CMPT 825 Homework 0, Fall 2020

Coulton Fraser (301405411)

October 16, 2020

1. Probability 10pt

(a)

	Red	Green
Box A	1/3	2/3
Box B	1/3	2/3
Box C	1/3	2/3

$$
\begin{aligned}
P(Red) &= \frac{\sum_{P(Red)}}{\sum_{P(Red)} + \sum_{P(Green)}} \\
&= \frac{3(1/3)}{3(1/3) + 3(2/3)} \\
&= \frac{1}{3}
\end{aligned}
\tag{1}
$$

(b)

	Red	Green
Box A	5/15	0
Box B	4/15	0
Box C	6/15	0

$$
\begin{aligned}
P(Red|Box1) &= \frac{P(Red|BoxA)}{P(Red|BoxA) + P(Red|BoxB) + P(Red|BoxC)} \\
&= \frac{1}{3}
\end{aligned}
\tag{2}
$$

2. Probability 15pt

(a)

$$
\begin{aligned}
P(Pos) &= P(TruePos) + P(FalseNeg) \\
&= (0.005 \cdot 0.98) + (0.995 \cdot 0.03) \\
&= 0.03475 \\
&= 3.475\%
\end{aligned}
\tag{3}
$$

(b)

$$P(Disease|Pos) = \frac{P(Pos|Disease) + P(Disease)}{P(Pos)}$$

$$= \frac{0.98 \cdot 0.005}{0.03475} \tag{4}$$

$$= 0.141007$$

$$\approx 14.101\%$$

3. Probability 10pt

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

$$= \int_{1}^{2} x \cdot 2x^{-2} dx$$

$$= 2 \int_{1}^{2} \frac{1}{x} dx \tag{5}$$

$$= 2 \Big[\ln|x| \Big]_{1}^{2}$$

$$= 2 \Big[\ln|2| - \ln|1| \Big]$$

$$= 2 ln|2|$$

$$\approx 1.386$$

$$Var(X) = E(X^2) - E(X)^2$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

$$= \int_{1}^{2} x^2 \cdot 2x^{-2} dx$$

$$= 2 \int_{1}^{2} x^0 dx \tag{6}$$

$$= 2 \Big[x \Big]_{1}^{2}$$

$$= 2 \Big[2 - 1 \Big]$$

$$= 2$$

$$Var(X) = 2 - (2\ln(2))^2$$

$$\approx 0.078$$

4. Derivatives 10pt

(a)

$$f'(x) = 5(3x + 1)^4 \tag{7}$$

(b)

$$f'(x) = \frac{8x}{4x^2 + 9} \tag{8}$$

(c)

$$f'(u) = e^{-u} \cdot u'$$
$$f'(x) = -e^{-x}$$

<div align="right">(9)</div>

(d)

$$d/dx\left[\frac{f}{g}\right] = \frac{gf' - fg'}{g^2}$$
$$f'(x) = \frac{4x^2 - 8x^2 - 2x}{x^4}$$
$$= -\frac{4x + 2}{x^3}$$

<div align="right">(10)</div>

(e)

$$f'(x) = -\frac{1}{x^2}$$

<div align="right">(11)</div>

5. Matrix 10pt

(a)

$$1 \cdot 0 + -4 \cdot 5, 1 \cdot 3 + -4 \cdot 8$$
$$-2 \cdot 0 + 1 \cdot 5, -2 \cdot 3 + 1 \cdot 8$$
$$\begin{bmatrix} -20 & -29 \\ 5 & 2 \end{bmatrix}$$

<div align="right">(12)</div>

(b)

$$3 \cdot 1 + -2 \cdot -5 + 2 \cdot 8, 3 \cdot -3 + -2 \cdot 2 + 2 \cdot 3$$
$$7 \cdot 1 + 1 \cdot -5 + -2 \cdot 4, 7 \cdot 3 + 1 \cdot 2 + -2 \cdot 3$$
$$\begin{bmatrix} 29 & -11 \\ -6 & -25 \end{bmatrix}$$

<div align="right">(13)</div>

6. Matrix 10pt

(a)

$$\vec{y} = W\left[C \times D\right] \cdot \vec{x}\left[D \times 1\right]$$
$$\vec{y}\left[C \times 1\right]$$

<div align="right">(14)</div>

$\vec{y}$ Will have a length of C rows by one column.

(b)

$$\vec{y_i} = W_{i,d}\vec{x_d}$$
$$\vec{y_3} = W_{3,7}\vec{x_7}$$
$$\frac{\partial_{y3}}{\partial_{x7}} = W_{3,7}\partial_{x7}$$

<div align="right">(15)</div>

(c)

$$M' = \begin{bmatrix} 1 & 1 \\ 2x & e^x \\ 0 & 0 \end{bmatrix} \quad (16)$$

7. Information Theory 10pt

(a)

$$
\begin{aligned}
H(X) &= -\sum_{x \in X} p(x) log p(x) \\
H(P) &= -\left[0.1 log(0.1) + 0.98 log(0.98) 0.1 log(0.1) \right] \\
&\approx 0.6929 \\
H(Q) &= -\left[0.15 log(0.15) + 0.60 log(0.60) 0.25 log(0.25) \right] \\
&\approx 1.3527
\end{aligned}
\quad (17)
$$

(b)

$$
\begin{aligned}
H(P,Q) &= -\sum_{x \in X} p(P) log p(Q) \\
&= -\left[0.1 log(0.15) + 0.98 log(0.60) 0.1 log(0.25) \right] \\
&\approx 0.7696
\end{aligned}
\quad (18)
$$

(c) It is minimized when P=Q. It is maximized when P and Q have the least amount of mutual information or difference.

8. Information Theory 10pt

(a)

$$
\begin{aligned}
H(X) &= -\sum_{x \in X} p(x) log p(x) \\
&= -\left[2(\frac{1}{8} log(\frac{1}{2})) + 6(\frac{1}{16} log(\frac{1}{4})) + 4(\frac{1}{32} log(\frac{1}{4})) + \frac{1}{4} log(\frac{1}{4}) \right] \\
&= 1.75 \\
H(X|Y) &= \sum_{x,y} p(x,y) log \frac{1}{p(x|y)} \\
&= 1.375 \\
MI &= H(X) - H(X|Y) \\
&= 0.375
\end{aligned}
\quad (19)
$$

(b) $0 \geq MI \leq \min(H(X), H(Y))$

9. Cosine Similarity 10pt

$$\cos\theta = \frac{2 \times 1 + 2 \times 3}{\sqrt{2^2 + 2^2} \times \sqrt{1^2 + 3^2}} \tag{20}$$

$$= 0.8944$$