Michael Reiter
V00831568

SENG 474 Assignment 2

1\. a) $\text{Success rate} = \underset{\text{red}}{(1)(\tfrac{1}{3})} + \underset{\text{blue}}{(0)(\tfrac{1}{3})} + \underset{\text{yellow}}{(0)(\tfrac{1}{3})} = \tfrac{1}{3}$

$\text{Error rate} = 1 - \text{success rate} = 1 - \tfrac{1}{3} = \boxed{\tfrac{2}{3}}$

b) $\text{Success rate} = \underset{\text{red}}{(0.7)(\tfrac{1}{3})} + \underset{\text{blue}}{(0.3)(\tfrac{1}{3})} + \underset{\text{yellow}}{(0)(\tfrac{1}{3})} = \tfrac{1}{3}$

$\text{Error rate} = 1 - \text{success rate} = 1 - \tfrac{1}{3} = \boxed{\tfrac{2}{3}}$

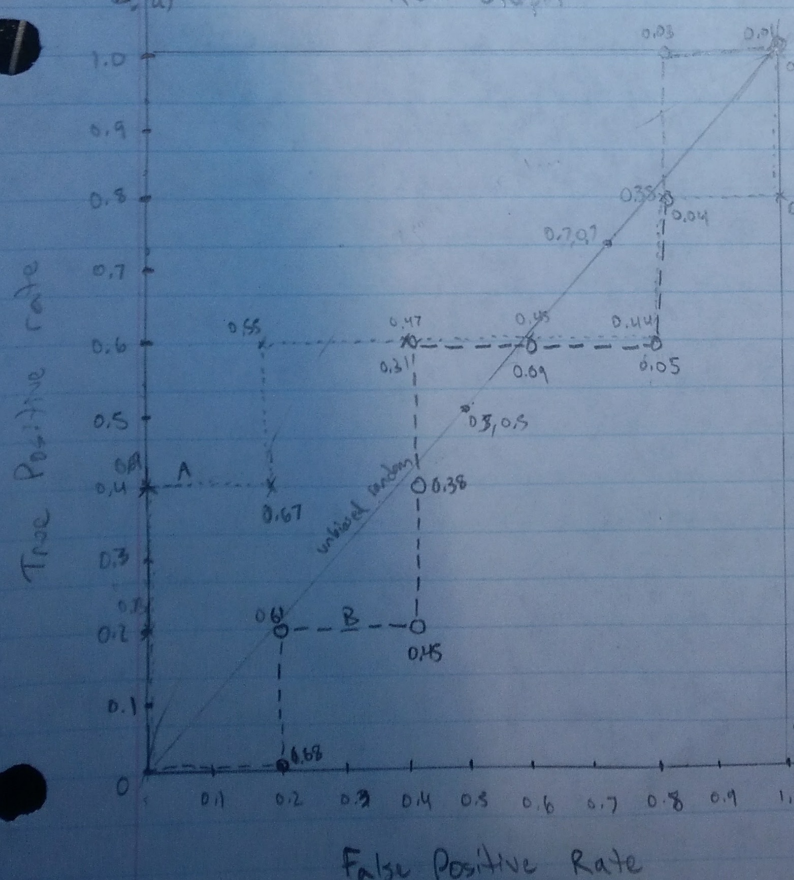
c) $\text{Success rate} = \underset{\text{red}}{(1)(\tfrac{1}{2})} + \underset{\text{blue}}{(0)(\tfrac{1}{4})} + \underset{\text{yellow}}{(0)(\tfrac{1}{4})} = \tfrac{1}{2}$

$\text{Error rate} = 1 - \text{success rate} = 1 - \tfrac{1}{2} = \boxed{\tfrac{1}{2}}$

d) $\text{Success rate} = (0.7)(\tfrac{1}{2}) + (0.3)(\tfrac{1}{4}) + (0)(\tfrac{1}{4}) = 0.425$

$\text{Error rate} = 1 - \text{success rate} = 1 - 0.425 = \boxed{0.575}$

2\. a) ROC Graph



Instance	Class	A
1	P	0.73
2	P	0.69
5	N	0.67
4	P	0.55
6	N	0.47
9	N	0.45
3	N	0.44
10	P	0.35
8	N	0.15
7	P	0.08

A : x·········x

B o— — — —o

random ————

Instance	Class	B
3	N	0.68
1	P	0.61
5	N	0.45
7	P	0.38
4	P	0.31
6	N	0.09
8	N	0.05
10	P	0.04
2	P	0.03
9	N	0.01

b) $t = 0.5$

Precision: $\frac{3}{3+1} = 75\%$ Precise

Recall: $\frac{3}{3+2} = 60\%$ Recall

F-measure: $2 \times (P \times R)/(P+R) = 2\frac{(0.75)(0.6)}{(0.75)+(0.6)} = \frac{2}{3} = 66.\bar{6}\%$

c) $t = 0.5$

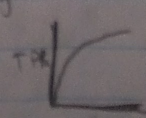
Precision $= \frac{1}{1+1} = 50\%$ Precise

Recall $= \frac{1}{1+4} = 20\%$ Recall

F-measure $= 2 \times (P \times R)/(P+R) = 2\frac{(0.5)(0.2)}{(0.5+0.2)} = \frac{2}{7} = 28.6\%$

Since A has a greater F-measure (F1) than B, A is a better classifier. This is expected based on the ROC from part a) since classifier A clearly has a better precision and recall for $t = 0.5$.

d) A performs better than the unbiased random classifier between $t = 0.45$ and $t = 1$. B performs better than the unbiased random classifier between $t = 0.01$ and $t = 0.04$ as well as between $t = 0.09$ and $t \geq 0.38$. Unbiased random is a straight line from bottom left to top right.

e) This is plotted approximately as follows. 

See graph for details

3. a) $P(D|\theta) = \prod_{i=1}^{N} P(x_i|\theta)$

αT is # True $= 7$ (from D)
αF is # false $= 3$

$$= \prod_{i=1}^{N} \theta^{x_i}(1-\theta)^{1-x_i}$$

$$= \theta^{\alpha T}(1-\theta)^{\alpha F}$$

solve for max θ
$\Rightarrow$ set derivative to 0

$$\frac{d}{d\theta} P(D|\theta) = \frac{d}{d\theta} \theta^{\alpha T}(1-\theta)^{\alpha F}$$

$$0 = \alpha T \theta^{\alpha T-1}(1-\theta)^{\alpha F} - \theta^{\alpha T} \alpha F (1-\theta)^{\alpha F-1}$$

$$\alpha_T \theta^{\alpha T-1}(1-\theta)^{\alpha F} = \theta^{\alpha T} \alpha F (1-\theta)^{\alpha F-1}$$

$$7\theta^6(1-\theta)^3 = 3\theta^7(1-\theta)^2$$

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

$$7 - 7\theta = 3\theta$$

$$7 = 10\theta$$

$$P(x=T) = \theta = \frac{7}{10}$$

3.b) $\text{Beta} \sim \theta^{\beta_1 - 1}(1-\theta)^{\beta_2 - 1}$

$\beta_1 = 4$, $P(X=T) = 1/2$, $\alpha_T = 7$, $\alpha_F = 3$

$$P(X=T) = \frac{\beta_1 - 1}{(\beta_1 - 1) + (\beta_2 - 1)} = \frac{1}{2}$$

$$\frac{4-1}{(4-1)+(\beta_2 - 1)} = \frac{1}{2}$$

$$\frac{3}{\beta_2 + 2} = \frac{1}{2}$$

$$6 = \beta_2 + 2$$
$$\beta_2 = 4$$

$$p(\theta|D) \propto p(D|\theta)P(\theta) = \theta^{\alpha_T}(1-\theta)^{\alpha_F}\,\theta^{\beta_1 - 1}(1-\theta)^{\beta_2 - 1}$$
$$= \theta^{7}(1-\theta)^{3}\,\theta^{4-1}(1-\theta)^{4-1}$$
$$= \theta^{10}(1-\theta)^{6}$$

$$0 = \frac{\partial}{\partial\theta}P(\theta|D) = \frac{\partial}{\partial\theta}\theta^{10}(1-\theta)^{6}$$

$$\frac{d}{d\theta}\log P(\theta|D) = \frac{d}{d\theta}\,10\log\theta + 6\log(1-\theta)$$

$$0 = \frac{10}{\theta} - \frac{6}{1-\theta}$$

$$\frac{10}{\theta} = \frac{6}{1-\theta}$$

$$\frac{1-\theta}{\theta} = \frac{6}{10}$$

$$\frac{1}{\theta} - 1 = \frac{6}{10}$$

$$\frac{1}{\theta} = \frac{16}{10}$$

$$\theta = \frac{10}{16} = \frac{5}{8}$$

4. a) $E(X) = \frac{1}{2N} \sum_{i=1}^{N} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2 \right)^2$

b) $\frac{\partial}{\partial w_0} E(X)$

$$= \frac{\partial}{\partial w_0} \frac{1}{2N} \sum_{i=1}^{N} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2 \right)^2$$

$$= \frac{1}{2N} \sum_{i=1}^{N} \frac{\partial}{\partial w_0} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2 \right)^2$$

$$= \frac{1}{2N} \sum_{i=1}^{N} 2 \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2 \right)(-1)$$

$$= -\frac{1}{N} \sum_{i=1}^{N} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2 \right)$$

so the update for w_0 is

$$w = w - K \left(-\frac{1}{N} \sum_{i=1}^{N} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2 \right) \right)$$

4. c) $\frac{\partial}{\partial w_1} E(X)$

$$= \frac{\partial}{\partial w_1} \frac{1}{2N} \sum_{i=1}^{N} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)^2$$

$$= \frac{1}{2N} \sum_{i=1}^{N} \frac{\partial}{\partial w_1} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)^2$$

$$= \frac{1}{2N} \sum_{i=1}^{N} 2\left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)\left(-x_{i,1}\right)$$

$$= \frac{-1}{N} \sum_{i=1}^{N} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)\left(x_{i,1}\right)$$

so the update for w_1 is

$$w = w - K\left[\frac{-1}{N} \sum_{i=1}^{N} (x_{i,1})\left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)\right]$$

d) $\frac{\partial}{\partial w_2} E(X)$

$$= \frac{\partial}{\partial w_2} \frac{1}{2N} \sum_{i=1}^{N} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)^2$$

$$= \frac{1}{2N} \sum_{i=1}^{N} \frac{\partial}{\partial w_2} \left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)^2$$

$$= \frac{1}{2N} \sum_{i=1}^{N} 2\left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)\left(-x_{i,2}^2\right)$$

$$= \frac{-1}{N} \sum_{i=1}^{N} \left(x_{i,2}^2\right)\left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)$$

so the update for w_2 is

$$w = w - K\left[\frac{-1}{N} \sum_{i=1}^{N} (x_{i,2}^2)\left(y_i - w_0 - w_1 x_{i,1} - w_2 x_{i,2}^2\right)\right]$$