SENG 474 - Assignment 1

Derrop Banerjee
V00837868

1. a) Entropy $= \sum_{i=0}^{n} P(x_i) \cdot \log_2 P(x_i)$

① Age -
- $P(\text{young}) = 1/3$ · $P(\text{pre-presbyopic}) = 1/3$ · $P(\text{presbyopic}) = 1/3$

② Spectacle prescription -
- $P(\text{myope}) = 1/2$ · $P(\text{hypermetrope}) = 1/2$

③ ~~Astigm~~ Astigmatism -
- $P(\text{no}) = 1/2$ · $P(\text{yes}) = 1/2$

④ Tear-prod-rate -
- $P(\text{reduced}) = 1/2$ · $P(\text{normal}) = 1/2$

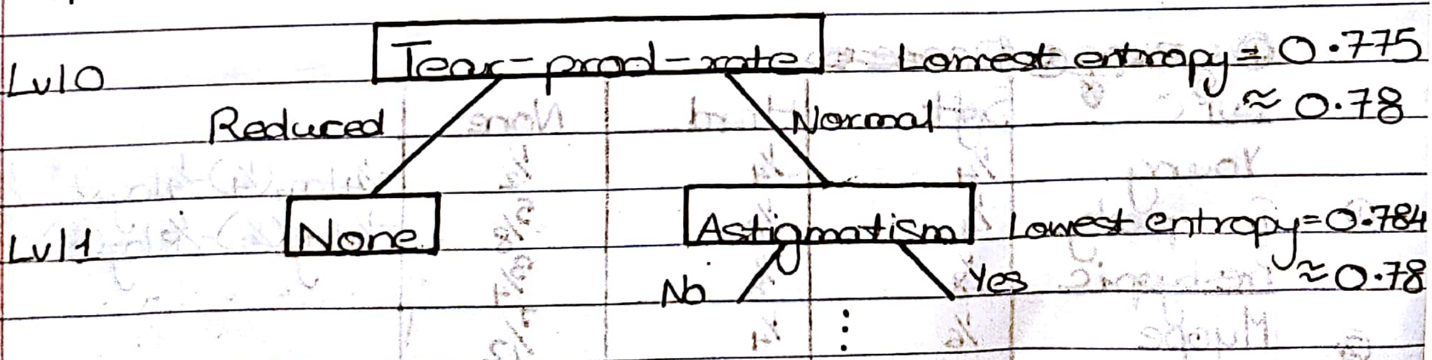
⑤ Contact Lenses -
- $P(\text{soft}) = 5/24$ · $P(\text{hard}) = 1/6$ · $P(\text{none}) = 5/8$

	LVO	Soft	Hard	None	
①	Young	1/4	1/4	1/2	$-\frac{1}{4}\log_2(\frac{1}{4}) - \frac{1}{4}\log_2(\frac{1}{4}) - \frac{1}{2}\log_2(\frac{1}{2}) = 1.5$
	Pre-presby.	1/4	1/8	5/8	$-\frac{1}{4}\log_2(\frac{1}{4}) - \frac{1}{8}\log_2(\frac{1}{8}) - \frac{5}{8}\log_2(\frac{5}{8}) = 1.3$
	Presbyopic	1/8	1/8	3/4	$= 1.06$
②	Myope	1/6	1/4	7/12	$= 1.38$
	Hypermetrope	1/4	1/12	2/3	$= 1.19$
③	No	5/12	0/12	7/12	$= 0.98$
	Yes	0/12	1/3	2/3	$= 0.92$
④	Reduced	0/12	0/12	12/12	$= 0$
	Normal	5/12	1/3	1/4	$= 1.55$

Expected Age = (1·5)(⅓)+(1·3)(⅓)+(1·06)(⅓) = 1·29.
Expected Specs = (1·38)(½)+(1·18)(½) = 1·28
Expected Astigmatism = (0·98)(½)+(0·92)(½) = 0·95
Expected Tears = (0)(½)+(1·55)(½) = 0·78.

Lvl 1	Soft	Hard	None	
Young	½	½	0/4	$-\frac{1}{2}\log_2(\frac{1}{2})-\frac{1}{2}\log_2(\frac{1}{2}) = 1$
Pre-presbyopic	½	¼	¼	= 1·5
Presbyopic	¼	¼	½	= 1·5
Myope	⅓	½	⅙	= 1·46
Hypermetrope	½	⅙	⅓	= 1·46
No	5/6	0/6	⅙	= 0·65
Yes	0/6	2/3	⅓	= 0·92
Reduced	–	–	–	
Normal	–	–	–	

Expected Age1 = (1)(⅓)+(1·5)(⅓)+(1·5)(⅓) = 1·33
Expected Specs1 = (1·46)(½)+(1·46)(½) = 1·46
Expected Astigmatism1 = (0·65)(½)+(0·92)(½) = 0·78
~~Expected Tears~~



b) The difference in my calculated entropy values and the values obtained from Scikit can be explained using multiple reasons such as rounding errors, calculation errors, human errors. All errors aside, computers and humans have different calculation algorithms (not neccessarily, but computers are known to use more efficient algorithms which results in negligable errors in calculation.

2. a) $Pr(R) = 33.33\%$
$Pr(Y) =$ "
$Pr(B) =$ "

Prob(R)	Pred¬R	Prob(Y)	Pred¬Y	Prob(B)	Pred¬B
1/3	0	1/3	1	1/3	1

$\Rightarrow (1/3 \times 0) + (1/3 \times 1) + (1/3 \times 1)$
$\Rightarrow 2/3$

$\therefore$ Expected error rate $= 2/3$

b)

Pr(R)	Pred(¬R)	Pr(Y)	Pred(¬Y)	Pr(B)	Pred(¬B)
1/3	0.3	1/3	1	1/3	0.7

$Pred(\neg R) = Pr(Y) \vee Pr(B) = 0 + 0.3 = 0.3$
$Pred(\neg Y) = Pr(R) \vee Pr(B) = 0.7 + 0.3 = 1$
$Pred(\neg B) = Pr(Y) \vee Pr(R) = 0 + 0.7 = 0.7$

$\Rightarrow (1/3 \times 0.3) + (1/3 \times 1) + (1/3 \times 0.7)$
$\Rightarrow 1/10 + 1/3 + 7/30$
$\Rightarrow 2/3$

$\therefore$ Expected error rate $= 2/3$

c)

Pr(R)	Pred(¬R)	Pr(Y)	Pred(¬Y)	Pr(B)	Pred(¬B)
1/2	0	1/4	1	1/4	1

$\Rightarrow (1/2 \times 0) + (1/4 \times 1) + (1/4 \times 1)$
$\Rightarrow 1/2$

$\therefore$ Expected error rate $= 1/2$

d)

$Pr(R)$	$Pred(\neg R)$	$Pr(Y)$	$Pred(\neg Y)$	$Pr(B)$	$Pred(\neg B)$
$1/2$	0.3	$1/4$	1	$1/4$	0.7

$\Rightarrow (1/2 \times 0.3) + (1/4 \times 1) + (1/4 \times 0.7)$

$\Rightarrow 3/20 + 1/4 + 7/40$

$\Rightarrow 23/40$

$\therefore$ Expected error rate $= 23/40$.

3. a) $\theta = Pr(x = T)$ $\qquad D = \{T, T, T, T, T, T, T, F, F, F\}$

~~[illegible]~~ $\alpha_T = \# T \qquad \alpha_F = \# F$

$P(D) = P(T \wedge T \wedge T \wedge T \wedge T \wedge T \wedge T \wedge F \wedge F \wedge F)$

$= \theta^7 (1-\theta)^3$

$\log_e(P(D)) = \log_e(\theta^7 (1-\theta)^3)$

$= 7 \log_e \theta + 3 \log(1-\theta)$

$\frac{d \log(P(D))}{d\theta} = 7/\theta + \left(-\frac{3}{(1-\theta)}\right)$

$\frac{7}{\theta} - \frac{3}{(1-\theta)} = 0$

$\Rightarrow \frac{7(1-\theta) - 3\theta}{\cancel{\theta(1-\theta)}} = 0 \qquad \cancel{(\times \theta(1-\theta))}$

$\Rightarrow 7 - 7\theta - 3\theta = 0$

$\Rightarrow 7 - 10\theta = 0$

$\Rightarrow 10\theta = 7$

$\Rightarrow \theta = 7/10$

b) $\theta^{(\beta_1-1)} \times (1-\theta)^{(\beta_2-1)}$ $\boxed{\theta = 1/2}$

$\theta = \frac{\beta_1}{\beta_2 - 1 + \beta_1 - 1}$

$P(\theta) \propto \theta^{(\beta_1-1)}(1-\theta)^{(\beta_2-1)}$.

$\Rightarrow 1/2 = \frac{4}{\beta_2 - 1 + 4 - 1}$

$\Rightarrow P(\theta|D) = \theta^7(1-\theta)^3\,\theta^{(\beta_1-1)}\theta^{(\beta_2-1)}$ $\boxed{\beta_1 = 4}$

$\Rightarrow 1/2 = \frac{4}{\beta_2+2}$

$\Rightarrow P(\theta|D) = (\frac{1}{2})^7(\frac{1}{2})^3(\frac{1}{2})^3(\frac{1}{2})^5$

$\Rightarrow \beta_2 + 2 = 8$

$\Rightarrow P(\theta|D) = (\frac{1}{2})^{18}$

$\Rightarrow \boxed{\beta_2 = 6}$

$\Rightarrow P(\theta|D) = \frac{1}{2^{18}}$

4. a) $\hat{y}_i = w_0 + w_1 x_{i,1} + w_2 x_{i,2}^4$

$$E(x) = \frac{1}{2N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$
$$= \frac{1}{2N}\sum_{i=1}^{N}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)^2$$

b) $w_0 \leftarrow w_0 - K\frac{d}{dw_0}E(w_0, w_1, w_2)$
$$= w_0 - K\frac{d}{dw_0}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)^2\right)$$
$$= w_0 - K\frac{1}{2N}\sum_{i=1}^{N}\frac{d}{dw_0}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)^2$$
$$= w_0 + \frac{K}{N}\sum_{i=1}^{N}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)$$

c) $w_1 \leftarrow w_1 - K\frac{d}{dw_1}E(w_0, w_1, w_2)$
$$= w_1 - K\frac{d}{dw_1}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)^2\right)$$
$$= w_1 + \frac{K}{N}\sum_{i=1}^{N}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)\,x_{i,1}$$

d) $w_2 \leftarrow w_2 - K\frac{d}{dw_2}E(w_0, w_1, w_2)$
$$= w_2 - K\frac{d}{dw_2}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)^2\right)$$
$$= w_2 + \frac{K}{N}\sum_{i=1}^{N}(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^4)\,x_{i,2}^4$$