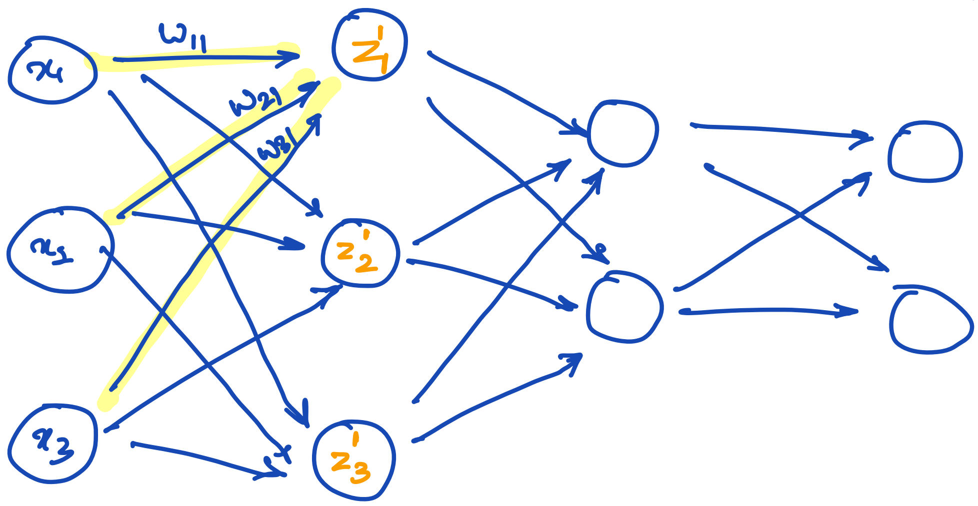
3 layer Network

3 features
2 Classes

layer 0	layer 1	layer 2	layer 3
Input layer	HL 1	HL 2	Output layer



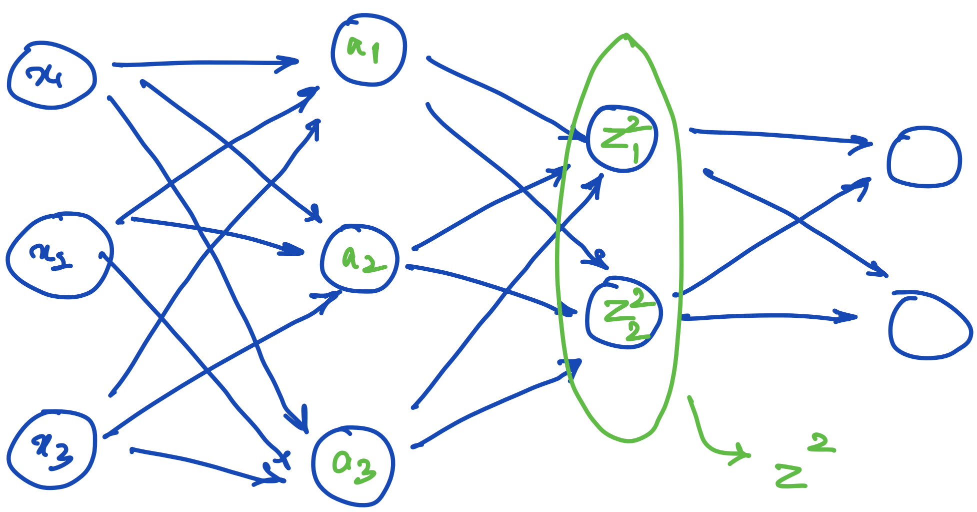
n = features

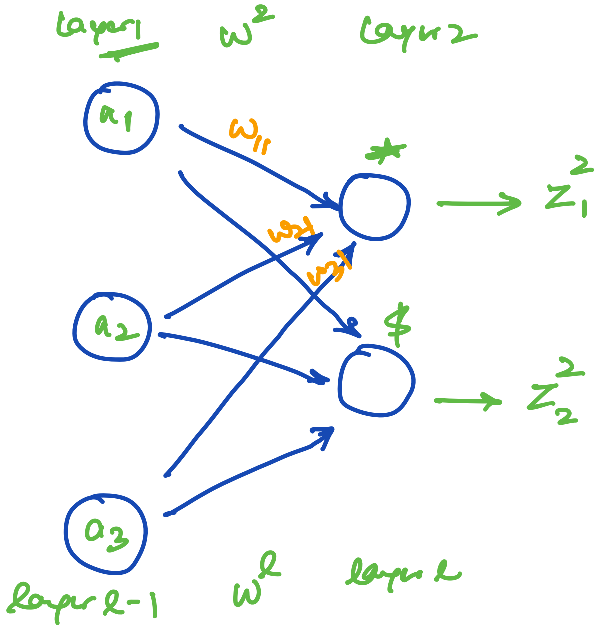
$$z^l_j = \sum_{i=1}^{n} w_{i,j}\, x_i + b \Rightarrow z^1_j = \sum_{i=1}^{n} w_{i,j}\, x_i + b$$

$$= w_{11}x_1 + w_{21}x_2 + w_{31}x_3$$

j = neuron no.

	layer 1		layer 2		layer 3
w^1	HL1	w^2	HL2	w^3	





$$w^2 = \begin{array}{cc} * & \$ \\ \end{array}$$

$$w^2 = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix}$$

3×2

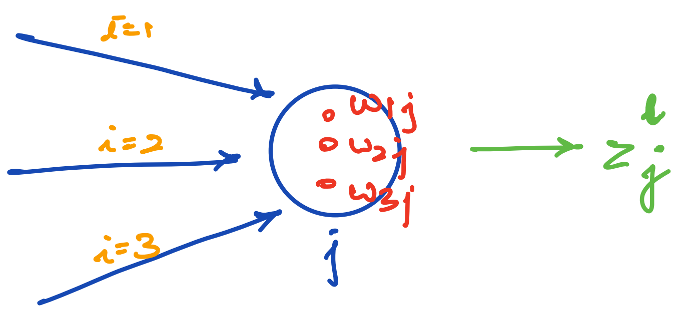
$$\boxed{z_j^l = \sum_i w_{ij}^l a_i^{l-1} + b}$$

i → no. of neurons in $l-1$ layer

j = neuron no.

$j = 1, 2$

↳ 2 neurons in layer 2



w_{ij} → weight connecting ith neuron in $l-1$ layer & jth neuron of l layer

$$w^l = \begin{bmatrix} w_{11}^l & w_{12}^l & w_{13}^l \\ w_{21}^l & w_{22}^l & w_{23}^l \\ w_{31}^l & w_{32}^l & w_{33}^l \end{bmatrix}$$

wts for neuron 1 in lth layer

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{matrix} \longrightarrow & z_1 \\ \longrightarrow & z_2 \end{matrix} \Longrightarrow \begin{bmatrix} \sigma(z_1) \\ \sigma(z_2) \end{bmatrix} z^l$$

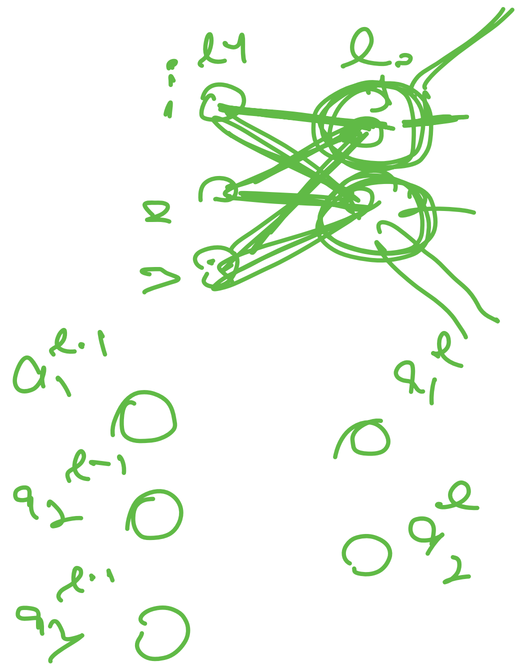
Vector Notation:

$$\boxed{z_j^l = \sum_i w_{ij}^l a_i^{l-1} + b^l}$$

every neuron

3X2 3X1

$$\boxed{z^l = [w^l]^T a^{l-1} + b^l}$$

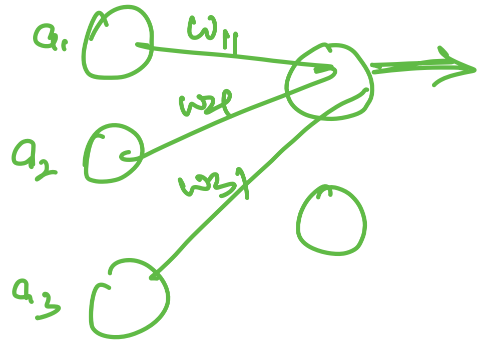


$$w^l = \begin{bmatrix} w_{11}^l & w_{12}^l \\ w_{21}^l & w_{22}^l \\ w_{31}^l & w_{32}^l \end{bmatrix}$$

3X2

$$(w^l)^T = \begin{bmatrix} w_{11}^l & w_{21}^l & w_{31}^l \\ w_{12}^l & w_{22}^l & w_{32}^l \end{bmatrix}_{2\times 3} \begin{bmatrix} a_1^{l-1} \\ a_2^{l-1} \\ a_3^{l-1} \end{bmatrix}_{3\times 1}$$

$$w_{11}^{\ell} a_1^{\ell-1} + w_{21}^{\ell} a_2^{\ell-1} + w_{31}^{\ell} a_3^{\ell-1} = z_1^{\ell}$$



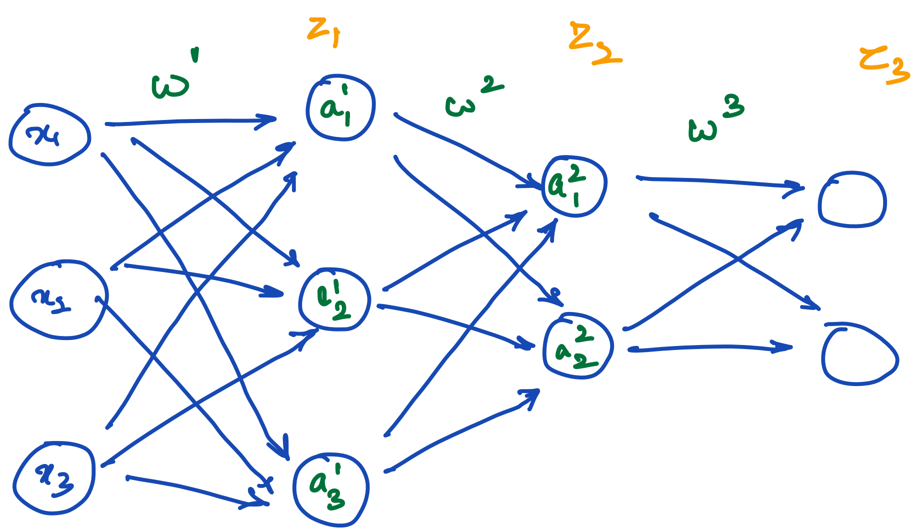
$$z^{\ell} = \begin{bmatrix} z_1^{\ell} \\ z_2^{\ell} \end{bmatrix}$$

$$\sigma(z^{\ell}) = \begin{bmatrix} \sigma(z_1^{\ell}) \\ \sigma(z_2^{\ell}) \end{bmatrix}$$

$$z^1 = (w^1)^T a + b^1$$

$$a^1 = \sigma(z^1)$$

$$z^2 = (w^2)^T a + b^2$$



$z^3 = (w^3)^T a^2 + b^3$

$\hat{y} = \text{softmax}(z^3)$

Exact softmax:

$[1, 2, 3]$

$$\left[\frac{e^1}{e^1+e^2+e^3}, \frac{e^2}{e^1+e^2+e^3}, \frac{e^3}{e^1+e^2+e^3} \right]$$

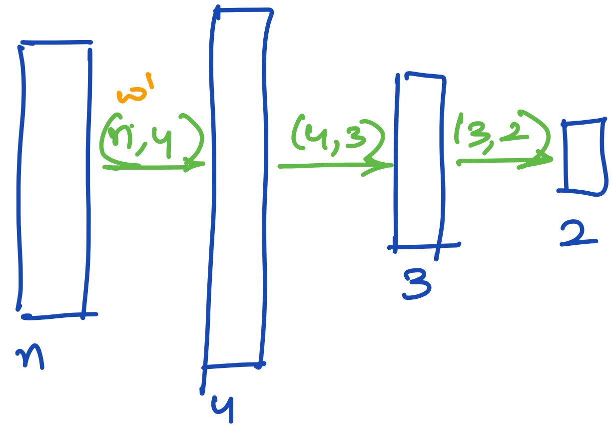
$$\cancel{\frac{1}{1+2+3}, \frac{2}{1+2+3}, \frac{3}{1+2+3}}$$

1 example $\rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

m examples

$$X = \begin{bmatrix} x_1^1 & x_2^1 & x_3^1 & \cdots & x_n^1 \\ & & x^{(i)} & & \end{bmatrix}_{m \times n}$$

Layer 0 Layer 1 Layer 2 Layer 3



$$z^1 = \underbrace{(w^1)^T}_{4\times n} \cdot \underbrace{x}_{m\times 4} + b^1 \qquad w^1 = \underline{n\times 4} \qquad x = \underline{m\times n}$$

$$z^1 = \underbrace{x}_{m\times n} \cdot \underbrace{w^1}_{n\times 4} + b^1$$

$$\frac{m\times n \quad n\times 4}{m\times 4}$$

$$\begin{bmatrix} z_1^1 \\ z_2^1 \\ z_3^1 \\ z_4^1 \end{bmatrix} \longrightarrow \begin{bmatrix} z_1^1 & z_2^1 & z_3^1 & z_4^1 \\ & & & \\ & & & \\ & & & \end{bmatrix}$$

1 example

2 example

$\vdots$

m

$\underline{m\times 4}$

$$Z^1 = \begin{bmatrix} z_1^{11} & z_2^{11} & z_3^{11} & z_4^{11} \\ z_1^{12} & z_2^{12} & z_3^{12} & z_4^{12} \\ \vdots & & & \\ z_1^{1m} & z_2^{1m} & z_3^{1m} & z_4^{1m} \end{bmatrix}_{m\times 4}$$

$z_j^{li} \longrightarrow$ ith example

l layer jth neuron output

$$a^1 = \sigma\begin{bmatrix} | \\ - z^1 - \\ | \end{bmatrix}$$

$$Z^1 = \begin{bmatrix} z_1^{11} & z_2^{11} & z_3^{11} & z_4^{11} \\ z_1^{12} & z_2^{12} & z_3^{12} & z_4^{12} \\ \vdots & & & \\ z_1^{1m} & z_2^{1m} & z_3^{1m} & z_4^{1m} \end{bmatrix}_{m\times 4} + \begin{bmatrix} b_1^1 & b_2^1 & b_3^1 & b_4^1 \end{bmatrix}_{1\times 4}$$

$$Z^1 = \begin{bmatrix} z_1^{11}+b_1^1 & z_2^{11}+b_2^1 & z_3^{11}+b_3^1 & z_4^{11}+b_4^1 \\ z_1^{12}+b_1^1 & z_2^{12}+b_2^1 & z_3^{12}+b_3^1 & z_4^{12}+b_4^1 \\ \vdots & & & \\ z_1^{1m}+b_1^1 & z_2^{1m}+b_2^1 & z_3^{1m}+b_3^1 & z_4^{1m}+b_4^1 \end{bmatrix}_{m\times 4}$$

Backpropagation