

Pierre Tobgu

Problem 2 CS 4662 Final

$$\text{CIN} = 306435914$$

$$d1=3, d2=0, d3=6, d4=4, d5=3, d6=5$$

$$d7=9, d8=1, d9=4$$

$$(X, y) = ((1, 1), 0) \quad \text{perform 1st backpropagation}$$

$\begin{matrix} & \times & \\ \uparrow & & \uparrow \\ x_1 & x_2 & \end{matrix}$ \uparrow output = y

$$\text{learning rate } \alpha = 0.1 \quad \text{bias} = -1$$

$$g(z) = \frac{1}{1 + e^{-z}}$$

$$* a_1(1) = g((3*1) + (1*6) + (1*3)) = g(-3+6+3)$$

$$g(6) = \frac{1}{1+e^{-6}} = \frac{1}{1+e^{-6}} = \frac{1}{1+(2.176752136 \times 10^{-3})}$$

$$a_1(1) = 0.9975$$

$$* a_2(1) = g((0*-1) + (4*1) + (5*1)) = g(4+5)$$

$$g(9) = \frac{1}{1+e^{-9}} = 0.99988$$

$$g((-1*9) + (0.9975*1) + (0.99988*4)) = g(-4.003)$$

$$\underline{0 = -4.003}$$

y is actual output
 \hat{y} output generated

$$\delta_{(1)} = (y - \hat{y}) \cdot (1 - \hat{y})$$

$$(4 - (-4.003)) * -4.003 (1 - (-4.003))$$

$$4.003 * (-4.003) * (5.003) = -80.168$$

$$\delta_1(3) = \underline{-80.168}$$

$$\delta_1(2) = [w_{11}(3) \delta_1(3)] a_{11}(1) (1 - a_{11}(1))$$

$$= [1 * -80.168] 0.9975 (1 - (0.9975))$$

$$\delta_1(2) = -0.1999$$

$$\delta_2(2) = [w_{12}(3) \delta_1(3)] a_{21}(1) (1 - a_{21}(1))$$

$$= [4 * -80.168] 0.99988 (1 - 0.99988)$$

$$\delta_2(2) = -0.0385$$

$$\delta_3(1) = \overset{\delta_2}{[9x - 80.168]} x (-1) x (1 - (-1)) =$$

$$= 1,443.024$$

$$\begin{aligned}
 & \begin{matrix} 18 \\ \downarrow \end{matrix} \\
 w_{11}(2)_{\text{new}} &= w_{11}(2) + \alpha \delta a_1(1) \\
 &= (1) + (0.1)(-80.168)(0.9975) \\
 w_{11}(2)_{\text{new}} &= -6.996758 \approx \textcircled{-7}
 \end{aligned}$$

$$\begin{aligned}
 w_{12}(2)_{\text{new}} &= w_{12}(2) + \alpha \delta a_1(1) \\
 &= (4) + (0.1)(-80.168)(0.9975) \\
 w_{12}(2)_{\text{new}} &= -3.996758 \approx \textcircled{-4}
 \end{aligned}$$

$$\begin{aligned}
 w_{13}(2)_{\text{new}} &= w_{13}(2) + \alpha \delta a_1(1) \\
 &= 9 + -7.996758 \\
 w_{13}(2)_{\text{new}} &= 1.003242 \approx \textcircled{1}
 \end{aligned}$$

$$\begin{aligned}
 w_{21}(1)_{\text{new}} &= w_{21}(1) + \alpha \delta_2(1) [\text{bias}] \\
 &= (0) + (0.1)(-0.0385)(-1) \\
 &= \boxed{3.85 \times 10^{-3}}
 \end{aligned}$$

$$\begin{aligned}
 w_{22}(1)_{\text{new}} &= w_{22}(1) + \alpha \delta_2 \text{ bias} * x_1 \\
 &= (4) + (0.1)(-0.0385)(-1)(1) \\
 &= \boxed{4.00385} \approx \textcircled{4}
 \end{aligned}$$

$$\begin{aligned}
 w_{23}(1)_{\text{new}} &= w_{23}(1) + \alpha \delta_2 \text{ bias} * x_2 \\
 &= (5) + (0.1)(-0.0385)(-1)(1) \\
 &= \boxed{5.00385} \approx \textcircled{5}
 \end{aligned}$$