Tema 1

$$x = [1,3,0]^T = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$$

$$W = \begin{bmatrix} 0.3 & 0.1 & -2 \\ -0.6 & -0.5 & 2 \\ -1 & -0.5 & 0.1 \end{bmatrix} \qquad b = [0.1, 0.1, 0.1]^{(T)}$$

$$y = [0,1,0]$$

1) $\boxed{z = W^T x + b}$

$$z = \begin{bmatrix} 0.3 & -0.6 & -1 \\ 0.1 & -0.5 & -0.5 \\ -2 & 2 & 0.1 \end{bmatrix} \cdot \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix} = \begin{pmatrix} -1,4 \\ -1,3 \\ 4,1 \end{pmatrix}$$

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

$$\hat{y} = \left(\frac{e^{-1,4}}{e^{-1,4}+e^{-1,3}+e^{4,1}} , \frac{e^{-1,3}}{e^{-1,4}+e^{-1,3}+e^{4,1}} , \frac{e^{4,1}}{e^{-1,4}+e^{-1,3}+e^{4,1}} \right)^T$$

$$\hat{y} = (0,004052, 0,004478, 0,991470)^T$$

Subject : Date :...../...../.....

3) $\nabla_z L = \hat{y} - y$

$\nabla_z L = (0,004052, 0,004478, 0,99147)^T - (0, 1, 0)^T$

$= (0,004052, -0,995522, 0,99147)^T$

4) $\nabla_W L = \nabla_z L \, x^T$

$= (0,004052, -0,995522, 0,99147)^T \cdot (-1, 3, 0)$

$= \begin{pmatrix} 0,004052 \\ -0,995522 \\ 0,99147 \end{pmatrix} \cdot (1,3,0) = \begin{pmatrix} 0,004052 & 0,012156 & 0 \\ -0,995522 & -2,986566 & 0 \\ 0,99147 & 2,97441 & 0 \end{pmatrix}$

5) $\nabla b_L = \nabla_z L$

$\nabla b_L = \begin{pmatrix} 0,004052 \\ -0,995522 \\ 0,99147 \end{pmatrix}$

6) $W \leftarrow W - \eta \nabla_W L$ ($\eta = 0,2$)

$W \leftarrow \begin{pmatrix} 0.3 & 0.1 & -2 \\ -0.6 & -0.5 & 2 \\ -1 & -0.5 & 0.1 \end{pmatrix} - 0,2 \cdot \begin{pmatrix} 0,004052 & 0,012156 & 0 \\ -0,995522 & -2,986566 & 0 \\ 0,99147 & 2,97441 & 0 \end{pmatrix}$

$W \leftarrow \begin{pmatrix} 0,299189 & 0,097568 & -2 \\ -0,400895 & 0,097313 & 2 \\ -1,198294 & -1,094882 & 0,1 \end{pmatrix}$ ✓

$b \leftarrow b - \eta \nabla_b L$

$b \leftarrow \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix} - 0,2 \cdot \begin{pmatrix} 0,004052 \\ -0,995522 \\ 0,99147 \end{pmatrix} = \begin{pmatrix} 0,099189 \\ 0,299104 \\ -0,098294 \end{pmatrix}$