

# TD Deep learning

## Perceptron et apprentissage : premiers pas

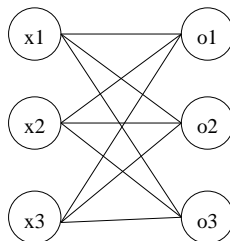
Ce premier exercice permet de vous familiariser avec le calcul *feed-forward* et de réaliser à la main une première étape de révision des poids du réseau, avant de commencer à l'automatiser.

Soit le tableau suivant décrivant une série de mesures réalisées sur 7 individus :

Individu	$x_1$	$x_2$	$x_3$	type
$i_1$	0.2	0.4	0.05	1
$i_2$	0.17	0.42	0.3	2
$i_3$	0.46	0.38	0.44	3
$i_4$	0.33	0.44	0.27	2
$i_5$	0.12	0.44	0.11	1
$i_6$	0.82	0.37	0.75	3
$i_7$	0.47	0.39	0.29	2

A partir de ces données, répondez aux questions suivantes :

1. Soit le perceptron sans couche cachée suivant :



Le tableau ci-dessous donne les poids du réseau,  $w_{ij}$  correspondant au poids de l'arc reliant la cellule  $x_i$  au neurone de sortie  $o_j$ , ainsi que les seuils  $\theta_i$  pour chaque neurone :

$w_{11}$	$w_{12}$	$w_{13}$	$w_{21}$	$w_{22}$	$w_{23}$	$w_{31}$	$w_{32}$	$w_{33}$	$\theta_1$	$\theta_2$	$\theta_3$
-0.4	0.1	0.5	0.1	-0.1	-0.1	-0.6	0.2	0.4	-0.1	0	0.3

En supposant que la fonction d'activation de ces 3 neurones est une fonction à seuil, dont le seuil  $\theta_i$  est précisé ci-dessus, calculez la sortie de cette machine apprenante pour les individus  $i_k$ ,  $k = 1, \dots, 7$ .

2. Que pouvez-vous en déduire sur la pertinence de cette machine ?
3. Donnez la matrice de confusion calculée sur ces 7 individus, en considérant qu'un individu classé dans plusieurs catégories est finalement associé à la mauvaise classe (approche pessimiste).
4. Comment calculez-vous l'erreur apparente à partir de cette matrice ? Donnez la valeur obtenue.
5. En utilisant un pas  $\lambda = 0.1$ , utilisez l'algorithme d'apprentissage vu en cours afin d'essayer d'améliorer les résultats obtenus par le réseau. Attention : vous n'effectuerez qu'une seule passe sur les exemples, ce qui veut dire que chaque exemple n'est considéré qu'une seule fois pour chacun des neurones.
6. Quelle est la nouvelle matrice de confusion et la nouvelle erreur ? Qu'en déduisez-vous ?
7. Comment pouvez-vous modifier le réseau pour obtenir la probabilité  $p(c/i_n)$  qu'un nouvel objet  $i_n$  soit classé dans l'une ou l'autre de ces trois classes  $c$  ?
8. A présent, implémentez avec le langage R ou Python la procédure d'apprentissage afin de pouvoir réaliser plusieurs pas et observer les résultats.