1\. 1. $Q(A, \text{Left}) = 1$ $\quad Q(B, \text{Right}) = 2$

2\. ~~two~~

1	-1
N/A	0
0.5	3
0.5	4

3\. $A \to \text{Right} \to A$ $\quad B \to \text{Left} \to B$

12 $\quad$ 10

iii

2\. True

False $\quad A \to C \to D$

False $\quad F \to G \to H$

True $\quad A \to C \to E$

False

3\. 1. $P(+c \mid +a, -d)$

$$= \frac{P(+c, +a, -d)}{P(+a, -d)}$$

$$= \frac{P(+c, +a, B, -d)}{P(+a, B, C, -d)}$$

$$= \cancel{P(+a)P(} \frac{P(+c, +a, +b, -d) + P(+c, +a, -b, -d)}{P(+a, +b, +c, -d) + P(+a, -b, +c, -d) + P(+a, +b, -c, -d) + P(+a, -b, -c, -d)}$$

$$= \frac{\frac{1}{4} \times \frac{1}{5} \times \frac{1}{2} \times \frac{5}{6} + \frac{1}{4} \times \frac{4}{5} \times \frac{3}{4} \times \frac{5}{6}}{\frac{1}{4} \times \frac{1}{5} \times \frac{1}{2} \times \frac{5}{6} + \frac{1}{4} \times \frac{4}{5} \times \frac{3}{4} \times \frac{5}{6} + \frac{1}{4} \times \frac{1}{5} \times \frac{1}{2} \times \frac{1}{8} + \frac{1}{4} \times \frac{4}{5} \times \frac{1}{4} \times \frac{1}{8}}$$

$$= \frac{\frac{1}{48} + \frac{1}{8}}{\frac{1}{48} + \frac{1}{8} + \frac{1}{320} + \frac{1}{160}}$$

$$= \frac{140}{149} = 0.9396$$

2. $\frac{2}{8} = \frac{1}{4}$

~~+a +b -c d +a~~

1

2

3 X

4 X

5 X

6

7 X

8 X

3. $\frac{1}{3} \times \frac{5}{6} = \frac{5}{18}$

$\frac{1}{5} \times \frac{5}{6} = \frac{1}{6}$

$\frac{1}{5} \times \frac{1}{8} = \frac{1}{40}$

$\frac{1}{3} \times \frac{1}{8} = \frac{1}{24}$

4. $P(-a \mid +b, -d)$

$$= \frac{\frac{5}{18} + \frac{1}{24}}{\frac{5}{18} + \frac{1}{6} + \frac{1}{40} + \frac{1}{24}}$$

$$= \frac{5}{8}$$

5. P(D|A) is better

~~Because Likelihood~~

Because evidence influences the choice of downstream variables, Liklihood conditions ~~on~~ is on upstream evidence

4. 1. $P(X_1=0, O_1=A) = 0.3 \times 0.9 = 0.27$

$P(X_1=1, O_1=A) = 0.7 \times 0.5 = 0.35$

$P(X_2=0, O_1=A, O_2=B) = 0.1 \times (0.4 \times 0.27 + 0.8 \times 0.35) = 0.0388$

$P(X_2=1, O_1=A, O_2=B) = 0.3 \times (0.6 \times 0.27 + 0.2 \times 0.35) = 0.116$

2. $P(X_1=1 \mid O_1=A, O_2=B)$

$$= \frac{P(X_1=1, O_1=A, O_2=B)}{P(O_1=A, O_2=B)}$$

$$\cancel{= P(X_1=1) \times P(O_1=A \mid X_1=1)}$$

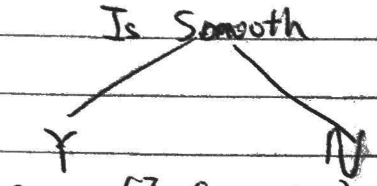
$$= \frac{P(X_1=1)\, P(O_1=A, O_2=B \mid X_1=1)}{P(O_1=A, O_2=B)}$$

$$= \frac{P(X_1=1)\,(0.35 \times 0.8 \times 0.1 + 0.35 \times 0.2 \times 0.5)}{P(X_2=0, O_1=A, O_2=B) + P(X_2=1, O_1=A, O_2=B)}$$

$= 0.28488$

① $-\frac{3}{8}\log\frac{3}{8} - \frac{5}{8}\log\frac{5}{8} \approx 0.95$

② Is Smooth

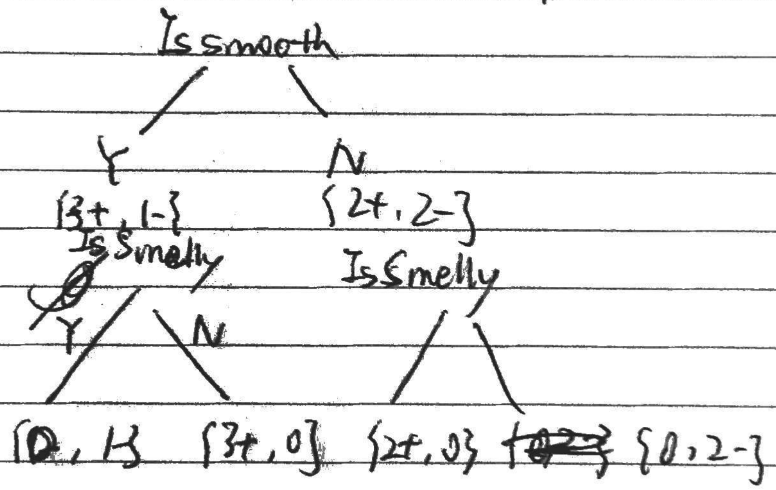


$$\text{Gain}(\text{IsSmooth}) = 0.95 - \frac{4}{8} \times \left(-\frac{3}{4}\log\frac{3}{4} - \frac{1}{4}\log\frac{1}{4}\right) - \frac{4}{8} \times \left(-\frac{2}{4}\log\frac{2}{4} - \frac{2}{4}\log\frac{2}{4}\right)$$

$$= 0.95 - 0.4013 - 0.5$$

$$= 0.0487$$

③



X: Not Poisonous

Y: Not poisonous

Z: poisonous

④

④ $P(IsPoisonous) = \frac{5}{8}$

$P(\sim IsPoisonous) = \frac{3}{8}$

$P(IsHeavy \mid IsPoisonous) = \frac{2}{5}$ $\quad P(IsHeavy \mid \sim IsPoisonous) = \frac{1}{3}$

$P(IsSmelly \mid IsPoisonous) = \frac{2}{5}$ $\quad P(IsSmelly \mid \sim Isp) = \frac{1}{3}$

$P(IsSpotted \mid Isp) = \frac{2}{5}$ $\quad P(IsSpotted \mid \sim Isp) = \frac{1}{3}$

$P(IsSmooth \mid Isp) = \frac{3}{5}$ $\quad P(IsSmooth \mid \sim Isp) = \frac{1}{3}$

$P(Isp \mid X) = \alpha \frac{2}{5} \times \frac{2}{5} \times \frac{2}{5} \times \frac{3}{5} \times \frac{5}{8} = \alpha\, 0.024$

$P(\sim Isp \mid X) = \alpha \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{3}{8} = \alpha\, 0.00463$

$\Rightarrow$ X is poisonous

$P(Isp \mid Y) = \alpha \cdot \frac{2}{5} \times \frac{3}{5} \times \frac{5}{8} = \alpha \cdot \frac{3}{20} = \alpha\, 0.15$

$P(\sim Isp \mid Y) = \alpha \cdot \frac{1}{3} \times \frac{1}{3} \times \frac{3}{8} = \alpha \cdot 0.041667$

$\Rightarrow$ Y is poisonous

$P(Isp \mid Z) = \alpha \frac{2}{5} \times \frac{2}{5} \times \frac{5}{8} = \alpha\, 0.1$

$P(\sim Isp \mid Z) = \alpha \frac{1}{3} \times \frac{1}{3} \times \frac{3}{8} = \alpha \cdot 0.041667$

$\Rightarrow$ Z is Poisonous

6 ~~Q~~. 1. $w_1 = 1$ $w_2 = 1$ $w_3 = 1$ $w_4 = 1$ $w_5 = 1$ $w_6 = 1$

$w_7 = 1$ $w_8 = -1$ $b_1 = -1.5$ $b_2 = -2.5$ $b_3 = -0.5$

2. In the hidden layer of this neural networks, the Left hidden unit will activate to be "1" if 2 or more inputs are "1". The right hidden unit will only activate to be "1" if three inputs are all "1". So given ~~$x_1 = 1$ x_2~~ $x_1 = 1$, $x_2 = 1$, $x_3 = 1$, h_1 will be activate $(1.5) = 1$ h_2 will be activate $(0.5) = 1$ then y will be activate $(-0.5) = 0$.

2. 1. Over fitting can cause the training model to be only accurate when it is being tested with traing data. While it has high acuracy with the training date, it will not guarantee a high acuracy ~~too~~ for testing data if testing data have more special cases. Two specific ways to deal with overfitting are regularization and dropout (turn some weights to 0)

2. If the validation error consistently goes up, it means that the model could be diverging because of high learning rate. To fix this, we can turn learning rate down a little bit.

3. False
4. False
5. 9