Question 1 (10 points):
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.

$P(c) = 0.008$, $P(-|c) = 0.02$
$P(-c) = 0.992$, $P(-|-c) = 0.97$
$P(+|c) = 0.98$, $P(+|-c) = 0.03$

---

* first test:

$$p(c|+) = \frac{p(+|c)\,p(c)}{p(+|c)\,p(c) + p(+|-c)\,p(-c)}$$

$$= \frac{0.98 \times 0.008}{0.98 \times 0.008 + 0.03 \times 0.992}$$

$$= 0.208 \simeq 0.21$$

* Second test:

Since the patient is already tested positive

$p(c) = 0.21$
$p(-c) = 0.79$

$\therefore$ for second test:

$$p(c|+) = \frac{p(+|c)\,p(c)}{p(+|c)\,p(c) + p(+|-c)\,p(-c)}$$

$$= \frac{0.98 \times 0.21}{0.98 \times 0.21 + 0.03 \times 0.79}$$

$$\simeq 0.90$$

$$\therefore\ p(-c|+) = 1 - p(c|+)$$
$$= 1 - 0.9$$
$$= \underline{\underline{0.1}}$$

Question 2 (5 points):

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 $Error_D(h)$?

$$error_S(h) = \frac{r}{n} = \frac{20}{100} = 0.20$$

95% confidence interval for $error_D(h)$

$$error_S(h) \pm 1.96 \sqrt{\frac{error_S(h)\,(1 - error_S(h))}{100}}$$

$$= 0.2 \pm 1.96 \times 0.04$$

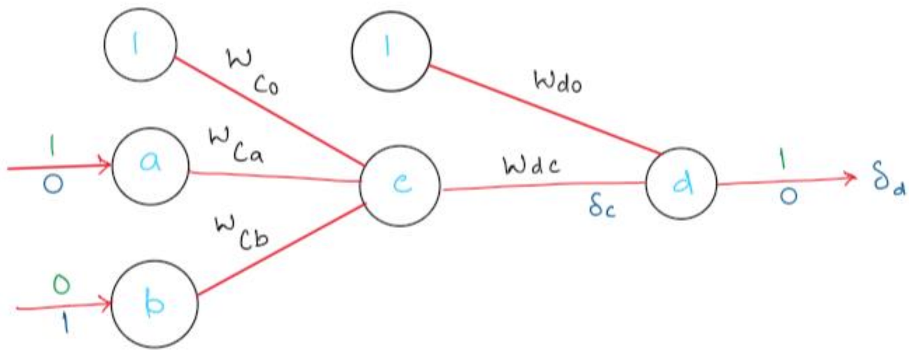
$$= 0.2 \pm 0.0784$$

$\therefore$ $error_D(h)$ lies between 0.1216 and 0.2784

Question 3 (15 points):

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



for first iteration:

$$
\begin{aligned}
O_c &= g(1\times W_{co} + a\times W_{ca} + b\times W_{cb}) \\
&= g(1\times 0.1 + 1\times 0.1 + 0) \\
&= g(0.1+0.1) \\
&= g(0.2) = \frac{1}{1+e^{-0.2}} \\
&= 0.55
\end{aligned}
$$

$$
\begin{aligned}
O_d &= g(1\times W_{do} + 0.1\times O_c) \\
&= g(0.1 + 0.05) \\
&= g(0.155) \\
&= \frac{1}{1+e^{-0.155}} = 0.539
\end{aligned}
$$

$$
\begin{aligned}
\delta_d &= O_d(t_d - O_d)(1-O_d) \\
&= 0.539(1-0.539)(1-0.539) \\
&= 0.114
\end{aligned}
$$

$$
\begin{aligned}
\delta_c &= O_c(1-O_c)\sum W_{dc}\delta_d \\
&= 0.55(1-0.55)\times 0.1\times 0.114 \\
&= 2.82\times 10^{-3} = 0.00282
\end{aligned}
$$

$\Delta w_{co} = \eta \delta_c \times 1 = 2.82 \times 10^{-3} \times 0.3 \times 1$
$= 8.46 \times 10^{-4}$

$\Delta W_{ca} = \eta \delta_c \times a = 2.82 \times 10^{-3} \times 0.3 \times 1$
$= 8.46 \times 10^{-4}$

$\Delta W_{cb} = \eta \delta_c \times b = 0$

$\Delta W_{do} = \eta \delta_d \times 1 = 0.3 \times 0.114 \times 1$
$= 0.0342$

$\Delta W_{dc} = \eta \delta_d \times O_c = 0.3 \times 0.114 \times 0.55$
$= 0.01881$

$\therefore$ new weights after iteration 1,

$W_{co} = W_{co} + \Delta W_{co}$
$= 0.1 + 8.46 \times 10^{-4}$
$= 0.1008$

$W_{ca} = W_{ca} + \Delta W_{ca}$
$= 0.1008$

$W_{cb} = 0.1 + 0 = 0.1$

$W_{do} = W_{do} + \Delta W_{do}$
$= 0.1342$

$W_{dc} = 0.1 + 0.01881$
$= 0.11881$

* for second iteration:

$O_c = g(1 \times W_{co} + a \times W_{ca} + b \times W_{cb})$

$= g(0.2008) = 0.55$

$O_d = g(1 \times W_{do} + O_c \times W_{dc})$

$= g(0.1995) = 0.5497$

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

$= (0 - 0.5497)(1 - 0.5497) 0.5497$

$= -0.136$

$\delta_c = O_c (1 - O_c) \Sigma W \delta_d$

$= 0.55 (1 - 0.55) \times 0.11881 \times -0.136$

$= -0.004$

$\Delta W_{do} = \alpha \Delta W_{do}(n-1) + \eta \delta_d \times 1$

$= 0.9 \times 0.0342 + 0.3 \times -0.136 \times 1$

$= -0.01$

$\Delta W_{dc} = \alpha \Delta W_{dc}(n-1) + \eta \delta_d \times O_c$

$= 0.9 \times 0.01881 + 0.3 \times -0.136 \times 0.55$

$= -0.0055$

$\Delta W_{co} = \alpha \Delta W_{co}(n-1) + \eta \delta_c \times 1$

$= -4.386 \times 10^{-4}$

$\Delta W_{ca} = \alpha \Delta W_{ca}(n-1) + \eta \delta_c \times a$

$= 0.9 \times 8.46 \times 10^{-4}$

$= 7.614 \times 10^{-4}$

$\Delta W_{cb} = \alpha \Delta W_{cb}(n-1) + \eta \delta_c \times b$

$= 0 + 0.3 \times -0.004 \times 1$

$= -1.2 \times 10^{-3}$

$\therefore$ new weights after iteration 2,

$$w_{do} = w_{do} + \Delta w_{do}$$
$$= 0.1342 + (-0.01) \quad = 0.1242$$

$$w_{dc} = 0.11881 + (-0.0055) \quad = 0.11331$$

$$w_{co} = 0.1008 + (-4.386 \times 10^{-4}) = 0.1003$$

$$w_{ca} = 0.1008 + (7.614 \times 10^{-4}) = 0.1015$$

$$w_{cb} = 0.1 + (-1.2 \times 10^{-3}) \quad = 0.0988$$