Boston University
Department of Electrical and Computer Engineering

***ENG EC 414 Introduction to Machine Learning***

**HW 10**



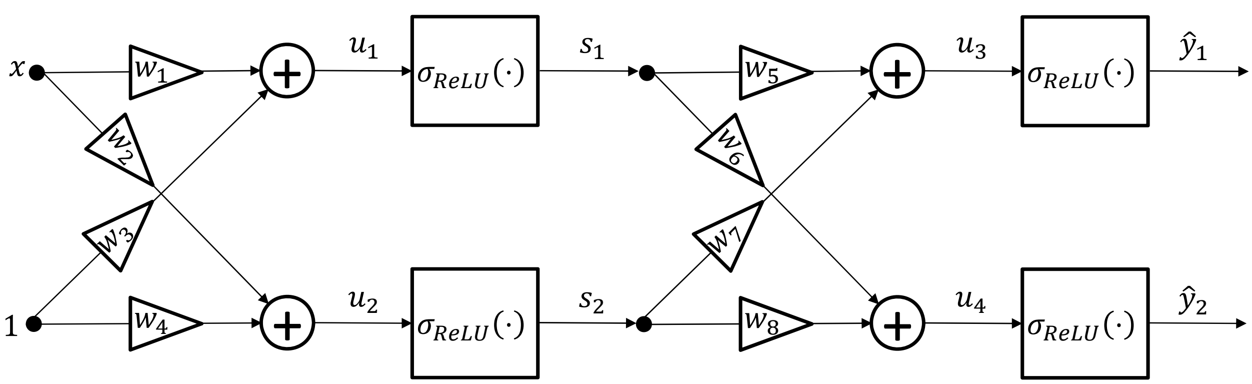
**Issued:** Fri 20 Nov 2020 **Due:** 10:00am Fri 4 Dec 2020 in Gradescope

**Important:** Before you proceed, please read the documents pertaining to *Homework formatting and submission guidelines* in the Homeworks section of Blackboard. **In particular, for computer assignments you are prohibited from using any online code or built-in MATLAB functions except as indicated in the problem or skeleton code (when provided).**

**Important:** To obtain full grade, please clearly motivate all your answers.

**Note:** Problem difficulty = number of coffee cups ☕

**Problem 10.1** [13pts] *(Backprop by Hand)* A two-layer feed-forward neural network with ReLU activation function, $\sigma_{ReLU}(t) = \max\{0, t\}$, is shown in the figure below. Let $\ell(\hat{y}_1, \hat{y}_2, y) = (\hat{y}_1 + \hat{y}_2 - y)^2$ be the loss function, $(x = 0.5, y = 1)$ be a training sample, and initial weights $w_1 = -1, w_2 = 1, w_3 = 1, w_4 = 1, w_5 = -1, w_6 = 1, w_7 = 2, w_8 = 2$. Note that the second input is fixed to 1, it is the bias term.




(a) [3pts] Compute the values of $u_1, u_2, s_1, s_2, u_3, u_4, \hat{y}_1, \hat{y}_2$, and $\ell(\hat{y}_1, \hat{y}_2, y)$ in the first forward pass iteration of the Backprop algorithm.

(b) [10pts] Compute the values of the partial derivatives of the loss with respect to $\hat{y}_1, \hat{y}_2, u_3, u_4, w_5, w_6, w_7, w_8, s_1, s_2$ in the first backward pass iteration of the Backprop algorithm.

**Problem 10.2** [13pts] *(Function Approximation with Neural Network)* Let

$$h(x) = \begin{cases} 2x + 2 & -1 \le x \le 0 \\ -x + 2 & 0 \le x \le 2 \\ 0 & \text{else} \end{cases}$$

Let $\sigma_{ReLU}(t) = \max\{0, t\}$ be the Rectifier Linear Unit activation function. Find values of $(\beta_1, \gamma_1)$, $(\alpha_2, \beta_2, \gamma_2)$, and $(\beta_3, \gamma_3)$ such that for all $x$,

$$h(x) = \underbrace{\sigma_{ReLU}\left(\beta_1 + \gamma_1 x\right)}_{h_1(x)} + \underbrace{\alpha_2 \cdot \sigma_{ReLU}\left(\beta_2 + \gamma_2 x\right)}_{h_2(x)} + \underbrace{\sigma_{ReLU}\left(\beta_3 + \gamma_3 x\right)}_{h_3(x)}$$

Sketch the graphs of $h_1(x)$, $h_2(x)$, $h_3(x)$, and $h(x)$ and properly label axes and key points.