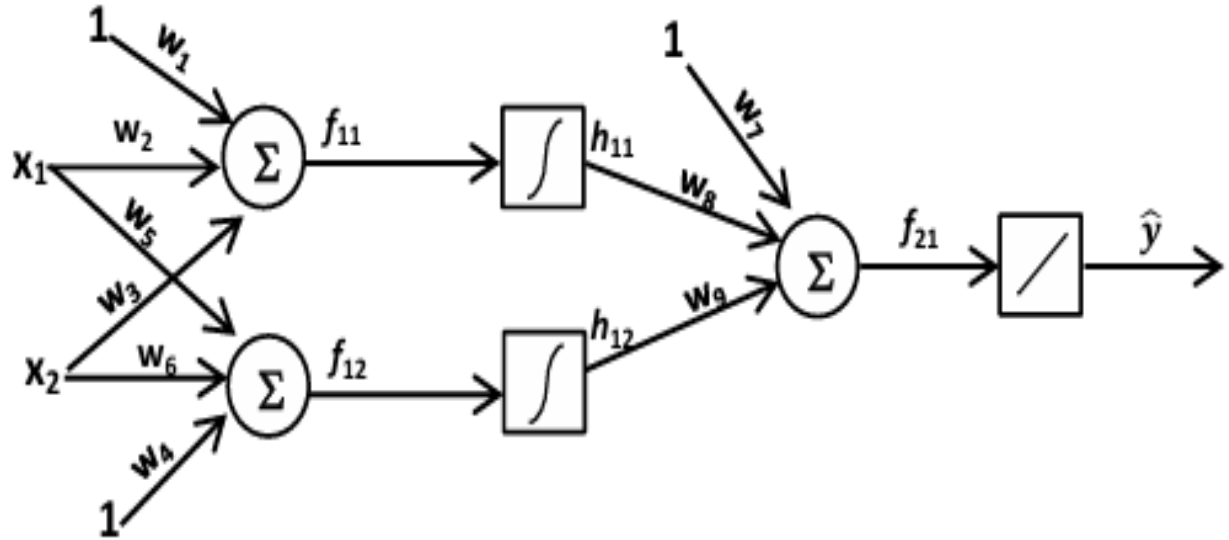
Centre Universitaire de Mila Institut: ST, Département: informatique

Apprentissage Automatique :Travaux dirigées

Année universitaire 2024-2025

Exercice01

Soit le réseau de neurones multicouches décrit par le graphe suivant :



1- Donner les formules mathématiques qui déterminent les sorties intermédiaires f_{11}, f_{12}, h_{11}, h_{12}, f_{21} ainsi que la sortie finale $\hat{y}$.

2- Soit la fonction d'erreur : $\mathbf{E}(\mathbf{w}) = (y - \hat{y})^2$

En appliquant l'algorithme de propagation en arrière (backpropagation), trouver les expressions des mises à jour des paramètres Δw_j pour **j = 1,…,9.**

Exercice01- SOLUTIONS

1- Propagation en avant (forward propagartion)

- $f_{11} = w_1 + w_2 x_1 + w_3 x_2$
- $f_{12} = w_4 + w_5 x_1 + w_6 x_2$
- $h_{11} = sigm(f_{11}) = \frac{1}{1+e^{-f_{11}}}$
- $h_{12} = sigm(f_{12}) = \frac{1}{1+e^{-f_{12}}}$
- $\boldsymbol{\hat{y} = f_{21} = w_7 + w_8 h_{11} + w_9 h_{12}}$

2- Propagation en arrière (backpropagartion algorithm)

La fonction d'erreur est donnée par $\mathbf{E(w)} = (\boldsymbol{y - \hat{y}})^2$

Donc, on aura $\frac{\partial E(w)}{\partial w_j} = -2(y - \hat{y})\frac{\partial \hat{y}}{\partial w_j}$

D'après la propagation en avant, on a : $\boldsymbol{\hat{y} = f_{21} = w_7 + w_8 h_{11} + w_9 h_{12}}$

Donc, les dérivées $\frac{\partial \hat{y}}{\partial w_j}$ peuvent être calculées par :

- $\frac{\partial \hat{y}}{\partial w_7} = 1$
- $\frac{\partial \hat{y}}{\partial w_8} = h_{11}$
- $\frac{\partial \hat{y}}{\partial w_9} = h_{12}$
- $\frac{\partial \hat{y}}{\partial w_1} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_1} = w_8 h_{11}(1 - h_{11})$
- $\frac{\partial \hat{y}}{\partial w_2} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_2} = w_8 h_{11}(1 - h_{11})x_1$

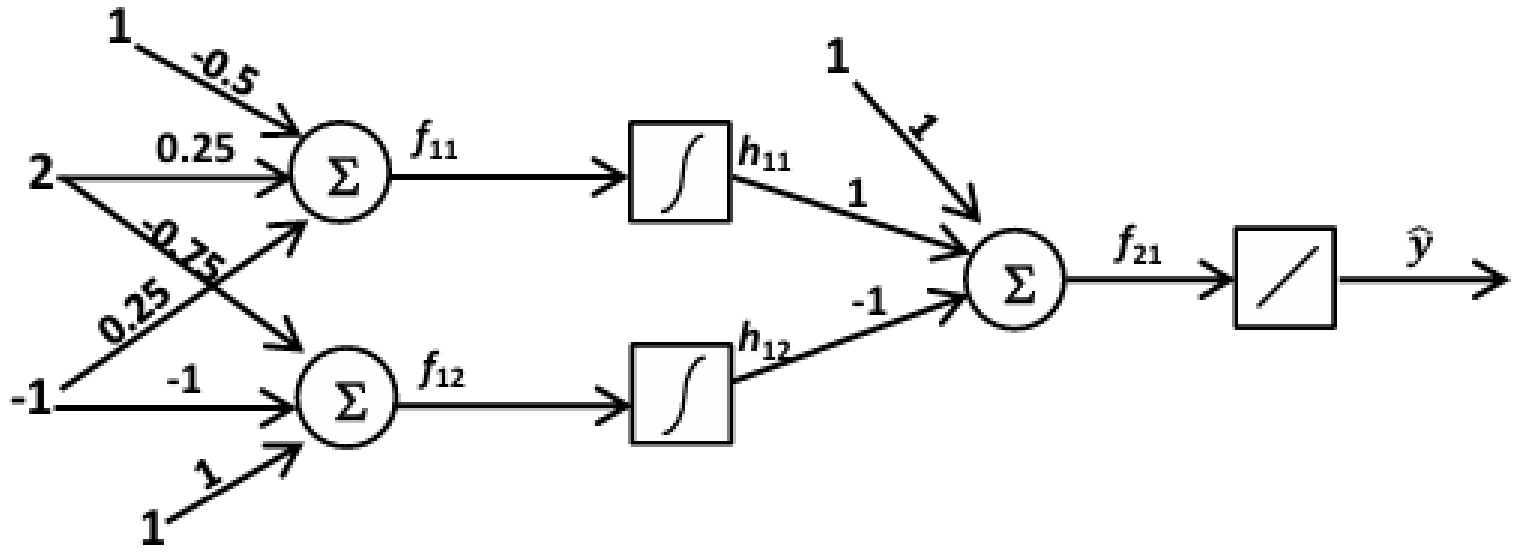
Exercice01- SOLUTIONS

- $\frac{\partial \hat{y}}{\partial w_3} = \frac{\partial \hat{y}}{\partial h_{11}} \frac{\partial h_{11}}{\partial f_{11}} \frac{\partial f_{11}}{\partial w_3} = w_8 h_{11}(1 - h_{11}) x_2$
- $\frac{\partial \hat{y}}{\partial w_4} = \frac{\partial \hat{y}}{\partial h_{12}} \frac{\partial h_{12}}{\partial f_{12}} \frac{\partial f_{12}}{\partial w_4} = w_9 h_{12}(1 - h_{12})$
- $\frac{\partial \hat{y}}{\partial w_5} = \frac{\partial \hat{y}}{\partial h_{12}} \frac{\partial h_{12}}{\partial f_{12}} \frac{\partial f_{12}}{\partial w_5} = w_9 h_{12}(1 - h_{12}) x_1$
- $\frac{\partial \hat{y}}{\partial w_6} = \frac{\partial \hat{y}}{\partial h_{12}} \frac{\partial h_{12}}{\partial f_{12}} \frac{\partial f_{12}}{\partial w_6} = w_9 h_{12}(1 - h_{12}) x_2$

En fin, la mise à jour de chaque paramétré est donnée par la formule : $\Delta w_j = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_j}$

Exercice02

- Soit le réseau de neurones multicouches de l'exercice 2 décrit par le graphe suivant :



- Soit la donnée **x = (2, -1), y = 1**

1- Calculer les sorties intermédiaires f_{11}, f_{12}, h_{11}, h_{12}, f_{21} ainsi que la sortie finale $\hat{y}$.

2- Calculer les mise à jour Δw_j ainsi que les paramètres w_j pour **j = 1,…,9** après u itération de mise à jour (en considérant le paramètre d'apprentissage $\alpha = 0.1$).

NB : La précision des calculs numériques est fixée à 4 chiffres après la virgule.

Exercice02- SOLUTIONS

1- Calcul des sorties intermédiaires $\mathbf{f_{11}}, \mathbf{f_{12}}, \mathbf{h_{11}}, \mathbf{h_{12}}, \mathbf{f_{21}}, \boldsymbol{et}\ \hat{\mathbf{y}}$:

- $\mathbf{f_{11}} = (-\mathbf{0.5}) * \mathbf{1} + \mathbf{0.25} * \mathbf{2} + \mathbf{0.25} * (-\mathbf{1}) = -\mathbf{0.25}$
- $\mathbf{f_{12}} = \mathbf{1} * \mathbf{1} + (-\mathbf{0.75}) * \mathbf{2} + (-\mathbf{1}) * (-\mathbf{1}) = \mathbf{0.50}$
- $\mathbf{h_{11}} = \mathbf{sigm}(\mathbf{f_{11}}) = \frac{\mathbf{1}}{\mathbf{1}+\mathbf{e}^{-\mathbf{f_{11}}}} = \frac{\mathbf{1}}{\mathbf{1}+\mathbf{e}^{-(-0.25)}} = \mathbf{0.4378}$
- $\mathbf{h_{12}} = \mathbf{sigm}(\mathbf{f_{21}}) = \frac{\mathbf{1}}{\mathbf{1}+\mathbf{e}^{-\mathbf{f_{21}}}} = \frac{\mathbf{1}}{\mathbf{1}+\mathbf{e}^{-0.5}} = \mathbf{0.6225}$
- $\mathbf{f_{21}} = \mathbf{1} * \mathbf{1} + \mathbf{1} * \mathbf{0.4378} + (-\mathbf{1}) * \mathbf{0.6225} = \mathbf{0.8154}$
- $\hat{\mathbf{y}} = \mathbf{f_{21}} = \mathbf{0.8154}$

Exercice02- SOLUTIONS

2- Calculer les paramètres w_j pour **j = 1, …, 9 :**

Soit le paramètre d'apprentissage $\alpha = 0.1$, On a $w_j = w_j + \Delta w_j$.

2-1- Calcul des $\frac{\partial \hat{y}}{\partial w_j}$ pour j = 1,…,9:

- $\frac{\partial \hat{y}}{\partial w_1} = 1 * 0.4378 * (1 - 0.4378) = 0.2461$
- $\frac{\partial \hat{y}}{\partial w_2} = 1 * 0.4378 * (1 - 0.4378) * 2 = 0.4923$
- $\frac{\partial \hat{y}}{\partial w_3} = 1 * 0.4378 * (1 - 0.4378) * (-1) = -0.2461$
- $\frac{\partial \hat{y}}{\partial w_4} = -1 * 0.6225 * (1 - 0.6225) = -0.2350$
- $\frac{\partial \hat{y}}{\partial w_5} = -1 * 0.6225 * (1 - 0.6225) * 2 = -0.4700$
- $\frac{\partial \hat{y}}{\partial w_6} = -1 * 0.6225 * (1 - 0.6225) * (-1) = 0.2350$
- $\frac{\partial \hat{y}}{\partial w_7} = 1$
- $\frac{\partial \hat{y}}{\partial w_8} = 0.4378$
- $\frac{\partial \hat{y}}{\partial w_9} = 0.6225$

2-2- Calcul des Δw_j pour j = 1,…,9:

- $\Delta w_1 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_1} = 0.1 * (1 - 0.8154) * 0.2461 = 0.0045$
- $\Delta w_2 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_2} = 0.1 * (1 - 0.8154) * 0.4923 = 0.0091$
- $\Delta w_3 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_3} = 0.1 * (1 - 0.8154) * (-0.2461) = -0.0045$
- $\Delta w_4 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_4} = 0.1 * (1 - 0.8154) * (-0.2350) = -0.0043$
- $\Delta w_5 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_5} = 0.1 * (1 - 0.8154) * (-0.4700) = -0.0087$
- $\Delta w_6 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_6} = 0.1 * (1 - 0.8154) * 0.2350 = 0.0043$
- $\Delta w_7 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_7} = 0.1 * (1 - 0.8154) * 1.0000 = 0.0185$
- $\Delta w_8 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_8} = 0.1 * (1 - 0.8154) * 0.4378 = 0.0081$
- $\Delta w_8 = \alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_8} = 0.1 * (1 - 0.8154) * 0.6225 = 0.0115$

Exercice02- SOLUTION

2-3- Calcul des w_j pour j = 1,…,9:

- $w_1 = w_1 + \Delta w_1 = -0.5 + 0.0045 = -0.4955$
- $w_2 = w_2 + \Delta w_2 = 0.25 + 0.0091 = 0.2591$
- $w_3 = w_3 + \Delta w_3 = 0.25 - 0.0045 = 0.2455$
- $w_4 = w_4 + \Delta w_4 = 1 - 0.0043 = 0.9957$
- $w_5 = w_5 + \Delta w_5 = -0.75 - 0.0087 = -0.7587$
- $w_6 = w_6 + \Delta w_6 = -1 + 0.0043 = -0.9957$
- $w_7 = w_7 + \Delta w_7 = 1 + 0.0185 = 1.0185$
- $w_8 = w_8 + \Delta w_8 = 1 + 0.0081 = 1.0081$
- $w_8 = w_8 + \Delta w_8 = -1 + 0.0115 = -0.9885$