Assignment 2 24133101 9 Wu Yifan

1. a. For Lay Eggs:

If lay Eggs is Yes: mammals: 0, non-mammals: 9; is No: mammals: 6, non-mammals: 1

$$Gini(Lay\ Eggs = Yes) = 1 - (\frac{0}{9})^2 - (\frac{9}{9})^2 = 0$$

$$Gini(Lay\ Eggs = No) = 1 - (\frac{6}{7})^2 - (\frac{1}{7})^2 = \frac{12}{49}$$

$$\text{Weighed Gini Index}: \frac{9}{16} \cdot 0 + \frac{7}{16} \cdot \frac{12}{49} \approx \boxed{0.1071}$$

For can Fly:

If can Fly is Yes: mammals: 1, non-mammals: 3; is No: mammals: 5, non-mammals: 7

$$Gini(can\ Fly = Yes) = 1 - (\frac{1}{4})^2 - (\frac{3}{4})^2 = \frac{6}{16}$$

$$Gini(can\ Fly = No) = 1 - (\frac{5}{12})^2 - (\frac{7}{12})^2 = \frac{70}{144}$$

$$\text{Weighed Gini Index}: \frac{4}{16} \cdot \frac{6}{16} + \frac{12}{16} \cdot \frac{70}{144} \approx \boxed{0.3646}$$

For Have legs:

If Have legs is Yes: mammals: 4, non-mammals: 7; is No: mammals: 2, non-mammals: 3

$$Gini(Have\ legs = Yes) = 1 - (\frac{4}{11})^2 - (\frac{7}{11})^2 = \frac{56}{121}$$

$$Gini(Have\ legs = No) = 1 - (\frac{2}{5})^2 - (\frac{3}{5})^2 = \frac{12}{25}$$

$$\text{Weighed Gini Index}: \frac{56}{121} \cdot \frac{11}{16} + \frac{12}{25} \cdot \frac{5}{16} = \boxed{0.1500}$$

Hence, the attribute "Lay Eggs" should use first to build a decision tree

b. Consider lay eggs = No

For can Fly: If can Fly is Yes: mammals: 1, non-mammals: 0; is No: mammals: 5, non-mammals: 1

$$Gini(can\ Fly = Yes) = 1 - 0 - 1^2 = 0$$

$$Gini(can\ Fly = No) = 1 - (\frac{5}{6})^2 - (\frac{1}{6})^2 = \frac{10}{36}$$

$$\text{Weighed Gini Index} = (\frac{1}{7}) \cdot 0 + (\frac{6}{7}) \cdot \frac{10}{36} \approx \boxed{0.2381}$$

For Have legs: If Have legs is Yes: mammals: 4, non-mammals: 0; is No: mammals: 2, non-mammals: 1

$$Gini(Have\ legs = Yes) = 1 - 1^2 - 0^2 = 0$$

$$Gini(Have\ legs = No) = 1 - (\frac{2}{3})^2 - (\frac{1}{3})^2 = \frac{4}{9}$$

$$\text{Weighed Gini Index} = (\frac{4}{7}) \cdot 0 + (\frac{3}{7}) \cdot \frac{4}{9} = \boxed{0.1905}$$

Hence, the attribute "Have legs" should use second to build decision tree c.

For lay Eggs:

$$P(\text{lay Eggs}=1 \mid \text{mammals}) = \frac{0+1}{6+2} = \frac{1}{8}$$

$$P(\text{lay Eggs}=-1 \mid \text{mammals}) = \frac{6+1}{6+2} = \frac{7}{8}$$

$$P(\text{lay Eggs}=1 \mid \text{non-mammals}) = \frac{9+1}{10+2} = \frac{10}{12}$$

$$P(\text{lay Eggs}=-1 \mid \text{non-mammals}) = \frac{1+1}{10+2} = \frac{2}{12}$$

$$\boxed{\begin{aligned} g_1(1) &= \log\left(\frac{1}{8} \times \frac{12}{10}\right) = \log\left(\frac{12}{80}\right) = \log\left(\frac{3}{20}\right) = -0.8239 \\ g_1(-1) &= \log\left(\frac{7}{8} \times \frac{12}{2}\right) = \log\left(\frac{42}{8}\right) = \log\left(\frac{21}{2}\right) = 1.0212 \end{aligned}}$$

For can Fly:

$$P(\text{can Fly}=1 \mid \text{mammals}) = \frac{1+1}{6+2} = \frac{2}{8}$$

$$P(\text{can Fly}=-1 \mid \text{mammals}) = \frac{5+1}{6+2} = \frac{6}{8}$$

$$P(\text{can Fly}=1 \mid \text{non-mammals}) = \frac{3+1}{10+2} = \frac{4}{12}$$

$$P(\text{can Fly}=-1 \mid \text{non-mammals}) = \frac{7+1}{10+2} = \frac{8}{12}$$

$$\boxed{\begin{aligned} g_2(1) &= \log\left(\frac{2}{8} \times \frac{12}{4}\right) = \log\left(\frac{12}{16}\right) = \log\left(\frac{3}{4}\right) \approx -0.1249 \\ g_2(-1) &= \log\left(\frac{6}{8} \times \frac{12}{8}\right) = \log\left(\frac{72}{64}\right) = \log\left(\frac{9}{8}\right) = 0.0512 \end{aligned}}$$

For have leg

$$P(\text{have leg}=1 \mid \text{mammals}) = \frac{4+1}{6+2} = \frac{5}{8}$$

$$P(\text{have leg}=-1 \mid \text{mammals}) = \frac{2+1}{6+2} = \frac{3}{8}$$

$$P(\text{have leg}=1 \mid \text{non-mammals}) = \frac{7+1}{10+2} = \frac{8}{12}$$

$$P(\text{have leg}=-1 \mid \text{non-mammals}) = \frac{3+1}{10+2} = \frac{4}{12}$$

$$\boxed{\begin{aligned} g_3(1) &= \log\left(\frac{5}{8} \times \frac{12}{8}\right) = \log\left(\frac{60}{64}\right) = \log\left(\frac{15}{16}\right) \approx -0.0280 \\ g_3(-1) &= \log\left(\frac{3}{8} \times \frac{12}{4}\right) = \log\left(\frac{36}{32}\right) = \log\left(\frac{9}{8}\right) = 0.0512 \end{aligned}}$$

$$\boxed{\alpha = \log\left(\frac{6+1}{16+2} \times \frac{16+2}{10+1}\right) = \log\left(\frac{7}{18} \times \frac{18}{11}\right) = \log\left(\frac{7}{11}\right) = -0.1963}$$

So the Naive Bayes classifier is

$$g(l,c,h) = -0.1963 + g_1(l) + g_2(c) + g_3(h)$$

2.

(a) Alice label: Spam: 1, 2, 3, 4, 5, 6

Legit: 7, 8, 9, 10, 11, 12, 13, 14

Bob True label: Spam: 1, 2, 3, 5, 10

Legit: 4, 6, 7, 8, 9, 11, 12, 13, 14

TP = 1, 2, 3, 5 FN = 10

FP = 4, 6 TN = 7, 8, 9, 11, 12, 13, 14

$$\boxed{TPR = \frac{TP}{TP+FN} = \frac{4}{4+1} = \frac{4}{5} = 0.8000 \quad FPR = \frac{FP}{FP+TN} = \frac{2}{2+7} = \frac{2}{9} = 0.2222}$$

(b)
$$AUC = \frac{\sum_{x\in AN}\sum_{y\in AP} 1_{f(y)>f(x)}}{\#AP \times \#AN} = \frac{C_3^1 \cdot C_9^1 + 8 + 4}{5\times 9} = \frac{39}{45} = \boxed{0.8667}$$

3. In $(x=2, y=3, z=2, w=7)$. Compute the gradient below:

$$\frac{\partial f}{\partial x} = \frac{\partial \max(x+y, z^2)}{\partial x} + \frac{\partial (x+w)^2}{\partial x} = 1 + 18 = 19$$

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

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

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

$\boxed{\text{The gradient in back propagation is } (19, 1, 0, 18)}$

4.

a. If a_3 is Yes: d_2_yes: 9 is No: d_2-yes: 1

d_2-no: 4 d_2-no: 6

$$Gini(a_3 = Yes) = 1 - \left(\frac{9}{13}\right)^2 - \left(\frac{4}{13}\right)^2 = \frac{72}{169}$$

$$Gini(a_3 = No) = 1 - \left(\frac{1}{7}\right)^2 - \left(\frac{6}{7}\right)^2 = \frac{12}{49}$$

$$\text{Weighted Gini Index} = \frac{72}{169}\cdot\frac{13}{20} + \frac{12}{49}\cdot\frac{7}{20} \approx \boxed{0.3626}$$

If a_5 is Yes: d_2-yes: 1 is No: d_2-yes: 9
d_2-no: 3 d_2-no: 7

$$Gini(a_5 = Yes) = 1 - (\frac{1}{4})^2 - (\frac{3}{4})^2 = \frac{6}{16}$$

$$Gini(a_5 = No) = 1 - (\frac{9}{16})^2 - (\frac{7}{16})^2 = \frac{126}{256}$$

$$\text{Weighted Gini Index} = \frac{6}{16} \cdot \frac{4}{20} + \frac{126}{256} \cdot \frac{16}{20} = \boxed{0.4689}$$

Hence, a_3 should use first

b. The sample d_1= yes : 9 (37.4, 37.6, 37.7, 37.7, 37.7, 37.9, 37.9, 38.1, 38.9)
d_1 = no : 11 (37.3, 37.5, 37.8, 38.0, 38.0, 38.3, 38.5, 38.7, 39.0, 39.4, 39.5)

$$AUC = \frac{1+2+2\times3+3\times2+5+8}{9\times11} = \frac{28}{99}$$

$$= \boxed{0.2828}$$

c. $P(d_2 = yes) = \frac{10}{20} = \frac{1}{2}$ $P(d_2 = no) = \frac{10}{20} = \frac{1}{2}$

$P(a_1 \geq 37.95 | d_2 = yes) = \frac{9}{10}$ $P(a_1 < 37.95 | d_2 = yes) = \frac{1}{10}$

$P(a_1 \geq 37.95 | d_2 = no) = \frac{1}{10}$ $P(a_2 < 37.95 | d_2 = no) = \frac{9}{10}$

$P(a_2 = yes | d_2 = yes) = 0$ $P(a_2 = no | d_2 = yes) = 1$

$P(a_2 = yes | d_2 = no) = 1$ $P(a_2 = no | d_2 = no) = 0$

$P(a_3 = yes | d_2 = yes) = \frac{9}{10}$ $P(a_3 = no | d_2 = yes) = \frac{1}{10}$

$P(a_3 = yes | d_2 = no) = \frac{4}{10}$ $P(a_3 = no | d_2 = no) = \frac{6}{10}$

$P(d_2 = yes | a_1 = 40, a_2 = no, a_3 = no)$

$\propto P(a_1 \geq 37.95 | d_2 = yes) \times P(a_2 = no | d_2 = yes) \times P(a_3 = no | d_2 = yes)$

$\propto \frac{9}{10} \times \frac{1}{10} = \frac{9}{100}$

$P(d_2 = no | a_1 = 40, a_2 = no, a_3 = no) \propto \frac{1}{10} \times \frac{6}{10} = \frac{6}{100}$

Normalize

$$\boxed{\begin{aligned}&P(d_2=\text{yes}\mid a_1,a_2,a_3)=\frac{9}{100}\times\frac{100}{9+6}=\frac{9}{15}=0.6\\&P(d_2=\text{no}\mid a_1,a_2,a_3)=\frac{6}{100}\times\frac{100}{9+6}=\frac{6}{15}=0.4\\&\text{The predict result of } d_2 \text{ is yes}\end{aligned}}$$

5、

(a) $\mathcal{E}_\lambda(c,b)=\frac{1}{2}(\max\{0,1-y_1(cx_1+b)\}+\max\{0,1-y_2(cx_2+b\}+\lambda c^2$

For samples $(x_1,y_1)=(3,1)$ and $(x_2,y_2)=(1,-1)$

At point $c=b=1$

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

$$\frac{\partial\mathcal{E}_\lambda(c,b)}{\partial c}=\frac{1}{2}+c=\frac{3}{2}=1.5\ ;\ \frac{\partial\mathcal{E}_\lambda(c,b)}{b}=\frac{1}{2}=0.5$$

$$\boxed{\text{The gradient is } \nabla\mathcal{E}_\lambda(1,1)=(1.5,0.5)}$$

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

If $1-3c-b\geq 0$, $1\geq 3c+b\geq c+b$

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

$$=\frac{1}{2}(2-2c)+\frac{1}{2}c^2=1-c+\frac{1}{2}c^2$$

The minimizer is $c^*=1$, $\mathcal{E}_\lambda(c,b)=\frac{1}{2}$

If $1-3c-b\leq 0$, $3c+b\geq 1$

If $1+c+b\leq 0$, $c+b\leq -1$

$\mathcal{E}_\lambda(c,b)=\frac{1}{2}c^2$, The minimizer is $c=0$

which make a contradictiono is $b\geq 1$, $b\leq -1$

If $1+c+b\geq 0$, $c+b\geq -1$

$\mathcal{E}_\lambda(c,b)=\frac{1}{2}(1+c+b)+\frac{1}{2}c^2$. is convex

The minimizer is $c = -\frac{1}{2}$ $b = \frac{1}{2}$

$\mathcal{E}_\lambda(c,b) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} \geq \frac{1}{2}$.

Hence, $\boxed{\text{let } c^* = 1}$, which achieve the global minimizer

6. When $x^*(\gamma, \xi, q) = 0$, $f(x) = (x-\xi)^2 + \gamma|x|^q \Big|_{x=0} = \xi^2$

$x^* = 0$ is a minimizer, which means for any $x \in \mathbb{R}$, we have

$$(x-\xi)^2 + \gamma|x|^q \geq \xi^2$$

$$\Leftrightarrow x^2 - 2x\xi + \xi^2 + \gamma|x|^q \geq \xi^2 \Leftrightarrow x^2 - 2x\xi + \gamma|x|^q \geq 0$$

Note that $0 < q < 1$, $\gamma \geq 0$, $f(x) = x^2 - 2x\xi + \gamma|x|^q$ is convex

If $x \geq 0$, $f(x) = x^2 - 2x\xi + \gamma x^q$, $f'(x) = 2x - 2\xi + q\gamma x^{q-1} = 0$

$$\xi = x + \frac{q\gamma}{2} x^{q-1}$$

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

$$x^2 - 2x^2 - \gamma q x^q + \gamma x^q \geq 0 \Rightarrow$$

$$-x^2 + (1-q)\gamma x^q \geq 0 \Rightarrow$$

$$x^{2-q} \leq \gamma(1-q) \Rightarrow x = [\gamma(1-q)]^{\frac{1}{2-q}}$$

Consider the original eq

$$[\gamma(1-q)]^{\frac{2}{2-q}} - 2\xi[\gamma(1-q)]^{\frac{1}{2-q}} + \gamma[\gamma(1-q)]^{\frac{q}{2-q}} \geq 0$$

$$\xi \leq \frac{[\gamma(1-q)]^{\frac{2}{2-q}} + \gamma[\gamma(1-q)]^{\frac{q}{2-q}}}{2[\gamma(1-q)]^{\frac{1}{2-q}}}$$

$$\leq \frac{1}{2}\left\{[\gamma(1-q)]^{\frac{1}{2-q}} + \gamma[\gamma(1-q)]^{\frac{q-1}{2-q}}\right\}$$

Similarly, for $x < 0$, $\xi \geq -\frac{1}{2}\left\{[\gamma(1-q)]^{\frac{1}{2-q}} + \gamma[\gamma(1-q)]^{\frac{q-1}{2-q}}\right\}$

Hence $\boxed{|\xi| \leq \frac{1}{2}\left\{[\gamma(1-q)]^{\frac{1}{2-q}} + \gamma[\gamma(1-q)]^{\frac{q-1}{2-q}}\right\}}$