

Date

Retele Neuronale : Tema 1

$$X = [1, 3, 0]$$

$$b = [0, 1, 0, 1]$$

$$W = \begin{bmatrix} 0,3 & 0,1 & -2 \\ -0,6 & -0,5 & 2 \\ -1 & -0,5 & 0,1 \end{bmatrix}$$

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

$$Z = W^T X + b = \begin{bmatrix} 0,3 & -0,6 & -1 \\ 0,1 & -0,5 & -0,5 \\ -2 & 2 & 0,1 \end{bmatrix} \cdot \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0,1 \\ 0,1 \\ 0,1 \end{pmatrix}$$

$$= \begin{pmatrix} 0,3 & -1,8 & 0 \\ 0,1 & -1,5 & 0 \\ -2 & 6 & 0 \end{pmatrix} + \begin{pmatrix} 0,1 \\ 0,1 \\ 0,1 \end{pmatrix} = \begin{pmatrix} -1,5 \\ -1,4 \\ 4 \end{pmatrix} + \begin{pmatrix} 0,1 \\ 0,1 \\ 0,1 \end{pmatrix}$$

$$= \begin{pmatrix} -1,4 \\ -1,3 \\ 4,1 \end{pmatrix}$$

$$\hat{y} = \text{softmax}(Z)$$

$$P_1 = \exp(-1,4) / (\exp(-1,4) + \exp(-1,3) + \exp(4,1)) \\ = 0,246 : (0,246 + 0,272 + 60,340) \\ = 0,246 : 60,858 = \underline{0,004}$$

$$\hat{y} = \begin{pmatrix} 0,004 \\ 0,004 \\ 0,991 \end{pmatrix}$$

$$P_2 = \exp(-1,3) : 60,858 = 0,272 : 60,858 \\ = \underline{0,004}$$

$$P_3 = \exp(4,1) : 60,858 = \underline{0,991}$$

①

Date

$$\nabla_z L = \hat{y} - y = \begin{pmatrix} 0,004 \\ 0,004 \\ 0,991 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,004 \\ -0,996 \\ 0,991 \end{pmatrix}$$

$$\nabla_w L = \nabla_z L X^T = \begin{pmatrix} 0,004 \\ -0,996 \\ 0,991 \end{pmatrix} \cdot (1, 3, 0) = \begin{pmatrix} 0,004 & 0,012 & 0 \\ -0,996 & -2,988 & 0 \\ 0,991 & 2,973 & 0 \end{pmatrix}$$

$$\nabla_b L = \begin{pmatrix} 0,004 \\ -0,996 \\ 0,991 \end{pmatrix}$$

$$W = W - \eta \nabla_w L$$

$$W \leftarrow \begin{pmatrix} 0,3 & 0,1 & -2 \\ -0,6 & -0,5 & 2 \\ -1 & -0,5 & 0,1 \end{pmatrix} - 0,3 \begin{pmatrix} 0,004 & 0,012 & 0 \\ -0,996 & -2,988 & 0 \\ 0,991 & 2,973 & 0 \end{pmatrix}$$

$$W \leftarrow \begin{pmatrix} 0,3 & 0,1 & -2 \\ -0,6 & -0,5 & 2 \\ -1 & -0,5 & 0,1 \end{pmatrix} - \begin{pmatrix} 0,001 & 0,003 & 0 \\ -0,298 & -0,896 & 0 \\ 0,297 & 0,891 & 0 \end{pmatrix}$$

$$W \leftarrow \begin{pmatrix} 0,299 & 0,097 & -2 \\ 0,302 & 0,396 & 2 \\ -1,297 & 1,391 & 0,1 \end{pmatrix}$$

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

(2)

Date

$$h \leftarrow \begin{pmatrix} 0,1 \\ 0,1 \\ 0,1 \end{pmatrix} - 0,1 \cdot \begin{pmatrix} 0,004 \\ -0,996 \\ 0,991 \end{pmatrix}$$

$$\leftarrow \begin{pmatrix} 0,1 \\ 0,1 \\ 0,1 \end{pmatrix} - \begin{pmatrix} 0 \\ -0,099 \\ 0,099 \end{pmatrix}$$

$$\leftarrow \begin{pmatrix} 0,1 \\ 0,199 \\ 0,001 \end{pmatrix}$$

Observation: a lower learning rate (η) means a more accurate result but a longer training time when it comes to iterations.

3