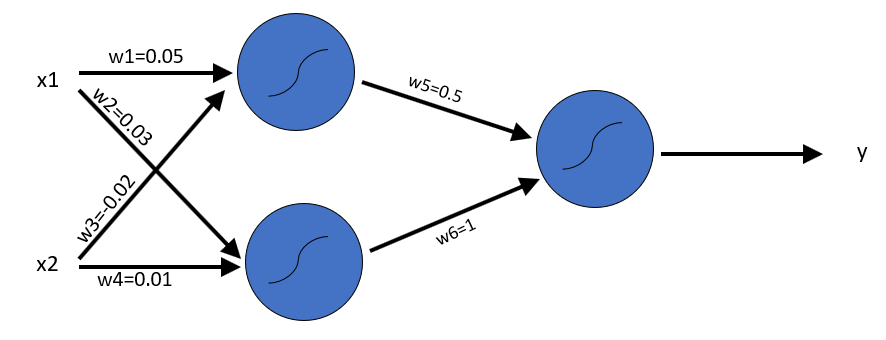
## ITC 6230: Deep Learning Homework Assignment #2

**DUE-DATE: Sunday, FEBRUARY 18, 2024, 23:59.**

Consider the following neural network:



Consider the following dataset:

|  |  |  |
| --- | --- | --- |
| **x1** | **x2** | **y (label)** |
| -1 | 1 | 0 |
| 0.5 | 0.5 | 1 |
| 1 | -1 | 1 |
| 0 | 0.5 | 0 |

**Question 1**

**a. (30 pts):** Using the Back-Propagation algorithm, compute the gradient (a vector of 6 components) of the loss function where the function is the output of the neural network when the input is the vector , and **all** nodes in the network use as activation function the ReLU function .

**b. (10 pts)** Check your answer in step a. by computing the difference quotient where and (the unit vector in the first dimension). What is the different between your computed value of and the difference quotient for the three proposed values of ?

**Question 2 (25 pts):** Using a step-size , use the GD equation:

to compute a new set of weights given the gradient computed at Question 1.a. Now, compute the value of the loss function of the neural network in the figure above, evaluated at the point . How does the new loss value compare to the original value ?

**Question 3 (35 pts):** For the network in the figure above, compute the gradient function of the single-instance mini-batch:

|  |  |  |
| --- | --- | --- |
| **x1** | **x2** | **y (label)** |
| -1 | 1 | 0 |

The activation function of the two hidden nodes is the rectified linear unit (ReLU): and the activation function for the output node is the hyperbolic tangent function . The loss function whose gradient we seek is again the same as in Question 1 (mean square error)