Question 1

Consider again the example application of Bayes rule in Section 6.2.1 of Tom Mitchell's textbook (or slide page 6 of Lecture 6-2). Suppose the doctor decides to order a second laboratory test for the same patient, and suppose the second test returns a positive result as well. What are the posterior probabilities of *cancer* and ¬*cancer* following these two tests? Assume that the two tests are independent.

Answer

Since,

$P(cancer \mid +) \Rightarrow P(+ \mid cancer) * P(cancer) / (P(+ \mid cancer) * P(cancer) + P(+ \mid \neg cancer) * P(\neg cancer))$.

$\Rightarrow 0.98 * 0.008 / (0.98 * 0.008 + 0.03 * 0.992)$
$\Rightarrow 0.0078 / (0.0078 + 0.0298)$
$\Rightarrow 0.21$

And,

$P(cancer \mid ++) \Rightarrow P(+ \mid cancer) * P(cancer \mid +) / (P(+ \mid cancer) * P(cancer \mid +) + P(+ \mid \neg cancer) * P(\neg cancer \mid +))$

We know that,

$P(cancer \mid +) \Rightarrow 0.21$ and,
$P(\neg cancer \mid +) \Rightarrow 1 - P(cancer \mid +) \Rightarrow 0.79$

Therefore,

$P(cancer \mid ++) \Rightarrow 0.98 * 0.21 / (0.98 * 0.21 + 0.03 * 0.79)$
$\Rightarrow 0.21 / (0.21 + 0.02)$
$\Rightarrow$ **0.91**

$P(\neg cancer \mid ++) \Rightarrow 1 - P(cancer \mid ++) \Rightarrow 1 - 0.91 \Rightarrow$ **0.09**

Question 2

Consider a learned hypothesis, *h*, for some Boolean concept. When *h* is tested on a set of 100 examples, it classifies 80 correctly. What is the 95% confidence interval for the true error rate for $Error_D(h)$?

Answer

M = Misclassified Examples
N = Total Examples

$error_D(h)$ => m/n
=> 20/100
=> **0.20**

95% confidence interval =>

$error_D(h) \pm 1.96 * \sqrt{[(error_D(h)*(1 - error_D(h)) / n]}$
=> $0.20 \pm 1.96 * \sqrt{[(0.20*(1-0.20)) / 100]}$
=> **0.1216 to 0.2784**

Question 3

Consider a two-layer feedforward ANN with two inputs a and b, one hidden unit c, and one output unit d. This network has five weights (w_{ca}, w_{cb}, w_{c0}, w_{dc}, w_{d0}), where w_{x0} represents the threshold weight for unit x. Initialize these weights to the values (.1, .1, .1, .1, .1), then give their values after each of the first two training iterations of the BACKPROPAGATION algorithm. Assume learning rate $\eta = .3$, momentum $\alpha = 0.9$, incremental weight updates, and the following training examples:

a	b	d
1	0	1
0	1	0

(Answer Below)

First Iteration

$$net_c = w_{c0} + a * w_{ca} + b * w_{cb}$$
$$= 0.1 + 0.1 + 0$$
$$\Rightarrow 0.2$$

$$O_c = \frac{1}{1+e^{-net_c}} = \frac{1}{1+e^{-0.2}} = 0.55$$

$$net_d = w_{d0} + O_c * w_{dc} = 0.1 + 0.55 * 0.1$$
$$= 0.155$$

Similarly,

$$O_d = \frac{1}{1+e^{-net_d}} = \frac{1}{1+e^{-0.155}} \approx 0.539$$

Using Backpropogation,

$$\delta_d = O_d * (1 - O_d) * (t_a - O_a)$$
$$= 0.539 * (1 - 0.539) * (1 - 0.539)$$
$$= 0.115$$

$$\Delta w_{dc} = \eta * \delta_d * O_c + \alpha * 0$$
$$= 0.3 * 0.115 * 0.55$$
$$= 0.019$$

$$\Delta w_{d0} = 0.034$$

$$\therefore w_{dc} = w_{dc} + \Delta w_{dc}$$
$$= 0.1 + 0.019 = 0.119$$

and,

$$\Delta w_{d0} = w_{d0} + \Delta w_{d0}$$
$$= 0.1 + 0.034 = 0.134$$

$\delta_c = O_c * (1 - O_c) * (w_c * \delta_d)$
$= 0.55 * (1-0.55) * (0.1 * 0.115)$
$= \underline{0.003}$

$\underline{\Delta w_{ca}} : \eta * \delta_c * x_a + \alpha * 0$
$= 0.3 * 0.003 * 1$
$= \underline{0.001}$

& $\underline{\Delta w_{co}} = 0.\underline{001}$
$\underline{\Delta w_{cb}} = \underline{0}$

$w_{co} = w_{co} + \Delta w_{co} = 0.1 + 0.001 = \underline{0.101}$
$\underline{w_{ca}} = w_{ca} + \Delta w_{ca} = 0.1 + 0.001 = \underline{0.101}$
$\underline{w_{cb}} = w_{cb} + \Delta w_{cb} = 0.1 + 0 = \underline{0.100}$

Second Iteration

$\underline{net_c} = w_{co} + a \cdot w_{ca} + b + w_{cb}$
$= 0.101 + 0 + 0.100$
$= \underline{0.201}$

$\underline{O_c} = \dfrac{1}{1+e^{-0.201}} = \underline{0.55}$

$\underline{net_d} = w_{do} + O_c * w_{dc}$
$= 0.134 + 0.55 * 0.119$
$= 0.\underline{1994}$

Simlarly,

$\underline{O_d} = \dfrac{1}{1+e^{-0.1994}} = 0.\underline{5496}$

Using Backpropogation,

$\delta_d = O_d * (1 - O_d) * (t_d - O_d)$

$= 0.5496(1 - 0.5496)(0 - 0.5496)$

$= -0.136$

$\Delta W_{dc} = \eta * \delta_d * O_c + \alpha * \Delta W_{dc}(old)$

$= 0.3 * (-0.136) * 0.55 + 0.9 * 0.019$

$= -0.0053$

$\& \Delta W_{d0} = -0.01$

$\therefore W_{dc} = W_{dc} + \Delta W_{dc}$

$= 0.119 + (-0.0053)$

$= 0.113$

$W_{d0} = W_{d0} + \Delta W_{d0}$

$= 0.134 + (-0.01)$

$= 0.124$

$\delta_c = O_c * (1 - O_c) * (W_{dc} * \delta_d)$

$= 0.55(1 - 0.55)(0.113 * (-0.136))$

$= -0.0038$

$\Delta W_{ca} = \eta * \delta_c * M_a + \alpha * \Delta W_{ca}(old)$

$= 0.3 * (-0.0038) * 0 + 0.9 * 0.101$

$= 0.0009$

$\& \Delta W_{c0} = 0$

$\Delta w_{cb} = -0.001$

$\therefore w_{co} = w_{co} + \Delta w_{co} = 0.101 + 0 = 0.101$

$w_{ca} = w_{ca} + \Delta w_{ca} = 0.101 + 0.0009 = 0.102$

$w_{cb} = w_{cb} = \Delta w_{cb} = 0.100 + (-0.001) = 0.099$

<u>Final values:-</u>

$w_{co} = 0.101$

$w_{ca} = 0.102$

$w_{cb} = 0.099$

$w_{do} = 0.124$

$w_{dc} = \underline{0.113}$