

1. Perform one iteration of gradient descent on w_{12} . (The weight going to h_1 from x_2 .) Use input vector $x = (1, 1, 1)$, ReLU as our nonlinearity, $L = (y - \hat{y})^2$ as our loss, $y = 1$ as our current label and $\eta = 0.1$ as our learning rate. (10 points)

