

Plan

- ➊ Présentation et exemples
- ➋ Classification par réseaux multicouches
- ➌ Apprentissage des réseaux multicouches

Plan

- ➊ Présentation et exemples
- ➋ Classification par réseaux multicouches
- ➌ Apprentissage des réseaux multicouches

L'analogie biologique

Le cerveau est un organe caractérisé par l'*interconnexion* d'un nombre élevé d'unités de traitement simples : les *neurones*.

Le comportement de ce *réseau de neurones* est déterminé par son *architecture* (le nombre des unités et la manière dont elles sont connectées) et par les poids affecté à chaque *connexion* ou *synapses*.

Le cerveau humain comporte environ 10^{11} neurones, chacun relié en moyenne à plus de 1000 autres.

Du biologique à l'artificiel

➔ Activation

Une perception est reçue sur les neurones capteurs.

Elle se propage dans le réseau selon la règle suivante : quand la somme des signaux reçus en entrée par un neurone dépasse un seuil, il émet à son tour un signal vers les neurones auxquels il est relié.

En sortie, une action résulte du calcul du réseau.

➔ Apprentissage

La capacité d'apprentissage de ces systèmes est liée à *modification des poids des connexions* à partir d'exemples.

C'est sur ce principe que sont basées les modélisations du cerveau sous la forme de *réseaux de neurones artificiels*, ou *réseaux connexionnistes* pour leur application à l'Apprentissage.

Le neurone formel.

Le neurone formel d'indice j est caractérisé par :

↳ l'ensemble de ses entrées

$$\mathbf{y}_i \text{ avec } i \in \text{source}(j)$$

↳ le poids des connexions qui mènent vers lui

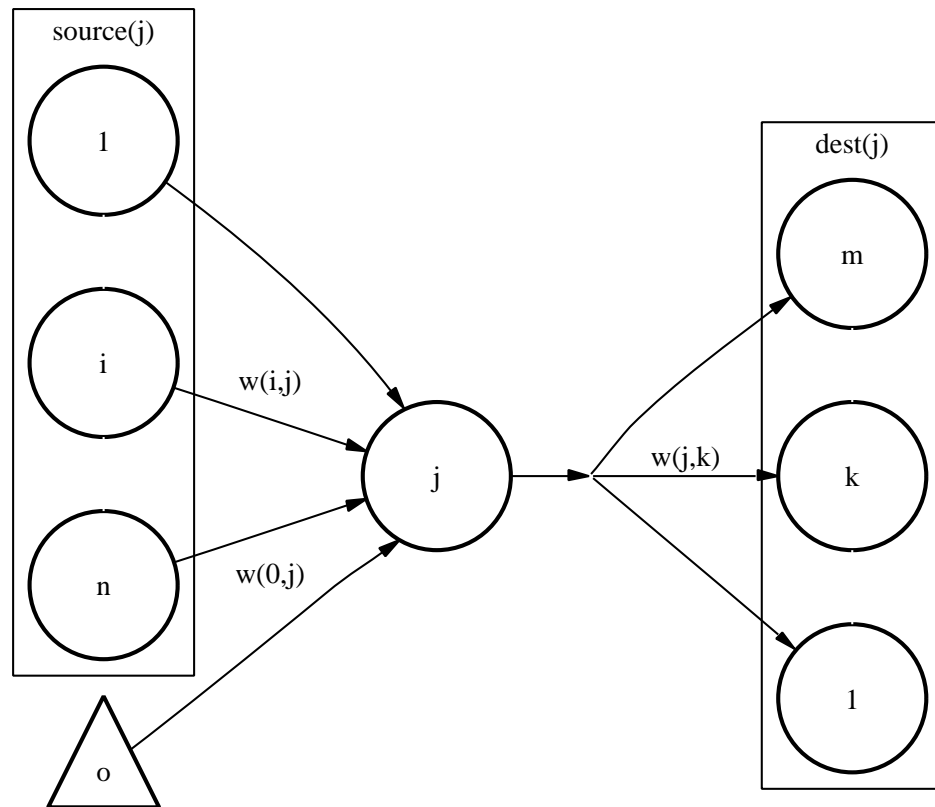
$$w(i, j) \text{ avec } i \in \text{source}(j)$$

↳ son état d'activation

$$\sigma_j = w(0, j) + \sum_{i \in \text{source}(j)} w(i, j) \mathbf{y}_i$$

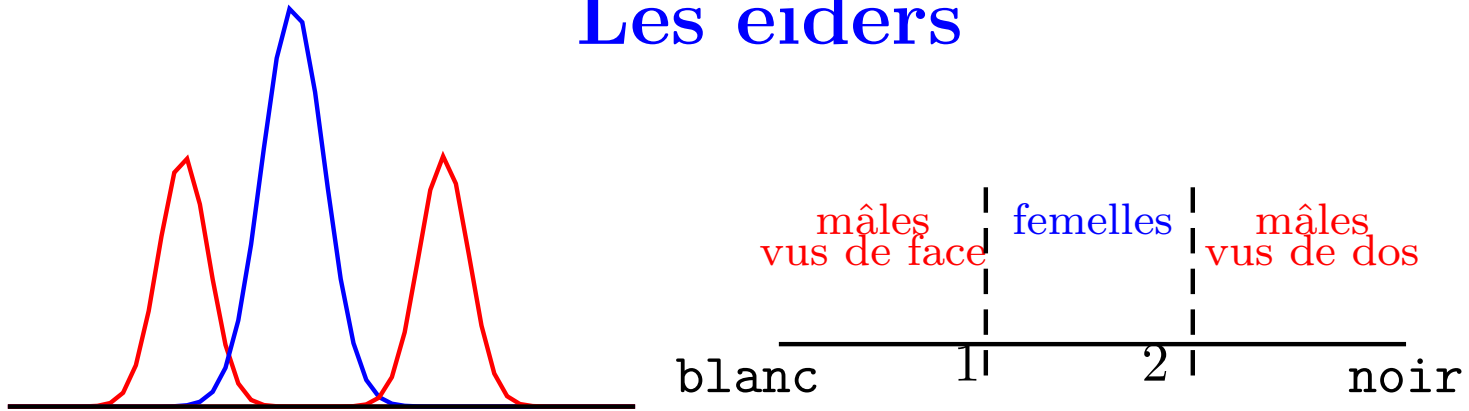
↳ sa fonction de sortie

$$\mathbf{y}_j = f(\sigma_j)$$



Le neurone formel : description et notations.

Les eiders



Répartition des eiders sur l'axe x : blanc / noir.

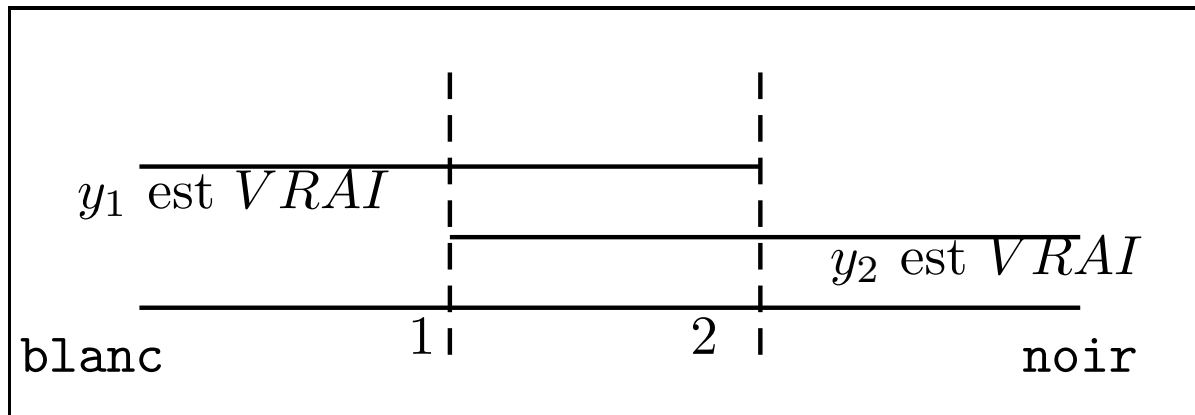
Notons x le niveau de gris d'un eider dont on cherche à déterminer le sexe. Les connaissances sont les suivantes :

➔ *Si $x \leq 1$ ou $x \geq 2$ alors mâle.*

➔ *Si $1 \leq x \leq 2$ alors femelle.*

Les eiders : variables intermédiaires

	$x \leq 1$	$1 \leq x \leq 2$	$x \geq 2$
y_1	<i>VRAI</i>	<i>VRAI</i>	<i>FAUX</i>
y_2	<i>FAUX</i>	<i>VRAI</i>	<i>VRAI</i>



Deux décisions élémentaires à combiner.

Les eiders

Ce problème a été présenté dans un cadre *logique*, mais il est facile de le transformer en un problème de décision *numérique*.

Il suffit d'attribuer les valeurs réelles 0 et 1 à y_1 et y_2 et au concept selon qu'ils sont *VRAI* ou *FAUX*.

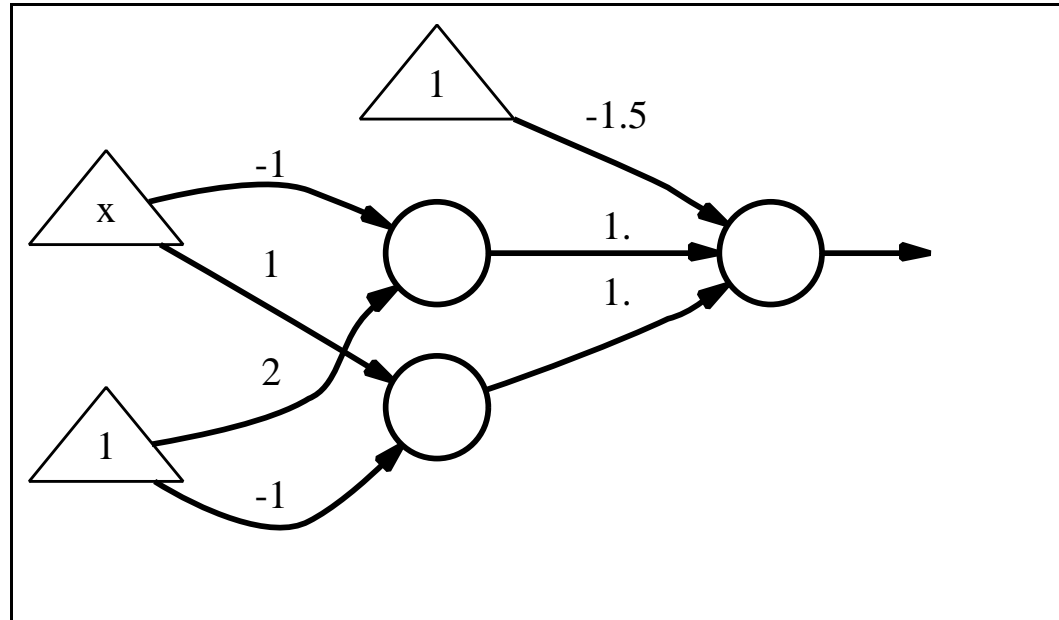
Il est en effet peu intuitif, mais immédiat de vérifier que la valeur numérique z

$$z = y_1 + y_2 - 1.5$$

est positive quand l'eider est femelle (elle vaut $z = 1 + 1 - 1.5 = 0.5$) et négative quand il est mâle (elle vaut $z = 1 + 0 - 1.5 = -0.5$ ou $z = 0 + 1 - 1.5 = -0.5$).

Donc, le concept est aussi exprimé par le signe de la valeur z .

Les eiders



Un réseau connexionniste qui distingue les eiders mâles des eiders femelles. La fonction de sortie des neurones formels est la fonction signe à valeurs 0 ou 1. La sortie 0 indique un eider mâle, la sortie 1 un eider femelle.

Plan

- ① Présentation et exemples
- ② Classification par réseaux multicouches
- ③ Apprentissage des réseaux multicouches

Architecture : les réseaux multicouches.

- ➔ Les neurones formels, ou *unités* d'une couche sont reliées à toutes celles de la couche supérieure, mais à aucune autre.
- ➔ À la connexion entre les unités i et j , on associe le *poids* $w(i, j)$.
- ➔ La *couche d'entrée* ne sert qu'à la lecture des données. Elle est composée de $d + 1$ éléments. Elle est activée par un vecteur \mathbf{x} de \mathbb{R} .
- ➔ La première *couche cachée* effectue le calcul pour chacun de ses neurones formels, puis la seconde couche fait ses calculs, etc...
- ➔ La *couche de sortie* sert à traduire la décision. Pour une classification, le neurone de la couche finale ayant la valeur de sortie la plus grande indique la classe calculée pour l'entrée.

Calculs de décision

Chaque neurone formel se met dans l'état :

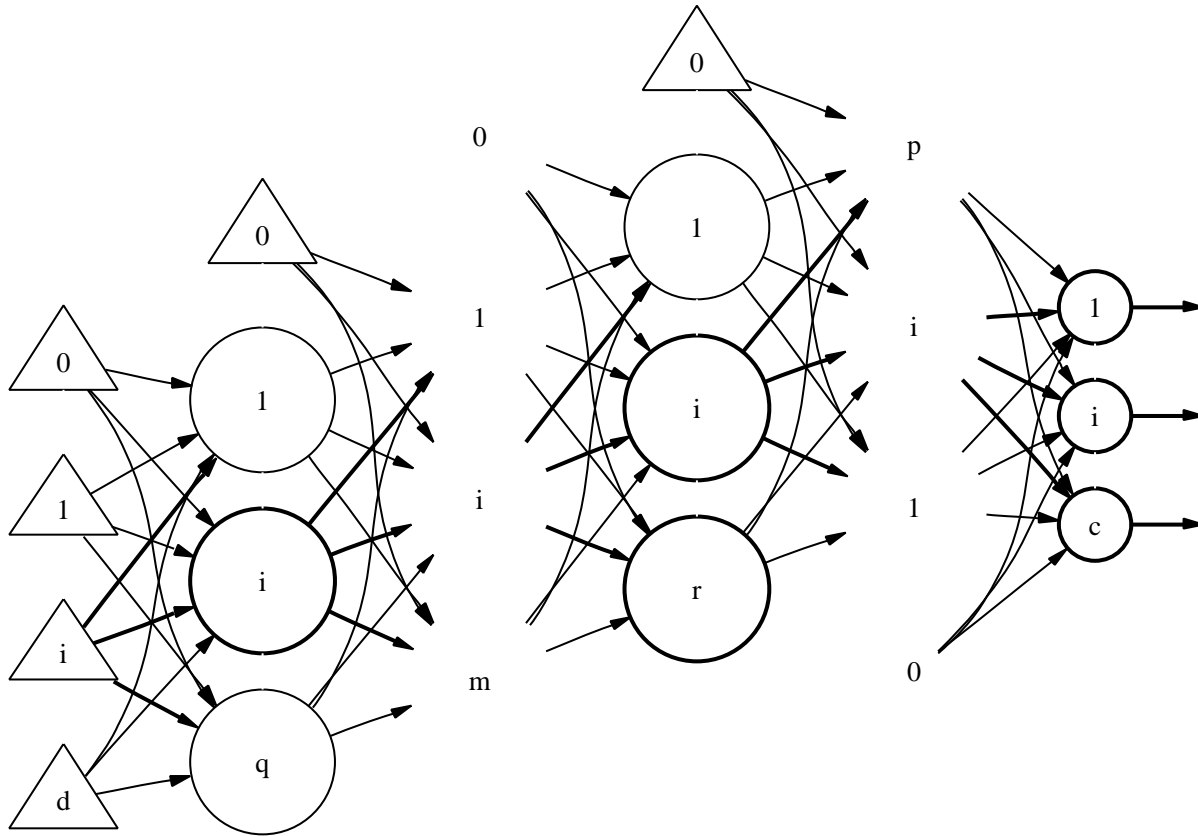
$$\sigma_j = \sum_{i \in source(j)} w(i, j) \mathbf{y}_i$$

puis transmet vers les autres neurones formels de $\{dest(j)\}$:

$$\mathbf{y}_i = f(\sigma_i)$$

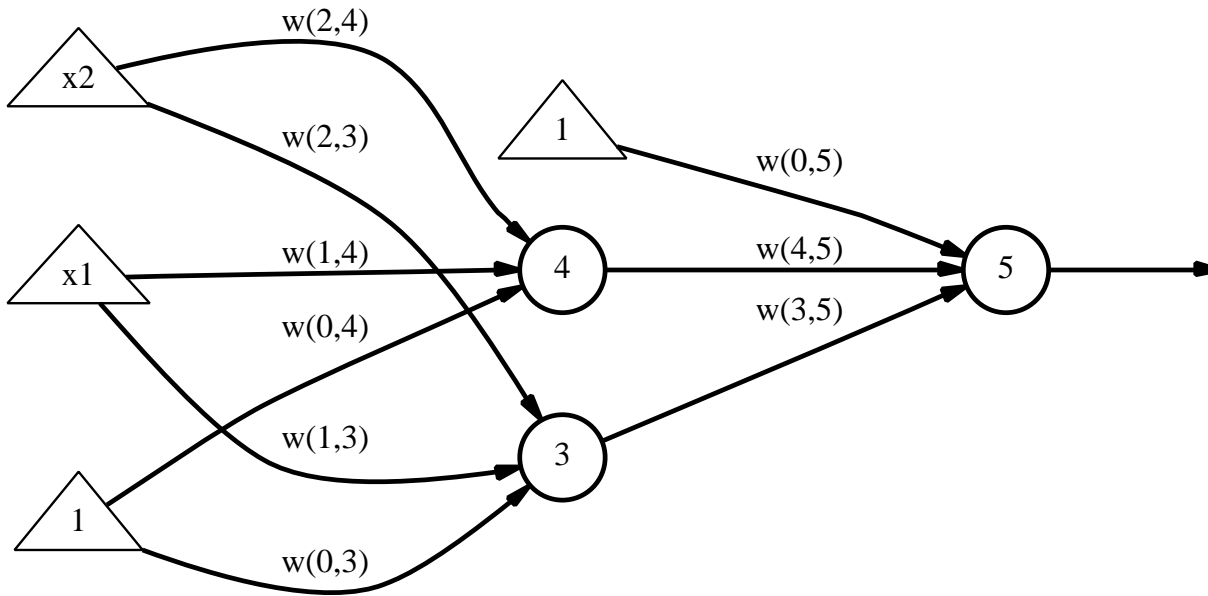
avec f fonction non linéaire : fonction *signe* ou *sigmoïde*.

Les réseaux multicouches.



Le réseau multicouche général.

Les réseaux multicouches.



Un exemple de réseau multicouche.

Exemple

Réseau à deux entrées \mathbf{x}_1 et \mathbf{x}_2 .

Une couche cachée composée des neurones formels numérotés 3 et 4.

Une couche de sortie composée d'un seul neurone formel, numéroté 5.

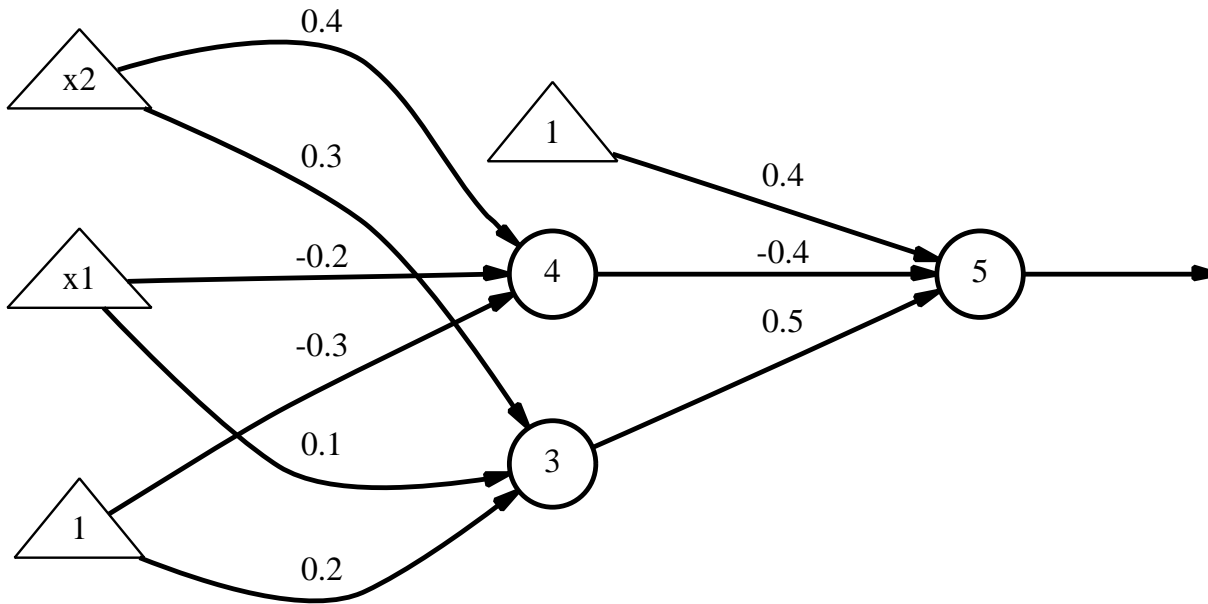
Valeur des poids :

$w(0, 3) = 0.2$	$w(1, 3) = 0.1$	$w(2, 3) = 0.3$
$w(0, 4) = -0.3$	$w(1, 4) = -0.2$	$w(2, 4) = 0.4$
$w(0, 5) = 0.4$	$w(3, 5) = 0.5$	$w(4, 5) = -0.4$

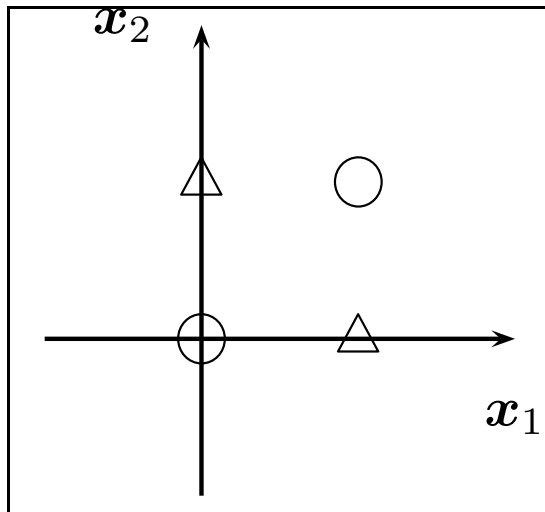
Valeur des entrées :

$$\mathbf{x}_1 = 1 \text{ et } \mathbf{x}_2 = 1$$

Les réseaux multicouches.

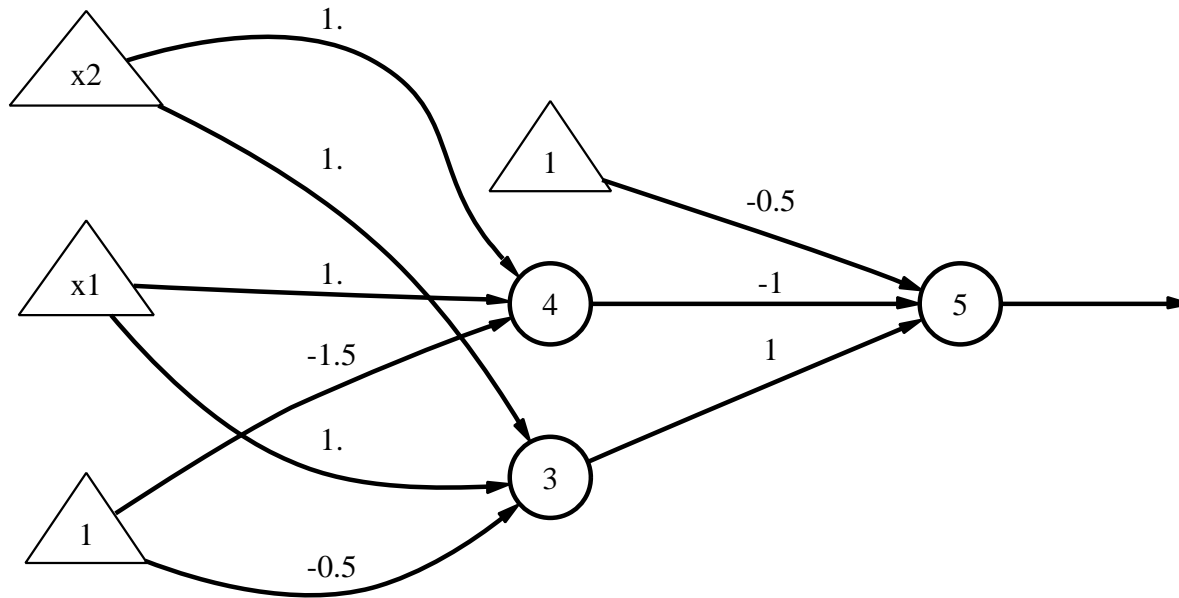


Le problème du OU exclusif.

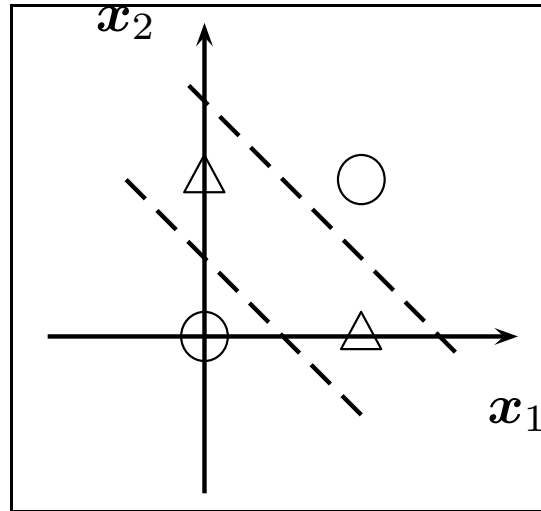


Le problème "XOR" : les deux points \triangle sont des exemples de la même classe, les deux points \bigcirc des exemples d'une autre classe. Les deux classes ne sont pas linéairement séparables.

Un réseau "XOR".



Un réseau "XOR".



Une des façons de résoudre le problème *XOR* avec un réseau connexionniste à une couche cachée. La première droite répond à l'équation $x_1 + x_2 - 0.5 = 0$ et réalise un *OU* logique. La seconde répond à l'équation $x_1 - x_2 - 0.5 = 0$ et réalise un *ET* logique. Elles sont faites par la première couche du réseau. La seconde couche combine les deux décisions linéaires en une décision non linéaire.