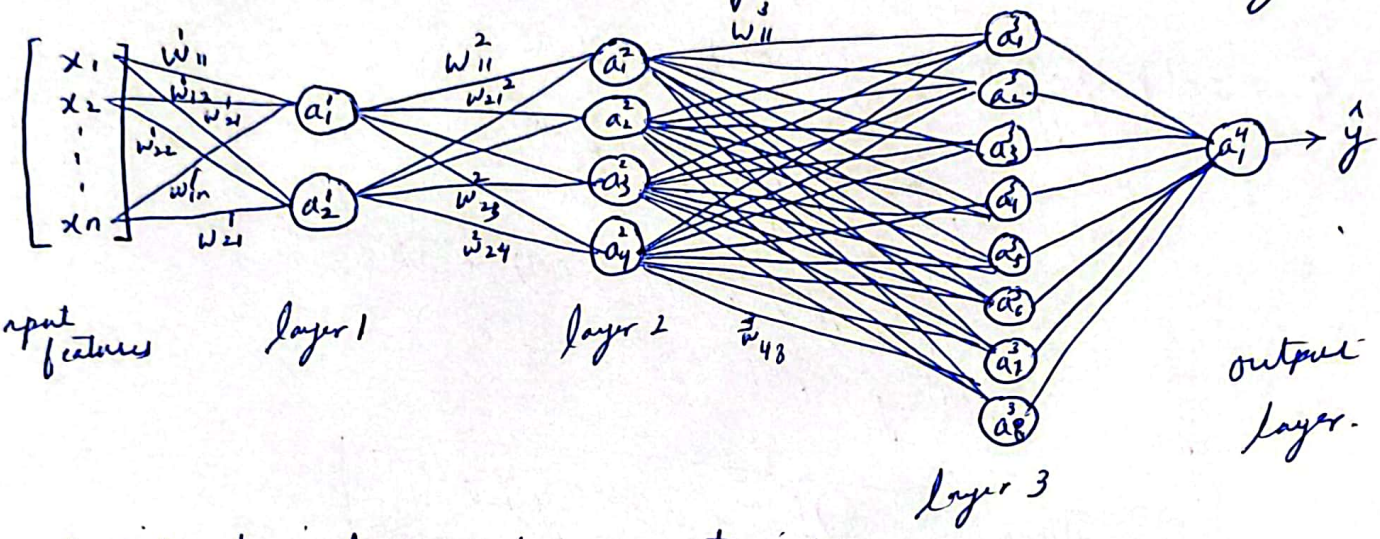
→ AI LAB NUMBER 8

Submitted to: Sir Awais
Submitted by: Ahmed Khalid 200148

1. Data collection, Data Preprocessing & Model Architecture Design :-



2. Initialization & forward propagation :-

- $Z^{[1]}_{2\times 1} = W^{[1]}_{2\times n} \times_{n\times 1} + b^{[1]}_{2\times 1}$
- $a^{[1]}_{2\times 1} = \delta(Z^{[1]})_{2\times 1}$
- $Z^{[2]}_{4\times 1} = W^{[2]}_{4\times 2}\, a^{[1]}_{2\times 1} + b^{[2]}_{4\times 1}$
- $a^{[2]}_{4\times 1} = \delta(Z^{[2]})_{4\times 1}$
- $Z^{[3]}_{8\times 1} = W^{[3]}_{8\times 4}\, a^{[2]}_{4\times 1} + b^{[3]}_{8\times 1}$
- $a^{[3]}_{8\times 1} = \delta[Z^{[3]}]_{8\times 1}$
- $Z^{[4]}_{1\times 1} = W^{[4]}_{1\times 8}\, a^{[3]}_{8\times 1} + b^{[4]}_{1\times 1}$
- $a^{[4]}_{1\times 1} = \delta(Z^{[4]})_{1\times 1}$

3. Loss Calculation :-

$$Loss = -[y \log(\hat{y}) + (1-y)\log(1-a^{[4]})]$$
$$= -[y \log(a^{[4]}) + (1-y)\log(1-a^{[4]})].$$

$$\frac{\partial L}{\partial a^{4}} = \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})}$$

4. Backpropagation for parameter update :-
→ weights

- $\frac{\partial L}{\partial w^{[4]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial w^{[4]}}$

$$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times a^{[3]}.$$

$$\frac{\partial L}{\partial w^{4}} = (a^{[4]} - y)\, a^{[3]}.$$
$1\times 8 \quad 1\times 1 \quad 8\times 1$

hence; $\frac{\partial L}{\partial w^{4}} = (a^{[4]} - y)\, a^{[3]T}$

- $\frac{\partial L}{\partial w^{[3]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial a^{[3]}} \times \frac{\partial a^{[3]}}{\partial z^{[3]}} \times \frac{\partial z^{[3]}}{\partial w^{[3]}}$

$$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times W^{[4]} \times a^{[3]}(1-a^{[3]}) \times a^{[2]}$$

$$\frac{\partial L}{\partial w^{[3]}} = w^{[4]} \times (a^{[4]} - y) \times a^{[3]}(1-a^{[3]}) \times a^{[2]}$$
$8\times 4 \quad 1\times 8 \quad 1\times 1 \quad 8\times 1 \quad 8\times 1 \quad 4\times 1$

$$= w^{[4]T} (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T} = \frac{\partial L}{\partial w^{[3]}}$$
$\frac{8\times 1}{8\times 1} \quad \frac{1\times 1}{8\times 1} \quad \frac{8\times 1}{1\times 8} \quad \frac{8\times 1}{1\times 1} \quad \frac{4\times 1}{1\times 4}$ $\qquad 8\times 4.$

hence; $\frac{\partial L}{\partial w^{(3)}} = w^{[4]T}(a^{[4]} - y) \times a^{(3)T}(1-a^{[3]}) \times a^{[2]T}.$

- $\frac{\partial L}{\partial w^{[2]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial a^{[3]}} \times \frac{\partial a^{[3]}}{\partial z^{[3]}} \times \frac{\partial z^{[3]}}{\partial a^{[2]}} \times \frac{\partial a^{[2]}}{\partial z^{[2]}} \times \frac{\partial z^{[2]}}{\partial w^{[2]}}$

$$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times W^{[4]} \times a^{[3]}(1-a^{[3]}) \times W^{[3]} \times a^{[2]}(1-a^{[2]}) \times a^{[1]}$$

$$\frac{\partial L}{\partial w^{[2]}} = W^{[3]T} W^{[4]T} \times (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T}(1-a^{[2]}) \times a^{[1]T}.$$
$4\times 2 \quad \frac{8\times 4 \quad 1\times 8}{4\times 1} \quad 1\times 1 \quad \frac{8\times 1 \quad 8\times 1}{1\times 1} \quad \frac{4\times 1}{1\times 4} \quad 4\times 1 \quad \frac{2\times 1}{1\times 2}$

$\frac{4\times 1}{4\times 4}$ $\qquad = 4\times 2$ ✓

hence; $\frac{\partial L}{\partial w^{[2]}} = w^{[3]T} w^{[4]T} \times (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T}(1-a^{[2]}) \times a^{[1]T}$

- $\frac{\partial L}{\partial w^{[1]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial a^{[3]}} \times \frac{\partial a^{[3]}}{\partial z^{[3]}} \times \frac{\partial z^{[3]}}{\partial a^{[2]}} \times \frac{\partial a^{[2]}}{\partial z^{[2]}} \times \frac{\partial z^{[2]}}{\partial a^{[1]}} \times \frac{\partial a^{[1]}}{\partial z^{[1]}} \times \frac{\partial z^{[1]}}{\partial w^{[1]}}$

$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times w^{[4]} \times a^{[3]}(1-a^{[3]}) \times w^{[3]} \times a^{[2]}(1-a^{[2]}) \times w^{[2]} \times a^{[1]}(1-a^{[1]}) \times X$

$\frac{\partial L}{\partial w^{[1]}} = w^{[2]T} w^{[3]T} w^{[4]T} \times (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T}(1-a^{[2]}) \times a^{[1]T}(1-a^{[1]}) \times X^T$

$2 \times n$

$4\times 2 \quad 8\times 4 \quad 1\times 8 \quad 1\times 1 \quad 8\times 1 \quad 8\times 4 \quad 4\times 1 \quad 4\times 1 \quad 2\times 1 \quad 2\times 4 \quad n\times 1$

$2\times 8 \quad 8\times 1 \quad 2\times 1 \quad 1\times 8 \quad 2\times 1 \quad 1\times 4 \quad 2\times 1 \quad 1\times 2$

$2\times 4 \quad 2\times 1 \quad 2\times 2$

$2\times 1 \cdot 1\times n \rightarrow 2\times n$

hence; $\frac{\partial L}{\partial w^{[1]}} = w^{[2]T} w^{[3]T} w^{[4]T} \times (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T}(1-a^{[2]}) \times a^{[1]T}(1-a^{[1]}) \times X^T$

→ biases

- $\frac{\partial L}{\partial b^{[4]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial b^{[4]}}$

$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times 1$

$\frac{\partial L}{\partial b^{[4]}} = a^{[4]} - y$

$1\times 1 \quad 1\times 1$

hence; remains same.

- $\frac{\partial L}{\partial b^{[3]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial a^{[3]}} \times \frac{\partial a^{[3]}}{\partial z^{[3]}} \times \frac{\partial z^{[3]}}{\partial b^{[3]}}$

$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times w^{[4]} \times a^{[3]}(1-a^{[3]}) \times 1$

$\frac{\partial L}{\partial b^{[3]}} = w^{[4]T} \times (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times 1$

$8\times 1 \quad 1\times 8 \quad 1\times 1 \quad 8\times 1 \quad 8\times 1$

$8\times 1 \quad 1\times 1 \quad 1\times 8 \quad 8\times 1$

$8\times 1 \quad 1\times 1$

hence; $\frac{\partial L}{\partial b^{[3]}} = w^{[4]T} \times (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]})$

- $\frac{\partial L}{\partial b^{[2]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial a^{[3]}} \times \frac{\partial a^{[3]}}{\partial z^{[3]}} \times \frac{\partial z^{[3]}}{\partial a^{[2]}} \times \frac{\partial a^{[2]}}{\partial z^{[2]}} \times \frac{\partial z^{[2]}}{\partial b^{[2]}}$

$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times w^{[4]} \times a^{[3]}(1-a^{[3]}) \times w^{[3]} \times a^{[2]} \times (1-a^{[2]}) \times 1$

$\frac{\partial L}{\partial b^{[2]}} = w^{[3]T} w^{[4]T} [a^{[4]} - y] \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T}(1-a^{[2]})$

$4\times1 \quad \frac{8\times4 \quad 1\times8}{4\times1} \quad 1\times1 \quad \frac{8\times1 \quad 8\times1}{1\times1} \quad \frac{4\times1 \quad 4\times1}{1\times1}$

hence; $\frac{\partial L}{\partial b^{[2]}} = w^{[3]T} w^{[4]T} (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \, a^{[2]T}(1-a^{[2]})$

- $\frac{\partial L}{\partial b^{[1]}} = \frac{\partial L}{\partial a^{[4]}} \times \frac{\partial a^{[4]}}{\partial z^{[4]}} \times \frac{\partial z^{[4]}}{\partial a^{[3]}} \times \frac{\partial a^{[3]}}{\partial z^{[3]}} \times \frac{\partial z^{[3]}}{\partial a^{[2]}} \times \frac{\partial a^{[2]}}{\partial z^{[2]}} \times \frac{\partial z^{[2]}}{\partial a^{[1]}} \times \frac{\partial a^{[1]}}{\partial z^{[1]}} \times \frac{\partial z^{[1]}}{\partial b^{[1]}}$

$= \frac{a^{[4]} - y}{a^{[4]}(1-a^{[4]})} \times a^{[4]}(1-a^{[4]}) \times w^{[4]} \times a^{[3]}(1-a^{[3]}) \times w^{[3]} \times a^{[2]}(1-a^{[2]}) \times w^{[2]} \times a^{[1]}(1-a^{[1]})$

$\frac{\partial L}{\partial b^{[1]}} = w^{[2]T} w^{[3]T} w^{[4]T} \times (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T}(1-a^{[2]}) \times a^{[1]T}(1-a^{[1]})$

$2\times1 \quad \frac{4\times2 \quad 8\times4 \quad 1\times8}{2\times4 \quad 4\times8 \quad 8\times1} \quad 1\times1 \quad \frac{8\times1 \quad 8\times1}{1\times1} \quad \frac{4\times1 \quad 4\times1}{1\times1} \quad \frac{2\times1 \quad 2\times1}{1\times1}$

2×1

hence; $\frac{\partial L}{\partial b^{[1]}} = w^{[2]T} w^{[3]T} w^{[4]T} (a^{[4]} - y) \times a^{[3]T}(1-a^{[3]}) \times a^{[2]T}(1-a^{[2]}) \times a^{[1]T}(1-a^{[1]})$