Centre Universitaire de Mila	Examen de rattrapage en « Apprentissage Automatique »
Institut : ST, Département : MI	Durée de l'examen : 1h30 / Mila le : 13-06-2018
Master 1 – STIC	Documents non autorisés

Exercice 1 (8 points).

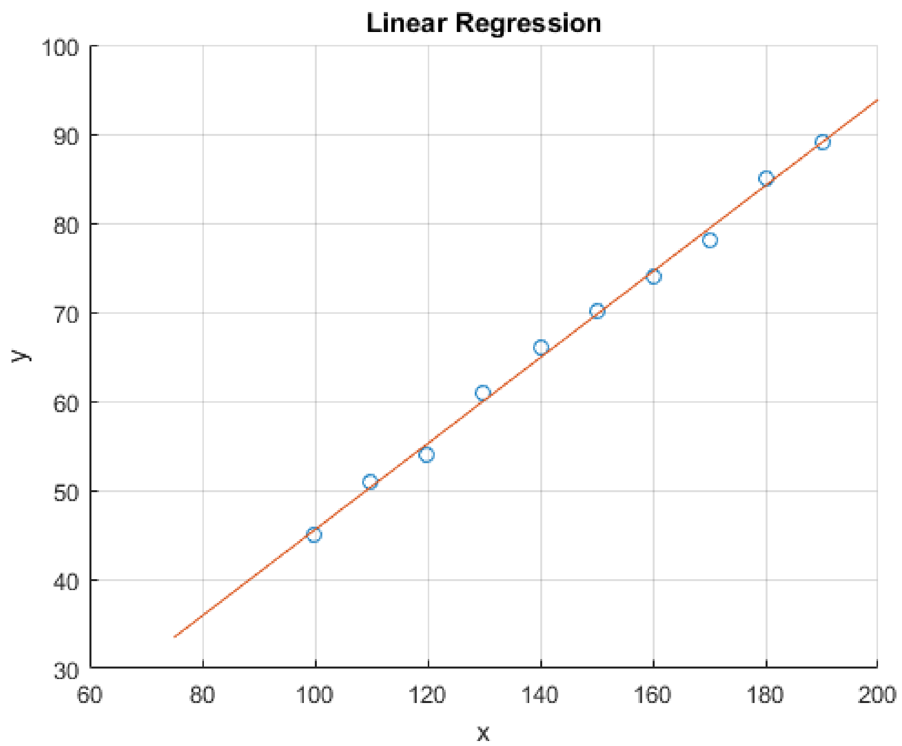
L'analyse de la température de fonctionnement d'un procédé chimique sur le rendement du produit a donné les valeurs suivantes pour la température et le rendement correspondant :

Température °C	Rendement %	Température °C	Rendement %
100	45	150	70
110	51	160	74
120	54	170	78
130	61	180	85
140	66	190	89

1) Donner une représentation graphique de ces données.
2) Trouver la fonction de régression linéaire par la méthode des moindres carrés, qui permet d'associer à la température la valeur de rendement correspondante.
3) Utilisant cette fonction de régression, prédire la valeur de rendement pour la température 80°C.
4) Déterminer (en utilisant la droite de régression) quand la valeur de rendement sera supérieure à 100.

Solution :

1) Représentation graphique de ces données.



2) Modèle de régression linéaire :

On a

$$x = [100; 110; 120; 130; 140; 150; 160; 170;\ 180; 190];$$

et l'étiquette y

$$y = [45;\ 51;\ 54;\ 61;\ 66;\ 70;\ 74;\ 78; 85; 89];$$

Application du modèle de régression linéaire :

Dans le modèle de la régression linéaire, l'ensemble des paramètres est calculé par la formule suivante :

$$\widetilde{\mathbf{w}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{y}$$

Application numérique sur les données de l'exercice :

$$\boldsymbol{X} = \begin{bmatrix} 100 & 1 \\ 110 & 1 \\ 120 & 1 \\ 130 & 1 \\ 140 & 1 \\ 150 & 1 \\ 160 & 1 \\ 170 & 1 \\ 180 & 1 \\ 190 & 1 \end{bmatrix} \quad \boldsymbol{y} = \begin{bmatrix} 45 \\ 51 \\ 54 \\ 61 \\ 66 \\ 70 \\ 74 \\ 78 \\ 85 \\ 89 \end{bmatrix}$$

$$\boldsymbol{X}^T\boldsymbol{X} = \begin{bmatrix} 100 & 110 & 120 & 130 & 140 & 150 & 160 & 170 & 180 & 190 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 100 & 1 \\ 110 & 1 \\ 120 & 1 \\ 130 & 1 \\ 140 & 1 \\ 150 & 1 \\ 160 & 1 \\ 170 & 1 \\ 180 & 1 \\ 190 & 1 \end{bmatrix} = \begin{bmatrix} 218500 & 1450 \\ 1450 & 10 \end{bmatrix}$$

$$\boldsymbol{X}^T\boldsymbol{y} = \begin{bmatrix} 100 & 110 & 120 & 130 & 140 & 150 & 160 & 170 & 180 & 190 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 45 \\ 51 \\ 54 \\ 61 \\ 66 \\ 70 \\ 74 \\ 78 \\ 85 \\ 89 \end{bmatrix} = \begin{bmatrix} 101570 \\ 673 \end{bmatrix}$$

On a $\boldsymbol{det}(\boldsymbol{X}^T\boldsymbol{X}) = \mathbf{82500}$**,** donc $(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \begin{bmatrix} 10/82500 & -1450/82500 \\ -1450/82500 & 218500/82500 \end{bmatrix} =$

$$\begin{bmatrix} 0.0001 & -0.0176 \\ -0.0176 & 2.6485 \end{bmatrix}$$

D'où $\widetilde{\mathbf{w}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \begin{bmatrix} 0.4830 \\ -2.7394 \end{bmatrix}$

$$\boldsymbol{y} = -\mathbf{2.7394} + \mathbf{0.4830}\boldsymbol{x}$$

3) Prédire la valeur de rendement pour la température 80°C :

$$\boldsymbol{y}_{\mathbf{80°C}} = -\mathbf{2.7394} + \mathbf{0.4830} * \mathbf{80} = \mathbf{35.9006}$$

4) La valeur de rendement sera supérieure à 100 si on a :

$$\boldsymbol{y} > \mathbf{100} \Rightarrow -\mathbf{2.7394} + \mathbf{0.4830}\boldsymbol{x} > \mathbf{100}$$

$$\Rightarrow \boldsymbol{x} > \frac{\mathbf{100} + \mathbf{2.7394}}{\mathbf{0.4830}}$$

$$\Rightarrow \boldsymbol{x} \geq \mathbf{212°C}$$

Exercice 2 (6 points).

Considérons un problème de classification binaire. Dans l'ensemble d'apprentissage, il y a 100 instances de classe $\mathbf{C_1}$ et 80 instances de classe $\mathbf{C_2}$. Supposons que, pour le classificateur appris, nous avons la matrice de confusion suivante liée à la classe $\mathbf{C_1}$.

Classe C_1		**La classification du classificateur appris**		
		C_1	**C_2**	**Total**
Le classement réelle	**C_1**	90	10	100
	C_2	20	60	80

1- Donner une définition de **la précision**, **le rappel** et de **la mesure F** par rapport à la classe c1.
2- Effectuer le calcul des 3 mesures d'évaluation.

Solution :

L'explication de la précision, du rappel et de la mesure F est en cours.

$$Prec(C_1) = \frac{TP_1}{TP_1 + FP_1} = \frac{90}{90 + 20} = 0.82$$

$$Rap(C_1) = \frac{TP_1}{TP_1 + FN_1} = \frac{90}{90 + 10} = 0.9$$

$$F(C_1) = \frac{2 * Prec(C_1) * Rap(C_1)}{Prec(C_1) + Rap(C_1)} = \frac{2 * 0.82 * 0.9}{0.82 + 0.9} = 0.86$$

Exercice 3 (6 points).

Soit $\mathbf{A}$ et $\mathbf{B}$ deux variables booléennes.

1) Concevoir un réseau de neurones à deux entrées permettant d'implémenter la fonction booléenne $\mathbf{A \land \neg B}$.

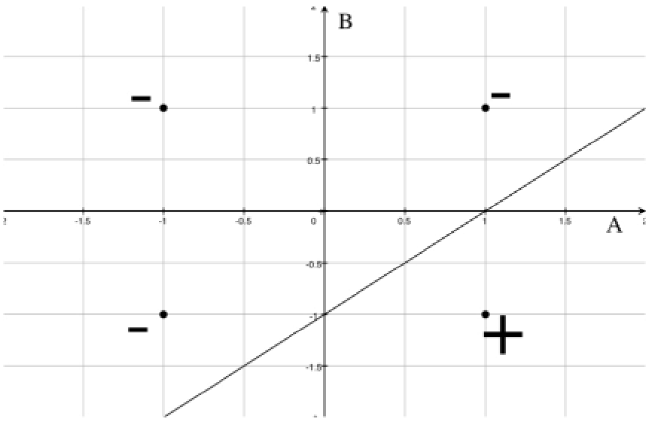
2) Concevoir un réseau de neurones à deux couches implémentant la fonction booléenne $\mathbf{A\ XOR\ B}$.

Solution :

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

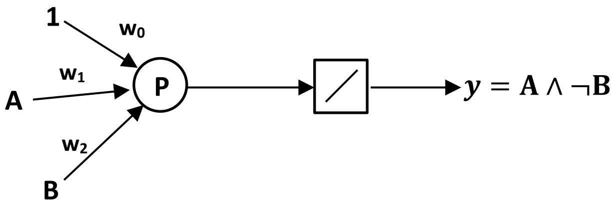
A	**B**	$\mathbf{y = A \land \neg B}$
-1	**-1**	**-1**
-1	**1**	**-1**
1	**-1**	**1**
1	**1**	**-1**

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 $\mathbf{A}$ en $\mathbf{1}$ et l'axe $\mathbf{B}$ en $-\mathbf{1}$. Donc l'équation de la droite est

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

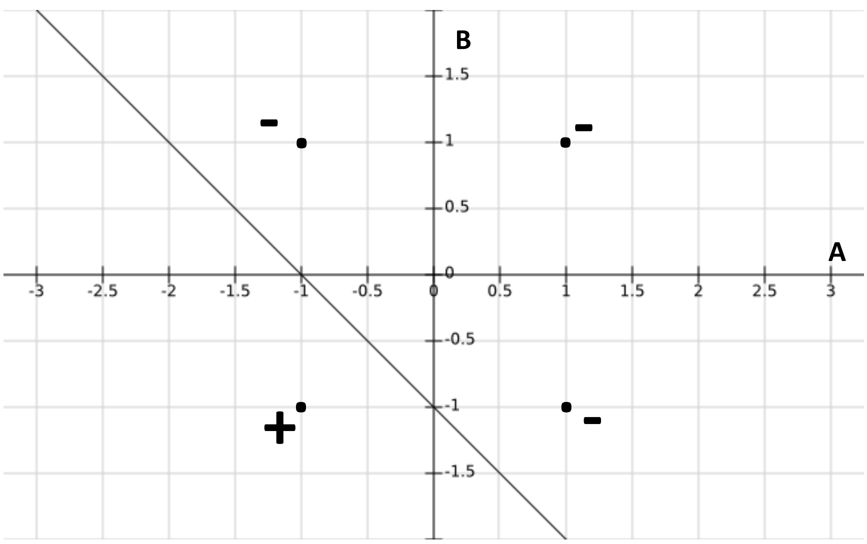


Les valeurs possibles pour les poids $\mathbf{w_0}$, $\mathbf{w_1}$ et $\mathbf{w_2}$ sont **1** et **-1**. En utilisant ces valeurs, la sortie du perceptron pour **A = 1, B = -1** est positive. Par conséquent, nous pouvons conclure que $\mathbf{w_0 = -1, w_1 = 1, w_2 = -1}$.

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

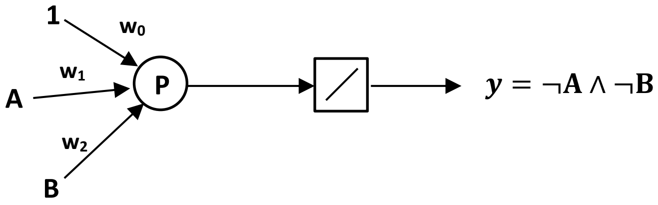
$\mathbf{A}$	$\mathbf{B}$	$\mathbf{y = \neg A \wedge \neg B}$
-1	**-1**	**1**
-1	**1**	**-1**
1	**-1**	**-1**
1	**1**	**-1**

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 $\mathbf{A}$ en $-\mathbf{1}$ et l'axe $\mathbf{B}$ en $-\mathbf{1}$. Donc l'équation de la droite est

$$\frac{\boldsymbol{A}-\mathbf{0}}{-\mathbf{1}-\mathbf{0}}=\frac{\boldsymbol{B}-(-\mathbf{1})}{\mathbf{0}-(-\mathbf{1})} \Longrightarrow \boldsymbol{A}=-\boldsymbol{B}-\mathbf{1} \Longrightarrow \mathbf{1}+\boldsymbol{A}+\boldsymbol{B}=\mathbf{0}$$



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.**

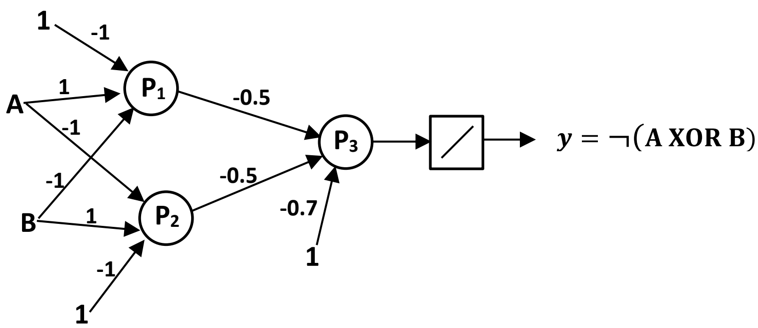
3) $\neg(\mathbf{A\ XOR\ B})$ ne peut pas être calculé par un perceptron unique, nous devons donc construire un réseau à deux couches de perceptrons. La structure du réseau peut être dérivée par :

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

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

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

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



<u>Exemple :</u>

Soit A = 1, B = -1

Résultat pour le perceptron P1 = 1 + 1 - 1 = 1

Résultat pour le perceptron P2 = -1 - 1 - 1 = -3

Résultat pour le perceptron P3 = -0.5*1 -0.5*-1 – 0.7 = -1.7 ⇒ le résultat final est faux (y = -1).