CS 4780 HW 2

1(a)i.) $P(X_i) = \begin{cases} p & X_i = 1 \\ 1-p & X_i = 0 \end{cases}$ $\quad p_s = \frac{n_1}{n} = \frac{139}{285} = 0.4877$

X_i corresponds to the doctor's diagnosis of patient i, the test data.

$\bar{X}$ is the sample probability of X_i being equal to 1.

n is the number of images evaluated by the doctor while n_1 is the number in this sample classified by the doctor as 1.

ii.) $\alpha = P(|\bar{X} - \mathbb{E}[\bar{X}]| > \epsilon) \leq 2\exp(-2n\epsilon^2)$

$$\alpha \leq 2\exp(-2n\epsilon^2)$$

$$\frac{\alpha}{2} \leq e^{-2n\epsilon^2}$$

$$\log\left(\frac{\alpha}{2}\right) \leq -2n\epsilon^2$$

$$-\frac{\log\left(\frac{\alpha}{2}\right)}{2n} \leq \epsilon^2$$

$$\sqrt{-\frac{\log\left(\frac{\alpha}{2}\right)}{2n}} \leq \epsilon$$

ϵ is deviation from mean thus interval is +/- this amount from sample mean

$$err_D(\hat{h}) \in \left[err_S(\hat{h}) - \sqrt{\frac{-0.5\log\left(\frac{\alpha}{2}\right)}{n}},\; err_S(\hat{h}) + \sqrt{\frac{-0.5\log\left(\frac{\alpha}{2}\right)}{n}} \right]$$

iii $err_p(h_A) = \left[err_s(h_A) - \sqrt{\frac{-0.5 \log\left(\frac{\alpha}{2}\right)}{n}},\ err_s(h_A) \sqrt{\frac{-0.5 \log\left(\frac{\alpha}{2}\right)}{n}} \right.$

$$err_p(h_A) = \left[\frac{30}{285} - \sqrt{\frac{-0.5 \log\left(\frac{0.025}{2}\right)}{285}},\ \frac{30}{285} + \sqrt{\frac{-0.5 \log\left(\frac{0.025}{2}\right)}{285}} \right.$$

$$err_p(h_A) = \left[0.0478,\ 0.16305 \right]$$ ← 97.5% CI for h_A

$$err_p(h_B) = \left[err_s(h_B) - \sqrt{\frac{-0.5 \log\left(\frac{\alpha}{2}\right)}{n}},\ err_s(h_B) + \sqrt{\frac{-0.5 \log\left(\frac{\alpha}{2}\right)}{n}} \right]$$

$$= \left[\frac{36}{285} - \sqrt{\frac{-0.5 \log\left(\frac{0.025}{2}\right)}{285}},\ \frac{36}{285} + \sqrt{\frac{-0.5 \log\left(\frac{0.025}{2}\right)}{285}} \right.$$

$$= \left[0.0853,\ 0.18410 \right]$$ ← 97.5% CI for h_B

(b) H_0: $err_p(\hat{h}_A) = err_p(\hat{h}_B)$
H_A: $err_p(\hat{h}_A) \neq err_p(\hat{h}_B)$

wins h_A over h_B : 15
wins h_B over h_A : 9

$P(W \leq w \mid p = 0.5, n = w+1)$

$P(W \leq 9 \mid p = 0.5, n = 24)$

$$= \sum_{i=0}^{i=9} \binom{24}{i} (0.5)^{24} = 2^9 (0.5)^{24} = 3.0518 \times 10^{-5}$$

as $p = 3.0518 \times 10^{-5} < 0.1$,
we reject the null hypothesis and conclude that the true error rates of h_A and h_B are significantly different

2.) (a) $h_{w,b}(x) = \begin{cases} +1, & \text{if } w \cdot x + b > 0 \\ -1, & \text{otherwise} \end{cases}$

$w = -1 \quad b = 0$

(b)

$h_{\vec{w},b}(\vec{x}) = \begin{cases} +1, & \text{if } \vec{w} \cdot \vec{x} + b > 0 \\ -1, & \text{otherwise} \end{cases}$

$\vec{w} = (1, \dots, 1)$ length d
$b = -n$

(c) $\vec{w} = (\underbrace{1, \dots, 1}_{1-10=1}, \underbrace{0, \dots, 0}_{11-90=0}, \underbrace{-1, \dots, -1}_{91-100=-1})$
$b = -5$

(d) $\vec{w} = (1, 1)$
$b = 1.5$

	0	1
0	-1	-1
1	-1	1

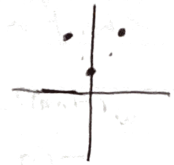
(e) $\vec{w} = (c_1 2^{t}, c_2 2^{t-1}, \dots, c_{t-1} 2^{t-(t-2)}, c_t 2^{t-(t-1)}, 0, \dots 0) \times (d-t+1)$
$b = -(d-t+1)$

3) (a) $\vec{w}^{(0)} = (0,0)$

While there is $i \in [1,2,3]$ s.t. $y_i(\vec{w}^{(t)} \cdot \vec{x}_i) \leq 0$

$1(0\cdot 1 + 0 \cdot 3) \leq 0$

$y_1(w^{(0)} \cdot \vec{x}_1) \leq 0$



$w^{(1)} = \vec{w}^{(0)} + y_1 x_1$

$= (0,0) + 1(1,3)$

$\vec{w}^{(1)} = (1,3)$

$y_1(w^{(1)} \cdot \vec{x}_1) = 1(1\cdot 1 + 3 \cdot 3) = 10 \not\leq 0$

$y_2(w^1 \cdot \vec{x}_2) = -1(1 \cdot -1 + 3 \cdot 3) = -8 \leq 0$

$w^{(2)} = w^{(1)} + y_2 \vec{x}_2 = (1,3) + -1(-1,3) = (2,0)$

$y_1(w^{(2)} \cdot \vec{x}_1) = 1(2 \cdot 1 + 0 \cdot 3) = 2 \not\leq 0$

$y_2(w^{(2)} \cdot \vec{x}_2) = -1(2 \cdot -1 + 0 \cdot 3) = 2 \not\leq 0$

$y_3(w^{(2)} \cdot \vec{x}_3) = 1(2 \cdot 0 + 0 \cdot 1) = 0 \leq 0$

$w^{(3)} = w^{(2)} + y_3 \vec{x}_3 = (2,0) + 1(0,1) = (2,1)$

$y_1(w^{(3)} \cdot \vec{x}_1) = 1(2 \cdot 1 + 1 \cdot 3) = 5 \not\leq 0$

$y_2(w^{(3)} \cdot \vec{x}_2) = -1(2 \cdot -1 + 1 \cdot 3) = -1 \leq 0$

$w^{(4)} = w^{(3)} + y_2 \vec{x}_2 = (2,1) + -1(-1,3) = (3,-2)$

$y_1(w^{(4)} \cdot \vec{x}_1) = 1(3 \cdot 1 + -2 \cdot 3) = -3 \leq 0$

$w^{(5)} = w^{(4)} + y_1 \vec{x}_1 = (3,-2) + 1(1,3) = (4,1)$

$y_1(w^{(5)} \cdot \vec{x}_1) = 1(4 \cdot 1 + 1 \cdot 3) = 7 \not\leq 0$

$y_2(w^{(5)} \cdot \vec{x}_2) = -1(4 \cdot -1 + 1 \cdot 3) = 1 \not\leq 0$

$y_3(w^{(5)} \cdot \vec{x}_3) = 1(4 \cdot 0 + 1 \cdot 1) = 1 \not\leq 0$

$w^{(0)} = (0,0)$

$w^{(1)} = (1,3)$

$w^{(2)} = (2,0)$

$w^{(3)} = (2,1)$

$w^{(4)} = (3,-2)$

$w^{(5)} = (4,1)$

(b) $w^{(t+1)} = w^{(t)} + y_i \vec{x}_i$, thus if we know the number of times each i is used to update, we can simply multiply the addition that many times and sum.

$$w^{(final)} = w^{(0)} + (1) y_1 \vec{x}_1 + (3) y_2 \vec{x}_2 + (1) y_3 \vec{x}_3 + (1) y_4 \vec{x}_4 + (1) y_5 \vec{x}_5$$

$$= (0,0,0,0,0) + (1)(1)(1,3,3,7,0) + 3(-1)(0,0,0,0,1)$$

$$+ 1(1)(0,0,0,1,0) + 1(1)(9,1,2,0,0) + 1(-1)(2,8,4,0,-2)$$

$$w^{(final)} = (8,-4,1,8,-1)$$