Q1.
Proof:

$$p(y=1|x) = \frac{p(x|y=1) * p(y=1)}{p(x)} = \frac{p(x|y=1) * p(y=1)}{p(x|y=1) * p(y=1) + p(x|y=0) * p(y=0)}$$

$$= \frac{1}{1 + \frac{p(x|y=0) * p(y=0)}{p(x|y=1) * p(y=1)}} = \frac{1}{1 + \exp\left(\ln\left(\frac{p(x|y=0) * p(y=0)}{p(x|y=1) * p(y=1)}\right)\right)}$$

$$= \frac{1}{1 + \exp\left(\ln\left(\frac{p(x|y=0)}{p(x|y=1)}\right) + \ln\left(\frac{p(y=0)}{p(y=1)}\right)\right)} = \frac{1}{1 + \exp\left(\ln\left(\frac{p(x|y=0)}{p(x|y=1)}\right) + \ln\left(\frac{1-\alpha}{\alpha}\right)\right)}$$

$$= \frac{1}{1 + \exp\left(\sum_{i=1}^{D} \ln\left(\frac{p(x_i|y=0)}{p(x_i|y=1)}\right) + \ln\left(\frac{1-\alpha}{\alpha}\right)\right)} \ (by\ conditionally\ independent)$$

$$= \frac{1}{1 + \exp\left(\sum_{i=1}^{D} \ln\left(\frac{\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu_{i0})^2\right)}{\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu_{i1})^2\right)}\right) + \ln\left(\frac{1-\alpha}{\alpha}\right)\right)}$$

$$= \frac{1}{1 + \exp\left(\sum_{i=1}^{D} \ln\left(\frac{\exp\left(-\frac{1}{2\sigma^2}(x_i - \mu_{i0})^2\right)}{\exp\left(-\frac{1}{2\sigma^2}(x_i - \mu_{i1})^2\right)}\right) + \ln\left(\frac{1-\alpha}{\alpha}\right)\right)}$$

$$= \frac{1}{1 + \exp\left(\sum_{i=1}^{D} \ln\left(\exp\left(\frac{1}{2\sigma^2}(x_i - \mu_{i1})^2 - \frac{1}{2\sigma^2}(x_i - \mu_{i0})^2\right)\right) + \ln\left(\frac{1-\alpha}{\alpha}\right)\right)}$$

$$= \frac{1}{1 + \exp\left(\sum_{i=1}^{D} \frac{(x_i - \mu_{i1})^2 - (x_i - \mu_{i0})^2}{2\sigma^2} + \ln\left(\frac{1-\alpha}{\alpha}\right)\right)}$$

$$= \frac{1}{1 + \exp\left(\sum_{i=1}^{D} \frac{(\mu_{i1}{}^2 - \mu_{i0}{}^2) + 2x_i(\mu_{i0} - \mu_{i1})}{2\sigma^2} + \ln\left(\frac{1-\alpha}{\alpha}\right)\right)}$$

$$= \frac{1}{1 + \exp\left(\sum_{i=1}^{D} \frac{(\mu_{i0} - \mu_{i1})}{\sigma^2} x_i + \ln\left(\frac{1-\alpha}{\alpha}\right) + \sum_{i=1}^{D} \frac{(\mu_{i1}{}^2 - \mu_{i0}{}^2)}{2\sigma^2}\right)}$$

From the equation above, it shows that p(y = 1|x) takes the form of a logistic function:

p(y = 1|x) = $\sigma(w^T x + b) = \frac{1}{1+\exp(-\sum_{i=1}^{D} w_i x_i - b)}$

where $w_i = -\frac{(\mu_{i0}-\mu_{i1})}{\sigma^2}$ b = $-\left(\ln\left(\frac{1-\alpha}{\alpha}\right) + \sum_{i=1}^{D} \frac{(\mu_{i1}{}^2 - \mu_{i0}{}^2)}{2\sigma^2}\right)$

Q2.

$$E(w,b) = -ln\left(\prod_{i=1}^{N} p(y^i = 0|x^i)^{1-y^i} * p(y^i = 1|x^i)^{y^i}\right)$$

$$= -\ln\left(\prod_{i=1}^{N} p(y^i = 0|x^i)^{1-y^i}\right) - \ln\left(\prod_{i=1}^{N} p(y^i = 1|x^i)^{y^i}\right)$$

$$= -(1-y^i)\ln\left(\prod_{i=1}^{N} p(y^i = 0|x^i)\right) - y^i * \ln\left(\prod_{i=1}^{N} p(y^i = 1|x^i)\right)$$

$$= \sum_{i=0}^{N} -(1-y^i)\ln\left(p(y^i = 0|x^i)\right) - \sum_{i=0}^{N} y^i \ln\left(p(y^i = 1|x^i)\right)$$

$$let\ z^i = w^T x_i + b, then$$

$$= \sum_{i=0}^{N} (1-y^i)\ln(1+\exp(-z^i)) + \sum_{i=0}^{N} (1-y^i)\ln(\exp(z^i)) + \sum_{i=0}^{N} y^i \ln(1+\exp(-z^i))$$

$$= \sum_{i=0}^{N} \ln(1+\exp(-z^i)) + \sum_{i=0}^{N} (1-y^i) z^i$$

Then derive expressions for the derivatives of E with respect to each of the model parameters:

$$\frac{E(w,b)}{\partial b} = -\sum_{i=0}^{N} \frac{\exp\left(-(wx^i + b)\right)}{1+\exp(-(wx^i+b))} + \sum_{i=0}^{N} (1-y^i) = -\sum_{i=0}^{N} \frac{\exp(-z^i)}{1+\exp(-z^i)} + \sum_{i=0}^{N} (1-y^i)$$

$$\frac{E(w,b)}{\partial w_j} = -\sum_{i=0}^{N} x_j{}^i \frac{\exp\left(-(wx^i + b)\right)}{1+\exp(-(wx^i+b))} + \sum_{i=0}^{N} (1-y^i) x_j{}^i$$

$$= x_j{}^i\left(-\sum_{i=0}^{N} \frac{\exp(-z^i)}{1+\exp(-z^i)} + \sum_{i=0}^{N} (1-y^i)\right)$$

Q3.

$$p(w,b|D) \propto p(D|w,b)p(w,b)$$

$$p(w,b) = \prod_{i=1}^{D} N\left(w_i \middle| 0, {}^{1}/_{\lambda}\right) \, N\left(b \middle| 0, {}^{1}/_{\lambda}\right)$$

$$let\ z = wx + b$$

$$p(D|w,b) = \prod_{i=1}^{N} \frac{1}{(1+\exp(-z))}^{y^i} \frac{\exp(-z)}{(1+\exp(-z))}^{1-y^i}$$

$$p(w,b|D) \propto \left(\prod_{i=1}^{N} \frac{1}{(1+\exp(-z))}^{y^i} \frac{\exp(-z)}{(1+\exp(-z))}^{1-y^i}\right) \prod_{i=1}^{D} N\left(w_i \middle| 0, {}^{1}/_{\lambda}\right) \, N\left(b \middle| 0, {}^{1}/_{\lambda}\right)$$

$$L(w,b)$$

$$= -\ln\left(\left(\prod_{i=1}^{N} \frac{1}{(1+exp(-z))}^{y^i} \frac{exp(-z)}{(1+exp(-z))}^{1-y^i}\right)\left(\prod_{i=1}^{D} N\left(w_i \middle| 0, {}^{1}/_{\lambda}\right) \, N\left(b \middle| 0, {}^{1}/_{\lambda}\right)\right)\right)$$

$$= -\ln\left(\prod_{i=1}^{N} \frac{1}{(1+exp(-z))}^{y^i} \frac{exp(-z)}{(1+exp(-z))}^{1-y^i}\right) - \ln\left(\prod_{i=1}^{D} N\left(w_i \middle| 0, {}^{1}/_{\lambda}\right) \, N\left(b \middle| 0, {}^{1}/_{\lambda}\right)\right)$$

$$= E(w,b) - \sum_{i=0}^{D} \ln\left(N\left(w_i \middle| 0, {}^{1}/_{\lambda}\right)\right) - \ln\left(N\left(b \middle| 0, {}^{1}/_{\lambda}\right)\right))$$

$$= E(w,b) - \sum_{i=0}^{D} -\frac{{w_i}^2 * \lambda}{2} - D * \ln\left(\frac{\sqrt{\lambda}}{\sqrt{2\pi}}\right) - \left(-\frac{b^2 * \lambda}{2}\right) - \ln\left(\frac{\sqrt{\lambda}}{\sqrt{2\pi}}\right)$$

$$= E(w,b) + \frac{\lambda}{2}\sum_{i=0}^{D} {w_i}^2 + \left(\frac{\lambda}{2}\right) b^2 - (D+1)\ln\left(\frac{\sqrt{\lambda}}{\sqrt{2\pi}}\right)$$

$$= E(w,b) + \frac{\lambda}{2}\sum_{i=0}^{D} {w_i}^2 + \frac{\lambda}{2} b^2 + c(\lambda)$$

Then derive expressions for the derivatives of L with respect to each of the model parameters:

$$\frac{L(w,b)}{\partial b} = -\sum_{i=0}^{N} \frac{\exp\left(-z^i\right)}{1+\exp\left(-z^i\right)} + \sum_{i=0}^{N}\left(1-y^i\right) + \lambda b$$

$$\frac{L(w,b)}{\partial w_j} = -\sum_{i=0}^{N} x_j{}^i \frac{\exp\left(-\left(wx^i + b\right)\right)}{1 + \exp\left(-\left(wx^i + b\right)\right)} + \sum_{i=0}^{N} \left(1 - y^i\right) x_j{}^i + \lambda \sum_{i=0}^{D} w_i$$

$$= x_j{}^i \left(-\sum_{i=0}^{N} \frac{\exp\left(-z^i\right)}{1 + \exp\left(-z^i\right)} + \sum_{i=0}^{N} \left(1 - y^i\right) \right) + \lambda \sum_{i=0}^{D} w_i$$

Q4.