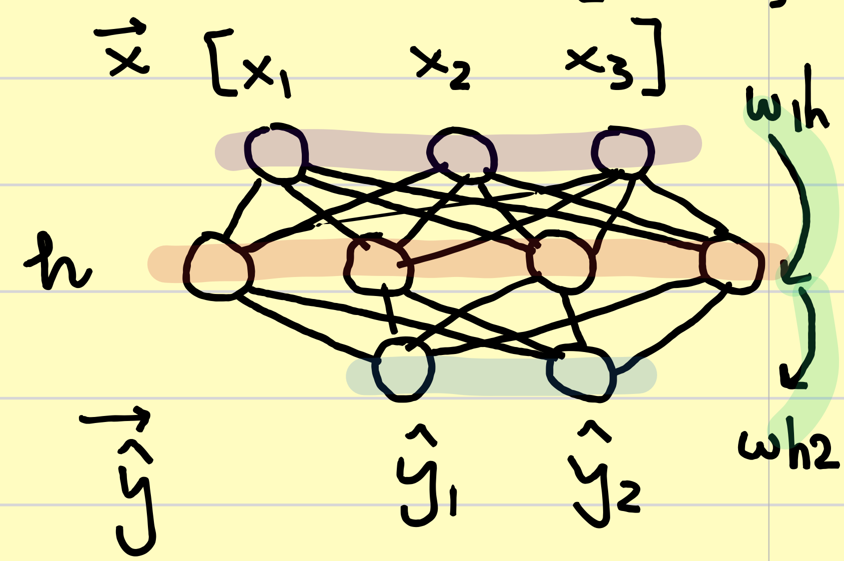
20250403_WIN6_Business_Intelligence

Example 3. 3 Layer Perception. FORWARD PASS [3,4,2]

$w_{ih} = [3 \times 4]$ $\quad w_{h2} = [4 \times 2]$



$\vec{x}$ [1x4]

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \begin{matrix} \leftarrow x_1 \\ \leftarrow x_2 \\ \leftarrow x_3 \end{matrix}$$

w_{ih} $\quad b$

$$\begin{bmatrix} 1 & -1 & 1 & -5 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & -2 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 3 \\ 5 \\ 3 \end{bmatrix} \underset{\sim}{\text{ReLU}} \begin{bmatrix} 0 \\ 3 \\ 5 \\ 3 \end{bmatrix}$$

w_{h2} $\quad b_2$ [1x2]

$$\begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix} \quad \begin{bmatrix} -2 \\ 3 \end{bmatrix} \underset{\sim}{\text{ReLU}} \begin{bmatrix} 0 \\ 3 \end{bmatrix} \begin{matrix} \hat{y}_1 \\ \hat{y}_2 \end{matrix}$$

1. Feed Forward $1 \rightarrow h$

$2 \cdot 1 + 1 \cdot (-1) + 3 \cdot 1 + (-5) = -1$

$2 \cdot 1 + 1 \cdot 1 + 3 \cdot 0 + 0 = 3$

$2 \cdot 0 + 1 \cdot 1 + 3 \cdot 1 + 1 = 5$

$2 \cdot 1 + 1 \cdot 0 + 3 \cdot 1 + (-2) = 3$

2. Feed Forward $h \rightarrow 2$

$0 \cdot 1 + 3 \cdot 1 + 5 \cdot (-1) + 3 \cdot 0 + 0 = -2$

$0 \cdot 0 + 3 \cdot 0 + 5 \cdot 1 + 3(-1) + 1 = 3$

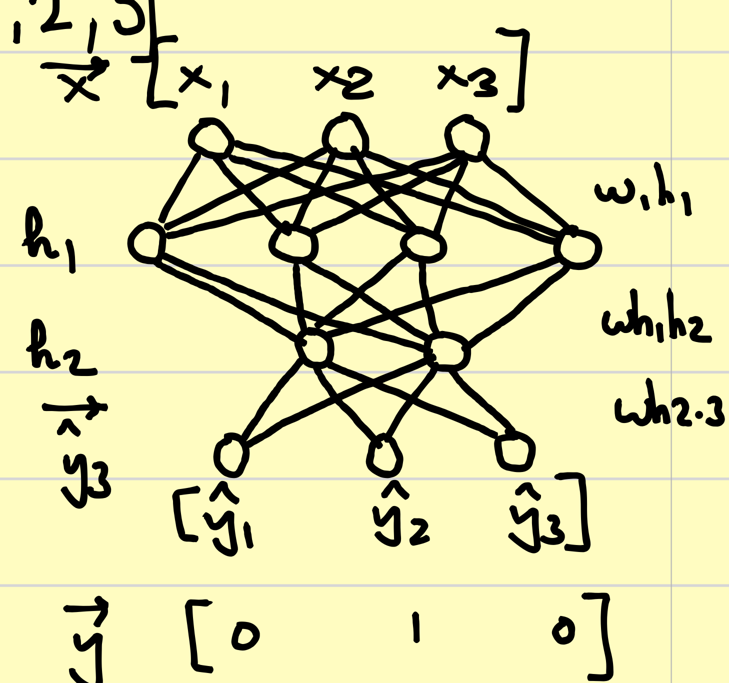
DECISION LAYER

The decision Layer has the property that transforms the output $\vec{\hat{y}}$ into a set of probabilities (PDF. Probability distribution function). All elements of the PDF add up to 1.

$$\text{SOFTMAX}\left[\vec{x}\right] = \begin{bmatrix} \left|\frac{x_1}{\sum |x_i|}\right| \\ \vdots \\ \left|\frac{x_n}{\sum |x_i|}\right| \end{bmatrix}$$

$$\text{SOFTMAX}\begin{bmatrix} 3 \\ -7 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{3}{3+7+4} \\ \frac{7}{3+7+4} \\ \frac{4}{3+7+4} \end{bmatrix} = \begin{bmatrix} 3/14 \\ 7/14 \\ 4/14 \end{bmatrix} = \begin{bmatrix} 0'15 \\ 0'5 \\ 0'35 \end{bmatrix}$$

Example 4. Perception with 2 hidden layers with Feed forward & Back Propagation. [3,4,2,3]



$\vec{x}$

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 2 & -4 & 4 & -2 \\ 1 & -2 & 2 & -1 \\ 3 & -6 & 6 & -3 \end{bmatrix} w_ih_1^* \qquad \begin{bmatrix} 1 & -2 & 2 & -1 \end{bmatrix} b_1^*$$

w_ih_1 b_1

$$\begin{bmatrix} 1 & -1 & 1 & -5 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & -2 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 3 \\ 5 \\ 3 \end{bmatrix} \underset{\text{ReLU}}{\simeq} \begin{bmatrix} 0 \\ 3 \\ 5 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 3 & -3 \\ 5 & -5 \\ 3 & -3 \end{bmatrix} wh_1h_2^*$$

$$\begin{bmatrix} 1 & -2 & 2 & -1 \end{bmatrix} \leftarrow \begin{bmatrix} 1 & -1 \end{bmatrix} b_2^*$$

ERROR$_1$

wh_1h_2 b_2

$$\begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 3 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 4 \end{bmatrix} \underset{\text{ReLU}}{\simeq} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad \begin{bmatrix} 0'46 & -1'54 & 1'08 \\ 0'92 & -3'08 & 2'16 \end{bmatrix} \text{3.b. } wh_{2,3}^*$$

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \xleftarrow{3.c} \begin{bmatrix} 0'23 & -0'77 & 0'54 \end{bmatrix} b_3^* \xleftarrow{3.a}$$

ERROR$_2$

$wh_{2\cdot3}$ b_3

$0'23 \cdot 2 + (-0'77) \cdot 0 + 0'54 \cdot 1 = 1$

$0'23 \cdot 0 + (-0'77) \cdot 2 + 0'54 \cdot 1 = -1$

$$\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -5 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 3 \\ 7 \end{bmatrix} \underset{\text{Soft Max}}{\simeq} \underbrace{\begin{bmatrix} 0'23 \\ 0'23 \\ 0'54 \end{bmatrix}}_{\hat{y}} - \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\vec{y}} = \underbrace{\begin{bmatrix} 0'23 \\ -0'77 \\ 0'54 \end{bmatrix}}_{\text{ERROR}_3}$$

1. Feed forward

$x \to h_1$: $2\cdot1 + 1\cdot(-1) + 3\cdot1 + (-5) = -1$

$2\cdot1 + 1\cdot1 + 3\cdot0 + 0 = 3$

$2\cdot0 + 1\cdot1 + 3\cdot1 + 1 = 5$

$2\cdot1 + 1\cdot0 + 3\cdot1 + (-2) = 3$

$h_1 \to h_2$: $0\cdot1 + 3(-1) + 5\cdot1 + 3\cdot0 + 0 = 2$

$0\cdot0 + 3\cdot1 + 5(-1) + 3\cdot1 + 3 = 4$

$h_2 \to \hat{y}_3$: $2\cdot2 + 4\cdot0 + (-1) = 3$

$2\cdot0 + 4\cdot2 + (-5) = 3$

$2\cdot1 + 4\cdot1 + 1 = 7$

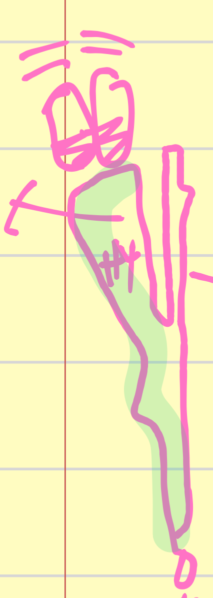
2. Soft Max

$$\text{SoftMax}\begin{bmatrix}3\\3\\7\end{bmatrix} = \begin{bmatrix}\frac{3}{3+3+7}\\ \frac{3}{3+3+7}\\ \frac{7}{3+3+7}\end{bmatrix} = \begin{bmatrix}0'23\\0'23\\0'54\end{bmatrix}$$

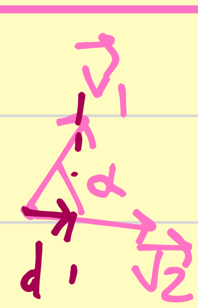
3. Back Propagation.

3.a. Transpose the ERROR VECTOR as new bias.

3.b. We "DOT.PRODUCT" of the trasposed Error (bias) and the output of the previous layer.



SCALAR PRODUCT



$\vec{v}_1 \cdot \vec{v}_2 = \|\vec{v}_1\| \, \|\vec{v}_2\| \cdot \cos\alpha = d$

DOT. PRODUCT

$\vec{v}_1 \cdot \vec{v}_2 = [x_1 \cdot x_2 , y_1 \cdot y_2]$

$[x_1 y_1][x_2 y_2]$

Dot Product.

2	0'46	-1'54	1'08
4	0'92	-3'08	2'16
	0'23	-0'77	0'54

2 · [0'23] = 0'46   4 · [0'23] = 0'92

2 · [-0'77] = -1'54   4 · [-0'77] = -3'08

2 · [0'54] = 1'08   4 [0'54] = 2'16

3.c. We multiply the transposed error with the weight matrix of the current layer. This way we get the new error of the layer above.

$\vec{x}$

2
1
3

$w^*_{1h_1}$			b^*_1			
2	1	3	1	15	ReLU	15
-4	-2	-6	-2	-30	≈	0
4	2	6	2	30		30
2	-1	3	-1	11		11

$w^*_{h_1h_2}$

0	3	5	3	1	184	ReLU	184
0	-3	-5	-3	-1	-184	≈	0

$w^*_{h_{2.3}}$

0'46	0'92	0'23	84'81
-1'54	-3'08	-0'77	-2
1'08	2'16	0'54	