Solutions to the ECE521 Midterm on 16 Feb 2017

Part I

1. F
2. T
3. F
4. T
5. F
6. T
7. F
8. F

Part II

1. d
2. d
3. b
4. c
5. b
6. b
7. b
8. d

Factory resistors question: Letting s represent a pass (success) and g represent the fact that a resistor is good, Bayes' theorem gives:

$$\mathrm{p}(g|s) = \frac{\mathrm{p}(s|g)\mathrm{p}(g)}{\mathrm{p}(s)} = \frac{0.8 \times 0.9}{0.72 + 0.03} = 0.96$$

Letting s represent two out of three passes,

$$\mathrm{p}(g|s) = \frac{\mathrm{p}(s|g)\mathrm{p}(g)}{\mathrm{p}(s)} = \frac{(0.8^2)(0.2) \times 0.9}{(0.8^2)(0.2) \times 0.9 + (0.3^2)(0.7) \times 0.1} \approx 0.95$$

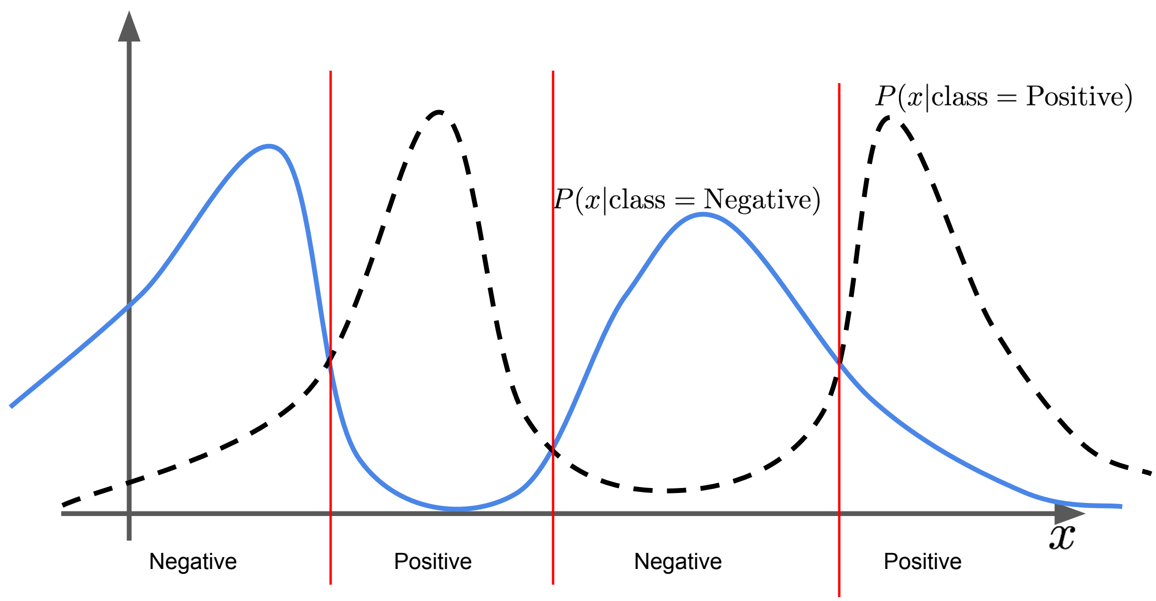
Part III

1.

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{2} \sum_{m=1}^{M} \left\{ \log \left[\frac{1}{1+\exp(-\mathbf{w}^T \mathbf{x}_m)} \right] (t_m + 1) + \log \left[\frac{1}{1+\exp(\mathbf{w}^T \mathbf{x}_m)} \right] (1 - t_m) \right\} \\
&= -\frac{1}{2} \sum_{m=1}^{M} 2 \log \left[\frac{1}{1+\exp(-t_m \mathbf{w}^T \mathbf{x}_m)} \right] \\
&= \sum_{m=1}^{M} \log \left[1 + \exp(-t_m \mathbf{w}^T \mathbf{x}_m) \right]
\end{aligned}$$

2. It has greater numerical stability, owing to the addends' lower dynamic range. (As an aside, note that Tensorflow takes advantage of this when implementing *tf.nn.sigmoid_cross_entropy_with_logits*.)

3.



Part IV

1. Yes it does. $\frac{\partial \tanh(z)}{\partial z} = 1 - \tanh^2(z)$

2. (a) $\mathbf{w} = (1, 1, 1)$. Many other answers possible.

(b) $\mathbf{w} = (1, 1, -2)$. Many other answers possible.

(c) $\{t\} = \{1, 1, 1, 0\}$ or $\{0, 0, 0, 1\}$.

3. (a) $\hat{y} = W_4\phi(\mathbf{z}_1) + b_4 + W_3\phi(\mathbf{z}_2) + b_3$

b) $\frac{\partial \mathcal{L}}{\partial W_4} = \frac{\partial \mathcal{L}}{\partial \hat{y}}\phi(\mathbf{z}_1)^T, \quad \frac{\partial \mathcal{L}}{\partial b_4} = \frac{\partial \mathcal{L}}{\partial \hat{y}}$

c) $\frac{\mathcal{L}}{\mathbf{z}_1} = \frac{\partial \phi(\mathbf{z}_1)}{\partial \mathbf{z}_1}(W_4^T \frac{\partial \mathcal{L}}{\partial \hat{y}} + W_2^T \frac{\partial \mathcal{L}}{\partial \mathbf{z}_2})$

Part V

[a] For (i), example suitable basis functions are $\phi_1(x) = x$ and $\phi_2(x) = x^2$. For (ii), example suitable basis functions are

$$\phi_1(x) = \cos(2\pi x/L) \quad \phi_2(x) = \sin(2\pi x/L)$$

if the periodic function is sinusoidal; expressions of more frequencies are fine as well.

[b] Yes. Suppose you combine the above two sets of basis functions (four functions total). Provided you have enough data, then the coefficients of the non-relevant basis functions for that source should drop down close to zero.