(a) keep all

$E[L(y=keep\_t)]$

$= P(t=Non\ Spam) \cdot 0 + P(t=Spam) \cdot 1$

$= 0.9 \times 0 + 0.1 \times 1$

$= 0.1$

remove all

$E[L(y=remove\_t)]$

$= P(t=Non\ Spam) \cdot 100 + P(t=Spam) \cdot 0$

$= 0.9 \times 100 + 0.1 \times 0$

$= 90$

(b) Let $p' = p(t=\text{spam}|x)$

$E[L(y=\text{keep}, t)]$

$= p' \times 1 + (1-p') 100 = p'$

$E(L(y=\text{remove}, t)] = (1-p') 100 + (1-p') \cdot 0 = 100 - 100p'$

$E(L(y=\text{keep}, t)] > E(L(y=\text{remove}, t)]$

$\Rightarrow \quad p' > 100 - 100p'$

$\Rightarrow \quad p' > \frac{100}{101}$

Thus, if $p(t=\text{spam}|x) > \frac{100}{101}$

remove all mail

else, i.e. $p(t=\text{spam}|x) \leq \frac{100}{101}$

keep all mail

(c) by (b) we know for a given feature x

if $P(t=\text{spam}|x) > \frac{100}{101}$

remove else: keep

$$P(t=\text{spam}|x_1=0, x_2=0) = \frac{P(t=\text{spam}) \cdot P(x_1, x_2=0 | t=\text{spam})}{P(x_1=0, x_2=0)}$$

$$= \frac{0.1 \times 0.4}{P(x_1, x_2 | t=\text{spam}) \cdot P(\text{spam}) + P(x_1, x_2 | t=\text{Nonspam})} = \frac{0.1 \times 0.4}{0.4 \times 0.1 + 0.8 \times 0.8}$$

$$\approx 0.0426$$

$$P(t=\text{spam}|x_1=0, x_2=1) = \frac{0.1 \times 0.3}{0.3 \times 0.1 + 0.001 \times 0.8} = \frac{300}{308} \approx 0.971$$

$$P(t=\text{spam}|x_1=1, x_2=0) = \frac{0.1 \times 0.2}{0.2 \times 0.1 + 0.001 \times 0.8} = \frac{200}{208} \approx 0.957$$

$$P(t=\text{spam}|x_1=1, x_2=1) = \frac{0.1 \times 0.1}{0.1 \times 0.1} = 1$$

Since only $P(t=\text{spam}|x_1=1, x_2=1) > \frac{100}{101}$

remove mail only if $x_1, x_2 = 1$ (both words appears)

else keep

(d) $E[L(y,t)] =$

$100 \times P(t = \text{Nonspam} \mid x_1, x_2 = (1,1)) \times P(x_1, x_2 = (1,1))$

$+ P(t = \text{spam} \mid x_1 = 0, x_2 = 0) \times P(x_1, x_2 = 0\_0)$

$+ P(t = \text{spam} \mid x_1 = 1, x_2 = 0) \times P(x_1, x_2 = 1\_0)$

$+ P(t = \text{spam} \mid x_1 = 0, x_2 = 1) \times P(x_1, x_2 = 0\_1)$

$= 100 \times P(t = \overbrace{\text{Spam}}^{\text{Non}}) \cdot P(x_1, x_2 = (1,1) \mid t = \text{Nonspam})$

$+ P(t = \text{spam}) \cdot P(x_1, x_2 = (0,0)\_(1,0)\_(0,1) \mid t = \text{spam})$

$= 0.1 \times (0.2 + 0.3 + 0.4)$

$= 0.09$

The following question please check

Q2 Q3 Q4 write up