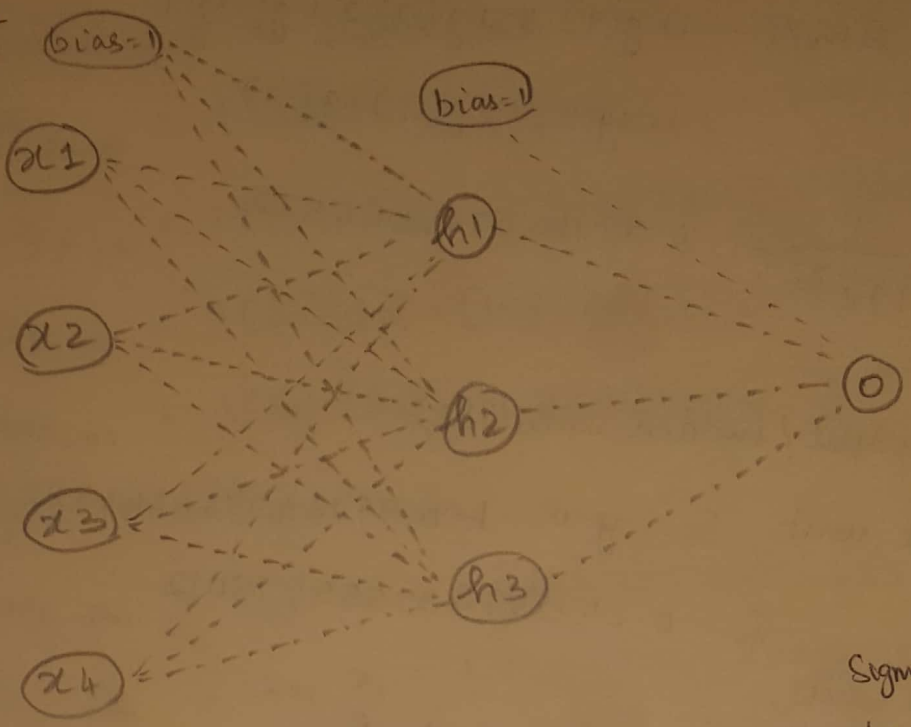
MACHINE LEARNING - ASSIGNMENT 4 (NEURAL NETS)

PART 1



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

Learning rate $\eta = 1$, Error:- Cross entropy, Hidden/Output Units :- Sigmoid

1) $h_1 = g(z_1)$

$z_1 = w_0 + \sum w_i x_i = 1 + 1(2) + 3(3) + 2(-2) + 1(1) = 9$

$h_1 = g(9) = \frac{1}{1+e^{-9}} = 0.9998766054240137$

2) $h_2 = g(z_2)$

$z_2 = w_0 + \sum w_i x_i = 2 + 1(3) + 3(1) + 2(4) + 1(1) = 17$

$h_2 = g(17) = \frac{1}{1+e^{-17}} = 0.9999999860006244$

3) $h_3 = g(z_3)$

$z_3 = w_0 + \sum w_i x_i = -1 + 1(1) + 3(-2) + 2(0) + 1(3) = -3$

$h_3 = g(-3) = \frac{1}{1+e^{3}} = 0.04742587317756678$

Output Unit, $O = g(z_o)$

$$z_o = w_o + \sum w_i x_i = 1 + g(9)\cdot 3 + g(17)\cdot 2 + g(-3)\cdot 1$$

$$= 1 + 3g(9) + 2g(17) + g(-3)$$

$$O = \frac{1}{1+e^{-z_o}} = 0.9976407644444522$$

2) Errors at output / hidden units (δ_j)

Error at output unit, $\delta_o = y - o = 1 - 0.9976407644444522$

$$\delta_o = 0.0023592355555478012$$

Error at hidden units,

$$\delta_j = o_j (1 - o_j) \sum_k w_{kj} \delta_k$$

$$\delta_{h1} = h1 (1 - h1) (\delta_o \cdot w_{h1 \to o})$$

$$= g(9)[1 - g(9)][1 - g(z_o)] \cdot 3$$

$$= 8.732428463315435\, e^{-07}$$

$$\delta_{h2} = h2 (1 - h2) (\delta_o \cdot w_{h2 \to o})$$

$$= g(17)[1 - g(17)](1 - g(z_o)) \cdot 2$$

$$= 1.9534174966644 16 e^{-10}$$

$$\delta_{h3} = h3 (1 - h3) (\delta_o \cdot w_{h3 \to o})$$

$$= g(-3)[1 - g(-3)][1 - g(z_o)] \cdot 1$$

$$= 0.0001065823819 14899 64$$

3) Gradients at various units

$$\Delta w_{ji} = \eta \delta_j o_i$$

From hidden units to output units,

$$\Delta w_{bh \to o} = 1(\delta_o)(1) = 1 - g(z_o) = 0.0023592355554780 12$$

$$\Delta w_{h1 \to o} = 1(\delta_o)(h1) = (1 - g(z_o))\, g(9) = 0.0023589444386069 92$$

$$\Delta w_{h2 \to o} = 1(\delta_o)(h2) = [1 - g(z_o)]\, g(17) = 0.0023592354578071332$$

$$\Delta w_{h3 \to o} = 1(\delta_o)(h3) = [1 - g(z_o)]\, g(-3) = 0.0001118888062501065 1$$

From input to hidden units,

$$\Delta w_{bi \to h1} = \delta_{h1} \cdot bi = 8.732428463315435\, e^{-07}$$

$$\Delta w_{x1 \to h1} = \delta_{h1} \cdot x_1 = 8.732428463315435\, e^{-07}$$

$$\Delta w_{x2 \to h1} = \delta_{h1} \cdot x_2 = 2.619728538994630 6\, e^{-06}$$

$$\Delta w_{x3 \to h1} = \delta_{h1} \cdot x_3 = 1.746485692 66\, e^{-06}$$

$$\Delta w_{x4 \to h1} = \delta_{h1} \cdot x_4 = 8.732428463315435\, e^{-07}$$

$$\Delta w_{bi \to h2} = \delta_{h2} \cdot bi = 1.953417496664416\, e^{-10}$$

$$\Delta w_{x1 \to h2} = \delta_{h2} \cdot x_1 = 1.953417496664416\, e^{-10}$$

$$\Delta w_{x2 \to h2} = \delta_{h2} \cdot x_2 = 5.860252489 99\, e^{-10}$$

$$\Delta w_{x3 \to h2} = \delta_{h2} \cdot x_3 = 3.906834993 33\, e^{-10}$$

$$\Delta w_{x4 \to h2} = \delta_{h2} \cdot x_4 = 1.953417496664416\, e^{-10}$$

$$\Delta w_{bi \to h3} = \delta_{h3} \cdot bi = 0.00010658238191489964$$

$$\Delta w_{x1 \to h3} = \delta_{h3} \cdot x_1 = 0.00010658238191489964$$

$$\Delta w_{x2 \to h3} = \delta_{h3} \cdot x_2 = 0.00031974714574469 89$$

$$\Delta w_{x3 \to h3} = \delta_{h3} \cdot x_3 = 0.00021316476382979928$$

$$\Delta w_{x4 \to h3} = \delta_{h3} \cdot x_4 = 0.00010658238191489964$$