

Data rețeaua din imagine, stabiliți care ar trebui să fie ponderile (weight-urile) noi după rularea algoritmului de backpropagation într-o singură iterație, folosind inputul (2,6) ce are clasa (target) 0 și rețeaua din imaginea alăturată. Detalii despre configurația rețelei:

- toate funcțiile de activare sunt de tip logistic (sigmoid),
- funcția de eroare este Cross-entropy.
- Neuronii nu au bias.
- Rata de învățare este:  $\eta = 0.5$
- Aproximați valoarea funcției sigmoid folosind graficul de mai jos.

Formule ajutătoare:

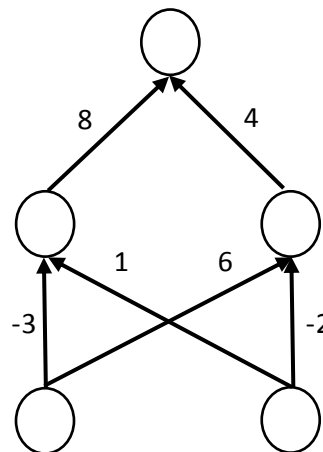
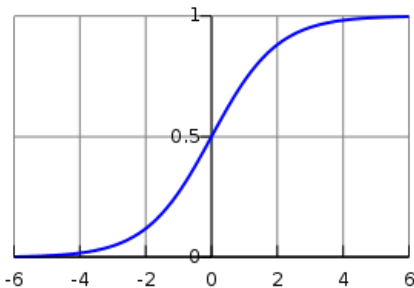
$$\delta_i^l = y_i^l - t_i^l$$

$$\delta_i^l = y_i^l(1 - y_i^l) \sum_k \delta_k^{l+1} \cdot w_{ik}^{l+1}$$

$$\frac{\partial C}{\partial w_{ij}^l} = \delta_j^l y_i^{l-1}$$

$$w_{ij}^l = w_{ij}^l - \eta * \frac{\partial C}{\partial w_{ij}^l}$$

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



### 1. Forward Pass:

Neuronii de pe stratul 1, stratul de input, nu au funcție de activare!

$$z_1^2 = 2 * (-3) + 6 * 1 = 0; y_1^2 = \sigma(0) = 0.5$$

$$z_2^2 = 2 * (6) + 6 * (-2) = 0; y_2^2 = \sigma(0) = 0.5$$

$$z^3 = 8 * (0.5) + 4 * (0.5) = 6; y^3 = 1$$

Calculăm eroarea:

$$\delta^3 = y^3 - t = 1 - 0 = 1$$

2a. Propagăm eroarea la neuronii din stratul anterior

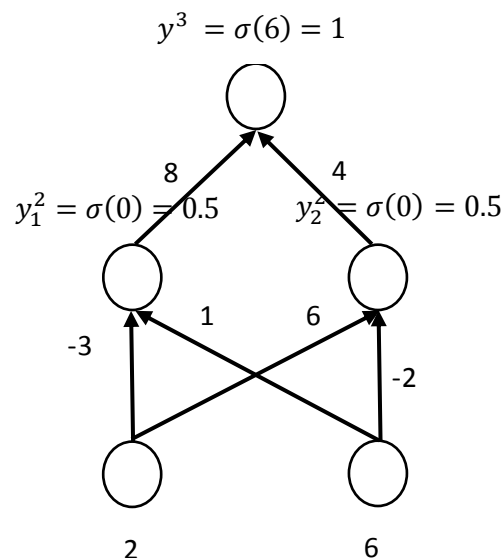
$$\delta_1^2 = y_1^2 * (1 - y_1^2) * (\delta^3 * w_{11}^3) = 0.5 * (1 - 0.5) * (1 * 8) = 2$$

$$\delta_2^2 = y_2^2 * (1 - y_2^2) * (\delta^3 * w_{21}^3) = 0.5 * (1 - 0.5) * (1 * 4) = 1$$

2b. Calculăm cum se modifică ponderile dintre stratul 2 și stratul 3

$$\frac{\partial C}{\partial w_{11}^3} = \delta^3 * y_1^2 = 1 * 0.5 = 0.5. \quad w_{11}^3 = 8 - 0.5 * 0.5 = 7.75$$

$$\frac{\partial C}{\partial w_{21}^3} = \delta^3 * y_2^2 = 1 * 0.5 = 0.5. \quad w_{21}^3 = 4 - 0.5 * 0.5 = 3.75$$



2c. Calculăm cum se modifică ponderile dintre stratul 1 și stratul 2 pe baza erorilor propagate (calculate la punctul 2a)

$$\frac{\partial C}{\partial w_{11}^2} = \delta_1^2 * y_1^1 = 2 * (2) = 4$$

$$\frac{\partial C}{\partial w_{12}^2} = \delta_1^2 * y_1^1 = 1 * (2) = 2$$

$$\frac{\partial C}{\partial w_{21}^2} = \delta_1^2 * y_2^1 = 2 * (6) = 12$$

$$\frac{\partial C}{\partial w_{22}^2} = \delta_2^2 * y_2^1 = 1 * (6) = 6$$

Pe baza gradientilor calculati mai sus calculam ponderile (weighturile) noi

$$\begin{aligned} w_{11}^2 &= -3 - 0.5 * 4 = -5 & w_{12}^2 &= 6 - 2 * 0.5 = 5 \\ w_{21}^2 &= 1 - 0.5 * 12 = -5 & w_{22}^2 &= -2 - 0.5 * 6 = -5 \end{aligned}$$