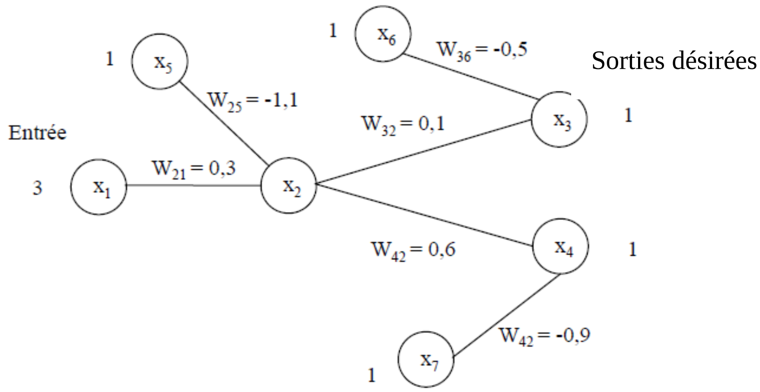
TD3 : Réseaux de Neurones

Exercice 1 :
Effectuez un tour de l'algorithme de rétropropagation des erreurs et indiquez la valeur des nouveaux poids pour : $\mathbf{w_{21}}$, $\mathbf{w_{25}}$, $\mathbf{w_{32}}$, $\mathbf{w_{42}}$. Vous devez présenter tous les calculs. La valeur de la constante d'apprentissage est $\eta = 0{,}05$. La fonction d'activation est la fonction suivante :

$$f = \begin{cases} 1 \text{ si } \sum_{i=0}^{n} w_i \cdot x_i > 0 \\ -1 \text{ sinon} \end{cases}$$



SOLUTION

6. Voici les calculs

$y_2 = (3*0{,}3) + (1 * -1{,}1) = -0{,}2 \qquad o_2 = -1$

$y_3 = (0{,}1*-1) + (1 * -0{,}5) = -0{,}6 \qquad o_3 = -1$

$y_4 = (0{,}6*-1) + (1 * -0{,}9) = -1{,}5 \qquad o_4 = -1$

$\delta_3 = -1(1 + 1)(1 + 1) = -4$

$\delta_4 = -1(1 + 1)(1 + 1) = -4$

$\delta_2 = -1(1 + 1)((0{,}1 * -4) + (0{,}6 * -4)) = 5{,}6$

$\Delta w_{21} = 0{,}05 * 5{,}6 * 3 = 0{,}84$

$\Delta w_{25} = 0{,}05 * 5{,}6 * 1 = 0{,}28$

$\Delta w_{32} = 0{,}05 * -4 * -1 = 0{,}2$

$\Delta w_{36} = 0{,}05 * -4 * 1 = -0{,}2$

$\Delta w_{42} = 0{,}05 * -4 * -1 = 0{,}2$

$\Delta w_{47} = 0{,}05 * -4 * 1 = -0{,}2$

$w_{21} = 0{,}3 + 0{,}84 = 1{,}14$

$w_{25} = -1{,}1 + 0{,}28 = -0{,}82$

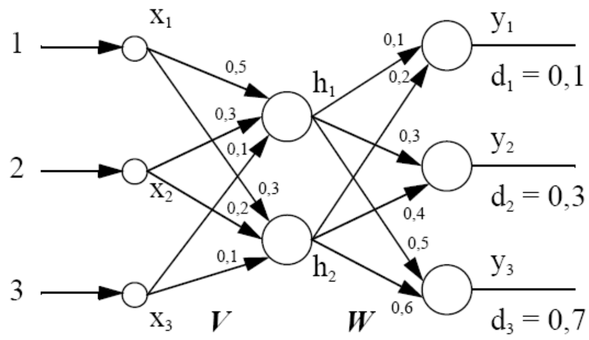
$w_{32} = 0{,}1 + 0{,}2 = 0{,}3$

$w_{36} = -0{,}5 + -0{,}2 = -0{,}7$

$w_{42} = 0{,}6 + 0{,}2 = 0{,}8$

$w_{47} = -0{,}9 + -0{,}2 = -1{,}1$

Exercice 2 :
Soit le perceptron multicouche suivant :



Dans l’unique but de simplifier les calculs, les neurones ne sont pas munis de l’habituel paramètre de polarisation (seuil). Les poids de connexion affichés directement sur la connexion sont résumés dans les deux matrices de connexion :

$$\boldsymbol{V} = \begin{bmatrix} 0,5 & 0,3 & 0,1 \\ 0,3 & 0,2 & 0,1 \end{bmatrix} \qquad \boldsymbol{W} = \begin{bmatrix} 0,1 & 0,2 \\ 0,3 & 0,4 \\ 0,5 & 0,6 \end{bmatrix}$$

Calculez les nouvelles valeurs de poids des matrices de connexion ***V*** et ***W*** après une passe complète de propagation directe - rétropropagation du gradient.
Les paramètres du réseau sont :

$$\eta = 1 \qquad f(net) = \frac{1}{1 + e^{-net}}$$

net = somme pondérée au niveau d’une cellule.

le stimulus $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ **doit donner la réponse** $\mathbf{t} = \begin{bmatrix} .1 \\ .3 \\ .7 \end{bmatrix}$

Cette activation est ensuite convertie en réponse. En utilisant la fonction logistique, on obtient:

$$\mathbf{h} = f(\mathbf{b}) = \begin{bmatrix} 0.8022 \\ 0.7311 \end{bmatrix} . \qquad \textbf{(VI.15)}$$

Cette activation est ensuite transmise aux cellules de la couche de sortie. Elles calculent leur activation :

$$\mathbf{a} = \mathbf{Zh} = \begin{bmatrix} 0.2264 \\ 0.5331 \\ 0.8397 \end{bmatrix} , \qquad \textbf{(VI.16)}$$

elles la transforment en réponse en utilisant la fonction logistique :

$$\mathbf{o} = f(\mathbf{a}) = \begin{bmatrix} 0.5564 \\ 0.6302 \\ 0.6984 \end{bmatrix} . \qquad \textbf{(VI.17)}$$

Les cellules de sortie peuvent évaluer leur *signal d'erreur*:

$$\begin{aligned} \boldsymbol{\delta}_{\text{sortie}} &= f'(\mathbf{a}) \circledast \mathbf{e} = \mathbf{o} \circledast (\mathbf{1} - \mathbf{o}) \circledast (\mathbf{t} - \mathbf{o}) \\ &= \begin{bmatrix} -0.1126 \\ -0.0770 \\ 0.0003 \end{bmatrix} . \end{aligned}$$

Réponse :

$$\boldsymbol{V} = \begin{bmatrix} 0,4946 & 0,2892 & 0,0837 \\ 0,2896 & 0,1791 & 0,0687 \end{bmatrix} \qquad \boldsymbol{W} = \begin{bmatrix} 0,0096 & 0,1177 \\ 0,2383 & 0,3437 \\ 0,5003 & 0,6002 \end{bmatrix}$$

Annexe : Algorithme de rétropropagation des erreurs

Pour chaque exemple d'entraînement

- Calculer les sorties du réseau
- Pour toutes les unités de sortie *k*, calculer l'erreur δ_k de la façon suivante:

$$\delta_k \leftarrow o_k(1-o_k)(t_k - o_k)$$

- Pour toutes les unités cachées *h*, calculer l'erreur δ_h de la façon suivante:

$$\delta_h \leftarrow o_h(1-o_h) \sum_{k \in sorties} w_{kh}\delta_k$$

- Mettre à jour tous les poids w_{ji} de la façon suivante:

$$w_{ji} \leftarrow w_{ji} + \Delta w_{ji}$$

ou

$$\Delta w_{ji} = \eta \delta_j x_{ji}$$