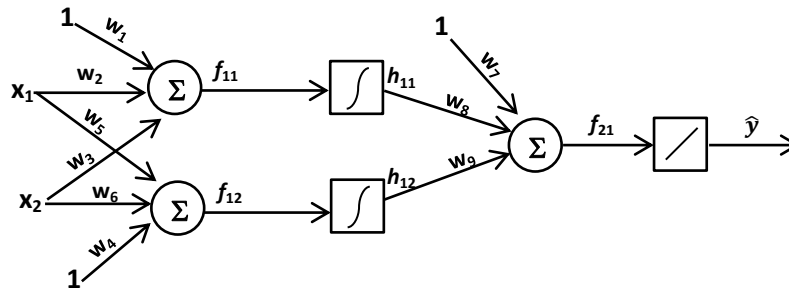


Exercice 1 : (Examen Final 2019)

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



- Donner les formules mathématiques qui déterminent les sorties intermédiaires f_{11} , f_{12} , h_{11} , h_{12} , f_{21} ainsi que la sortie finale $\hat{y}$.
- Soit la fonction d'erreur : $E(\mathbf{w}) = (\mathbf{y} - \hat{\mathbf{y}})^2$
En appliquant l'algorithme de propagation en arrière (backpropagation), trouver les expressions des mises à jour des paramètres $\Delta \mathbf{w}_j$ pour $j = 1, \dots, 9$.

Solution :

- Propagation en avant (forward propagation)

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

- Propagation en arrière (backpropagation algorithm)

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

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

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

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

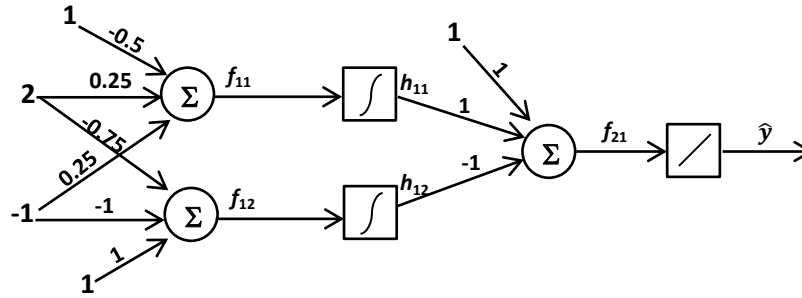
- $\frac{\partial \hat{y}}{\partial w_7} = 1$
- $\frac{\partial \hat{y}}{\partial w_8} = h_{11}$
- $\frac{\partial \hat{y}}{\partial w_9} = h_{12}$
- $\frac{\partial \hat{y}}{\partial w_1} = \frac{\partial \hat{y}}{\partial h_{11}} \frac{\partial h_{11}}{\partial f_{11}} \frac{\partial f_{11}}{\partial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

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

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

Exercice 2 : (Examen Final 2019)

- 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 $j = 1, \dots, 9$ après une 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.

Solution :

- Calcul des sorties intermédiaires f_{11} , f_{12} , h_{11} , h_{12} , f_{21} , et $\hat{y}$:
 - $f_{11} = (-0.5) * 1 + 0.25 * 2 + 0.25 * (-1) = -0.25$
 - $f_{12} = 1 * 1 + (-0.75) * 2 + (-1) * (-1) = 0.50$
 - $h_{11} = \text{sigm}(f_{11}) = \frac{1}{1+e^{-f_{11}}} = \frac{1}{1+e^{-(-0.25)}} = 0.4378$
 - $h_{12} = \text{sigm}(f_{12}) = \frac{1}{1+e^{-f_{12}}} = \frac{1}{1+e^{-0.5}} = 0.6225$
 - $f_{21} = 1 * 1 + 1 * 0.4378 + (-1) * 0.6225 = 0.8154$
 - $\hat{y} = f_{21} = 0.8154$
- Calculer les paramètres w_j pour $j = 1, \dots, 9$:

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 $j = 1, \dots, 9$:

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

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

2-2- Calcul des Δw_j pour $j = 1, \dots, 9$:

- $\Delta w_1 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_1} = 0.1 * (1 - 0.8154) * 0.2461 = 0.0045$
- $\Delta w_2 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_2} = 0.1 * (1 - 0.8154) * 0.4923 = 0.0091$
- $\Delta w_3 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_3} = 0.1 * (1 - 0.8154) * (-0.2461) = -0.0045$
- $\Delta w_4 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_4} = 0.1 * (1 - 0.8154) * (-0.2350) = -0.0043$
- $\Delta w_5 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_5} = 0.1 * (1 - 0.8154) * (-0.4700) = -0.0087$
- $\Delta w_6 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_6} = 0.1 * (1 - 0.8154) * 0.2350 = 0.0043$
- $\Delta w_7 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_7} = 0.1 * (1 - 0.8154) * 1.0000 = 0.0185$
- $\Delta w_8 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_8} = 0.1 * (1 - 0.8154) * 0.4378 = 0.0081$
- $\Delta w_9 = \alpha(y - \hat{y}) \frac{\partial \hat{y}}{\partial w_9} = 0.1 * (1 - 0.8154) * 0.6225 = 0.0115$

2-3- Calcul des w_j pour $j = 1, \dots, 9$:

- $w_1 = w_1 + \Delta w_1 = -0.5 + 0.0045 = -0.4955$
- $w_2 = w_2 + \Delta w_2 = 0.25 + 0.0091 = 0.2591$
- $w_3 = w_3 + \Delta w_3 = 0.25 - 0.0045 = 0.2455$
- $w_4 = w_4 + \Delta w_4 = 1 - 0.0043 = 0.9957$
- $w_5 = w_5 + \Delta w_5 = -0.75 - 0.0087 = -0.7587$
- $w_6 = w_6 + \Delta w_6 = -1 + 0.0043 = -0.9957$
- $w_7 = w_7 + \Delta w_7 = 1 + 0.0185 = 1.0185$
- $w_8 = w_8 + \Delta w_8 = 1 + 0.0081 = 1.0081$
- $w_9 = w_9 + \Delta w_9 = -1 + 0.0115 = -0.9885$

Exercice 3 (Rattrapage 2018)

Soit **A** et **B** deux variables booléennes.

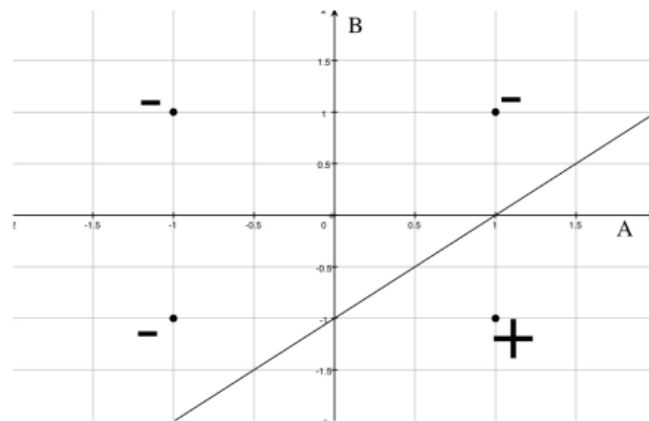
- 1) Concevoir un réseau de neurones à deux entrées permettant d'implémenter la fonction booléenne $A \wedge \neg B$.
- 2) Concevoir un réseau de neurones à deux couches implémentant la fonction booléenne **A XOR B**.

Solution :

- 1) Le perceptron demandé a **3** entrées : **A**, **B** et la constante **1**. Les valeurs de **A** et **B** sont **1** (vrai) ou **-1** (faux). Le tableau suivant décrit la sortie **y** du perceptron :

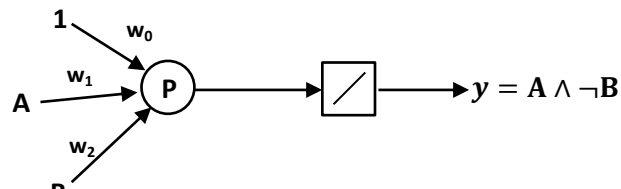
A	B	$y = A \wedge \neg B$
-1	-1	-1
-1	1	-1
1	-1	1
1	1	-1

Une des surfaces de décision correctes (n'importe quelle droite séparant le point positif des points négatifs serait correcte) est affichée dans la figure suivante :



La droite traversant l'axe **A** en **1** et l'axe **B** en **-1**. Donc l'équation de la droite est

$$\frac{A - 0}{1 - 0} = \frac{B - (-1)}{0 - (-1)} \Rightarrow A = B + 1 \Rightarrow 1 - A + B = 0$$

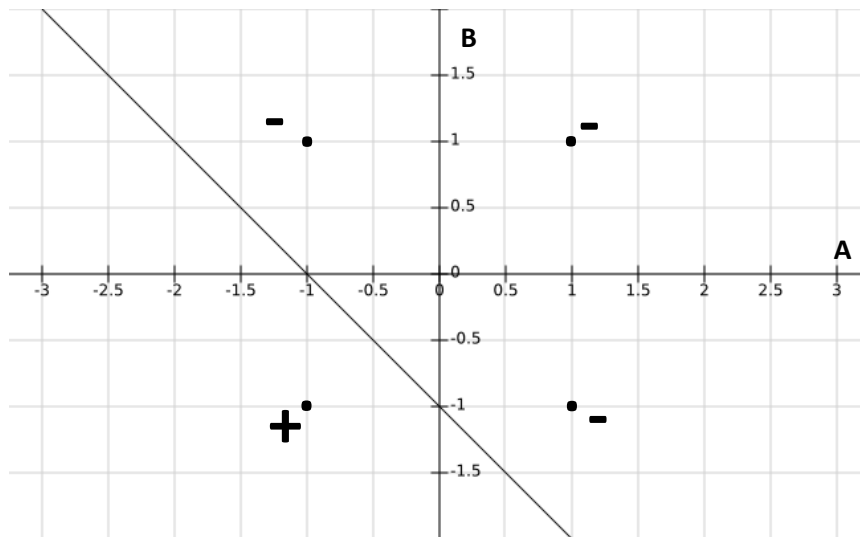


Les valeurs possibles pour les poids **w₀**, **w₁** et **w₂** sont **1** et **-1**. En utilisant ces valeurs, la sortie du perceptron pour **A = 1**, **B = -1** est positive. Par conséquent, nous pouvons conclure que **w₀ = -1**, **w₁ = 1**, **w₂ = -1**.

- 2) Le perceptron demandé a **3** entrées : **A**, **B** et la constante **1**. Les valeurs de **A** et **B** sont **1** (vrai) ou **-1** (faux). Le tableau suivant décrit la sortie **y** du perceptron :

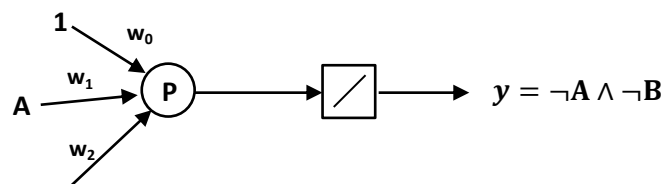
A	B	$y = \neg A \wedge \neg B$
-1	-1	1
-1	1	-1
1	-1	-1
1	1	-1

Une des surfaces de décision correctes (n'importe quelle droite séparant les points positifs des points négatifs serait correcte) est affichée dans la figure suivante :



La droite traversant l'axe **A** en -1 et l'axe **B** en -1 . Donc l'équation de la droite est

$$\frac{A - 0}{-1 - 0} = \frac{B - (-1)}{0 - (-1)} \Rightarrow A = -B - 1 \Rightarrow 1 + A + B = 0$$



Les valeurs possibles pour les poids w_0 , w_1 et w_2 sont 1 et -1 . En utilisant ces valeurs, la sortie du perceptron pour $A = -1$, $B = -1$ est positive. Par conséquent, nous pouvons conclure que $w_0 = -0.7$, $w_1 = -0.5$, $w_2 = -0.5$.

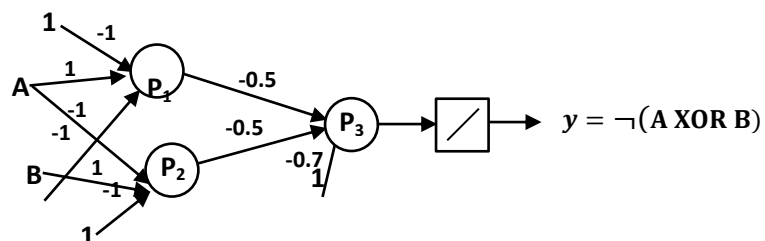
3) $\neg(A \text{ XOR } B)$ ne peut pas être calculé par un perceptron unique, nous devons donc construire un réseau à deux couches de perceptrons. La structure du réseau peut être dérivée par :

- Exprimer $\neg(A \text{ XOR } B)$ en fonction des composantes logiques :

$$\neg(A \text{ XOR } B) = \neg((A \wedge \neg B) \vee (\neg A \wedge B)) = \neg(A \wedge \neg B) \wedge \neg(\neg A \wedge B)$$

- Définir les perceptrons P_1 et P_2 pour $(A \wedge \neg B)$ et $(\neg A \wedge B)$
- Combiner les sorties de P_1 et P_2 dans P_3 qui implémente $\neg f(P_1) \wedge \neg f(P_2)$

En fin, le réseau demandé est donné dans la figure suivante :



Exemple :

Soit $A = 1$, $B = -1$

Résultat pour le perceptron $P_1 = 1 + 1 - 1 = 1$

Résultat pour le perceptron $P_2 = -1 - 1 - 1 = -3$

Résultat pour le perceptron $P_3 = -0.5 \cdot 1 - 0.5 \cdot (-3) - 0.7 = -1.7 \Rightarrow$ le résultat final est faux ($y = -1$).