{*q} \geq 0$, $x^* \geq 0$

$\Rightarrow 0 \leq x^* \leq \left(\gamma(1-q)\right)^{\frac{1}{2-q}}$

$\left.\frac{d\rho}{dx^*}\right|_{x_\rho} = 1 + \frac{1}{2}\gamma q(q-1) x_\rho^{q-2} = 0$, we know $\rho(x^*)$ is convex.

$\Rightarrow x_\rho = \left(\frac{2}{\gamma q(1-q)}\right)^{\frac{1}{q-2}} = \left(\frac{\gamma q(1-q)}{2}\right)^{\frac{1}{2-q}} < \left(\gamma(1-q)\right)^{\frac{1}{2-q}}$

$\Rightarrow \rho \geq \rho(x_\rho) = \left(\frac{\gamma q(1-q)}{2}\right)^{\frac{1}{2-q}} + \frac{1}{2}\gamma q \cdot \left(\frac{\gamma q(1-q)}{2}\right)^{\frac{q-1}{2-q}}$

② if $\forall x \leq 0$, $f(x) = x^2 - 2x\rho + \gamma(-x)^q$,

$\frac{df}{dx} = 2x - 2\rho - \gamma q(-x)^{q-1}$, $\frac{d^2 f}{dx^2} = 2 + \gamma q(q-1)(-x)^{q-2}$,

when $x \leq 0$, f is convex $\to$ concave

$\triangle$ we consider $\exists\, x^* \leq 0$, s.t. $f'(x^*) = 0$.

$\left.\frac{\partial f}{\partial x}\right|_{x^*} = 2x^* - 2\rho - \gamma q(-x^*)^{q-1} = 0 \Rightarrow \rho = x^* - \frac{1}{2}\gamma q(-x^*)^{q-1}$

substitute ρ into f. $f(x^*) = -x^{*2} + (1-q)\gamma(-x^*)^q \geq 0$

$\Rightarrow x^* \geq -\left((1-q)\gamma\right)^{\frac{1}{2-q}}$

日期: /

$$\frac{d\varsigma}{dx^*}\Big|_{x_\varsigma} = 1 + \frac{1}{2}\gamma q(q-1)(-x_\varsigma)^{q-2} = 0$$

$$\Rightarrow x_\varsigma = -\left(\frac{2}{\gamma q(1-q)}\right)^{\frac{1}{q-2}} = -\left(\frac{\gamma q(1-q)}{2}\right)^{\frac{1}{2-q}} \geq -\left((1-q)\gamma\right)^{\frac{1}{2-q}}$$

$$\Rightarrow \varsigma \leq -\left(\frac{\gamma q(1-q)}{2}\right)^{\frac{1}{2-q}} - \frac{1}{2}\gamma q\left(\frac{\gamma q(1-q)}{2}\right)^{\frac{q-1}{2-q}}$$

③ if $x \geq 0$, we assume that $\nexists\ x^*$, s.t $f'(x^*) = 0$

so that $f(x)$ is monotonically over $(0, +\infty)$.

we need $\begin{cases} \lim\limits_{x\to+\infty} f(x) \geq 0 \\ \frac{df}{dx} \leq 0 \end{cases}$ or $\frac{df}{dx} \geq 0$

$$\Rightarrow \begin{cases} \lim\limits_{x\to+\infty} f(x) \geq 0 \\ 2x - 2\varsigma + \gamma q \cdot x^{q-1} \leq 0 \end{cases} \text{ or } \quad 2x - 2\varsigma + \gamma q \cdot x^{q-1} \geq 0 \quad \text{for } \forall x \geq 0$$

$$\Rightarrow \min_{x\geq 0}\left(x + \frac{1}{2}\gamma q x^{q-1}\right) = 0$$

$$\Rightarrow \varsigma \leq 0$$

④ if $x \leq 0$, we assume that $\nexists\ x^*$, s.t. $f'(x^*) = 0$

so that $f(x)$ monotonic over $(-\infty, 0)$

we have $\begin{cases} \lim\limits_{x\to-\infty} f(x) \geq 0 \\ \frac{df}{dx} \geq 0 \end{cases}$ or $\frac{df}{dx} \leq 0$

$$\Rightarrow 2x - 2\varsigma - \gamma q(-x)^{q-1} \geq 0 \quad \text{or} \quad 2x - 2\varsigma - \gamma q(-x)^{q-1} \leq 0, \text{ for } \forall x \leq 0$$

$$\Rightarrow \varsigma \geq x - \frac{1}{2}\gamma q(-x)^{q-1}, \quad \forall x \leq 0$$

$$\max_{x\leq 0}\left(x - \frac{1}{2}\gamma q(-x)^{q-1}\right) = -\left(\frac{2}{\gamma q(1-q)}\right)^{\frac{1}{q-2}} - \frac{1}{2}\gamma q\left(\frac{2}{\gamma q(1-q)}\right)^{\frac{q-1}{q-2}}$$

$$\Rightarrow \varsigma \geq -\left(\frac{\gamma q(1-q)}{2}\right)^{\frac{1}{2-q}} - \frac{1}{2}\gamma q\left(\frac{2}{\gamma q(1-q)}\right)^{\frac{1-q}{2-q}}$$

日期: /

hence,

$$\boxed{\begin{aligned} &\rho \geq \left(\frac{\gamma q(1-q)}{2}\right)^{\frac{1}{2-q}} + \frac{1}{2}\gamma q \cdot \left(\frac{\gamma q(1-q)}{2}\right)^{\frac{q-1}{2-q}} \\ \text{or}\quad &\rho \leq -\left(\frac{\gamma q(1-q)}{2}\right)^{\frac{1}{2-q}} - \frac{1}{2}\gamma q\left(\frac{\gamma q(1-q)}{2}\right)^{\frac{q-1}{2-q}} \end{aligned}}$$