Instructor: Vishnu Boddeti
CSE 891-001: Deep Learning
Homework 1
September, 20

1 Hard-Coding a Multilayer Perceptron (2pts): There are many possible solutions, here is one.

$$\boldsymbol{W}^{(1)} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \boldsymbol{b}^{(1)} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \boldsymbol{W}^{(2)} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} b^{(2)} = -2.5 \tag{1}$$

2 Sparsifying Activation Function (3pts):

- $\frac{\partial L}{\partial w_1} = 0$ - Yes
- $\frac{\partial L}{\partial w_2} = 0$ - Yes
- $\frac{\partial L}{\partial w_3} = 0$ - No

3 Universal Approximation Theorem (5pts):

- $W_0 = \begin{bmatrix} -1 & 1 \end{bmatrix}^T$, $b_0 = \begin{bmatrix} b & -a \end{bmatrix}^T$, $W_1 = \begin{bmatrix} h & h \end{bmatrix}^T$, $b_1 = -h$