## **Partial Derivatives with respect to a**

### **Part 1**

How do we compute partial derivatives

1. The following neural network will be used to demonstrate the calculations
2. Here are the parameters of the network

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | 0.1 | 0.3 | 0.8 | -0.4 |  |
| W1= | -0.3 | -0.2 | 0.5 | 0.5 |
|  | -0.3 | 0 | 0.5 | 0.4 |
|  | 0.2 | 0.5 | -0.9 | 0.7 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| W2 = |  | 0.5 | 0.8 | 0.2 | 0.4 |  |
| 0.5 | 0.2 | 0.3 | -0.5 |

* 1. x = [ 2 5 3 3 ] true distribution y = [ 1 0 ]

1. Now, we want to find the partial derivative w.r.t w212 as highlighted in the figure
2. We will solve the above equation sequentially
   1. Consider square error loss function L
      1. Here, the y2 terms get cancelled, leaving