ECE/CS 559 - Fall 2017 - Midterm Part #2.

Full Name: **ID Number:**

Q3 (25 pts). Let u be the step activation function with $u(x) = 1$ if $x \geq 0$, and $u(x) = 0$, otherwise. Consider the perceptron $y = u(w_0 + w_1 x_1 + w_2 x_2)$, where w_1 and w_2 are the weights for inputs x_1 and x_2, respectively, w_0 is the perceptron bias, and y is the perceptron output. Let $\mathcal{C}_0 = \left\{ \begin{bmatrix} 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \end{bmatrix} \right\}$, and $\mathcal{C}_1 = \left\{ \begin{bmatrix} 1 & 1 \end{bmatrix} \right\}$. The desired output for class $\mathcal{C}_0$ is 0, and the desired output for class $\mathcal{C}_1$ is 1. Correspondingly, let $d(\mathbf{x}) = 0$ if $\mathbf{x} \in \mathcal{C}_0$, and otherwise, let $d(\mathbf{x}) = 1$ if $\mathbf{x} \in \mathcal{C}_1$.

(a) **(8 pts)** If possible, find w_0, w_1, w_2 that can separate $\mathcal{C}_0$ and $\mathcal{C}_1$ (i.e., provide the desired output for all 4 possible input vectors). Otherwise, prove that no choice of weights can separate the two classes.

(b) **(10 pts)** Recall that the perceptron training algorithm relies on the update $\mathbf{w} \leftarrow \mathbf{w} + \eta(d(\mathbf{x}) - y) \begin{bmatrix} 1 & \mathbf{x} \end{bmatrix}$, where $\mathbf{w} = \begin{bmatrix} w_0 & w_1 & w_2 \end{bmatrix}$ is the weight vector. Let $\eta = 1$ and the initial weight vector be given by $\mathbf{w} = \begin{bmatrix} -0.5 & 1 & 0 \end{bmatrix}$. Calculate the updated weights after two epochs of training.

.

(c) **(7 pts)** Will the weights provided by the algorithm (as setup in (b)) eventually converge after a sufficiently larger number of epochs? Justify your answer.

Q4 (15 pts). Let

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \phi\left(\begin{bmatrix} w_{10} & w_{11} & w_{12} \\ w_{20} & w_{21} & w_{22} \end{bmatrix}\begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}\right), \text{ and } \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \phi\left(\begin{bmatrix} u_{10} & u_{11} & u_{12} \\ u_{20} & u_{21} & u_{22} \end{bmatrix}\begin{bmatrix} 1 \\ y_1 \\ y_2 \end{bmatrix}\right) \tag{1}$$

with the understanding that the activation function ϕ is applied component-wise. These equations define a two-layer neural network with 2 input nodes, 2 neurons in the hidden layer, and 2 output nodes.

(a) **(5 pts)** Draw the block diagram of the neural network with all inputs, outputs, weights labeled.

(b) **(10 pts)** Let $E = (d_1 - z_1)^2 + (d_2 - z_2)^4 + u_{22}^2$. Write down the expressions for $\frac{\partial E}{\partial w_{10}}$ and $\frac{\partial E}{\partial u_{22}}$. You may use the backpropagation algorithm. Your expressions may contain intermediate variables that you shall clearly define on the feedforward/feedback graphs.

Q5 (10 pts). Consider the activation function $\phi(v) = \frac{v}{1+|v|}$ defined for all real numbers.

(a) **(5 pts)** Find $\phi'(v) = \frac{\partial \phi}{\partial v}$.

(b) **(5 pts)** Express $\phi'(v)$ in terms of $\phi(v)$ only.