

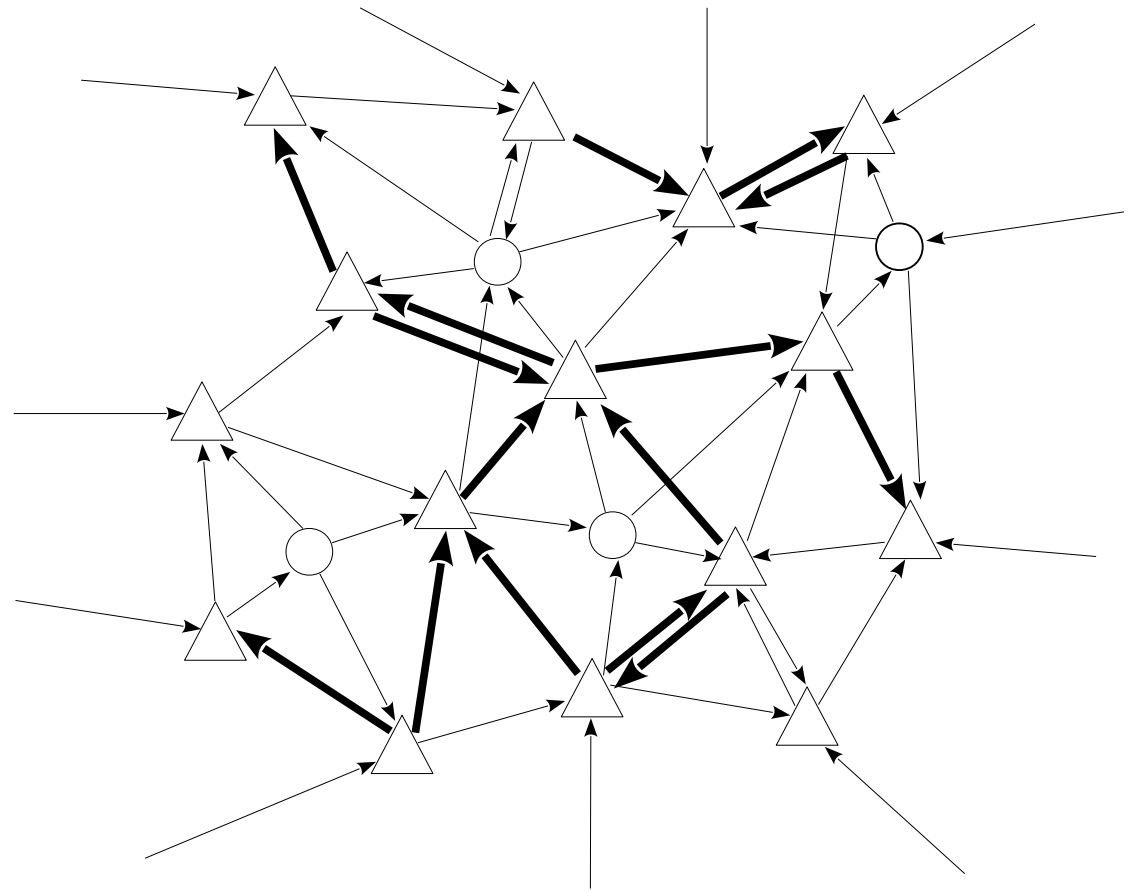
Neurosciences Computationnelles : CM6 Apprentissage Automatique



Michael Graupner
(michael.graupner@parisdescartes.fr)

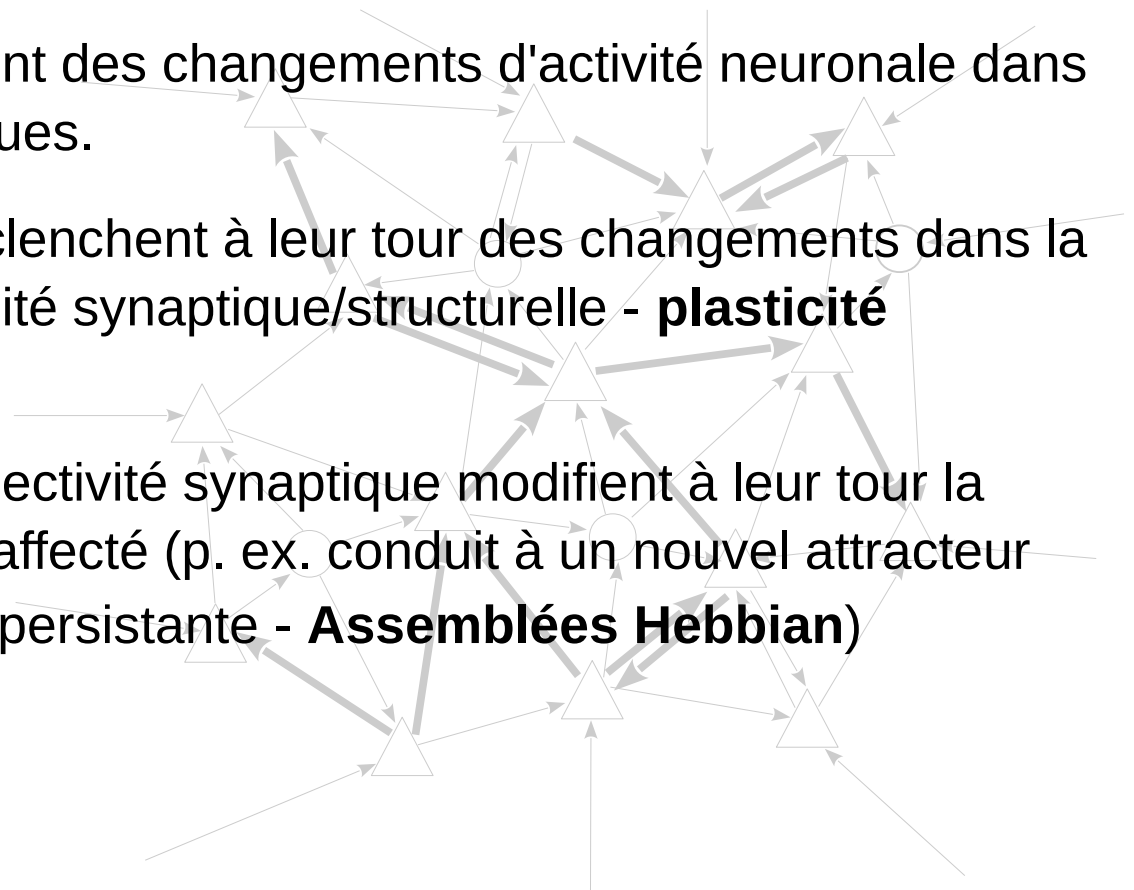
Mécanismes d'apprentissage et de mémoire

- Connectivité synaptique = superposition des traces laissées par les inputs externes



Apprentissage et mémoire - le scénario "Hebbian"

- Des stimuli externes déclenchent des changements d'activité neuronale dans des circuits neuronaux spécifiques.
- Les changements d'activité déclenchent à leur tour des changements dans la connectivité synaptique (plasticité synaptique/structurelle - **plasticité Hebbian**)
- Les changements dans la connectivité synaptique modifient à leur tour la dynamique du circuit neuronal affecté (p. ex. conduit à un nouvel attracteur → activité de période de retard persistante - **Assemblées Hebbian**)



Différents types d'apprentissage

- **Apprentissage non supervisé :**

Les stimuli externes déclenchent des changements d'activité neuronale dans des circuits neuronaux spécifiques; Codage - construction d'une représentation neuronale.

- Apprentissage Hebbian

- **Apprentissage par renforcement :**

Apprendre à connaître les stimuli uniquement sur la base des récompenses et des punitions qui leur sont associées.

- Conditionnement classique

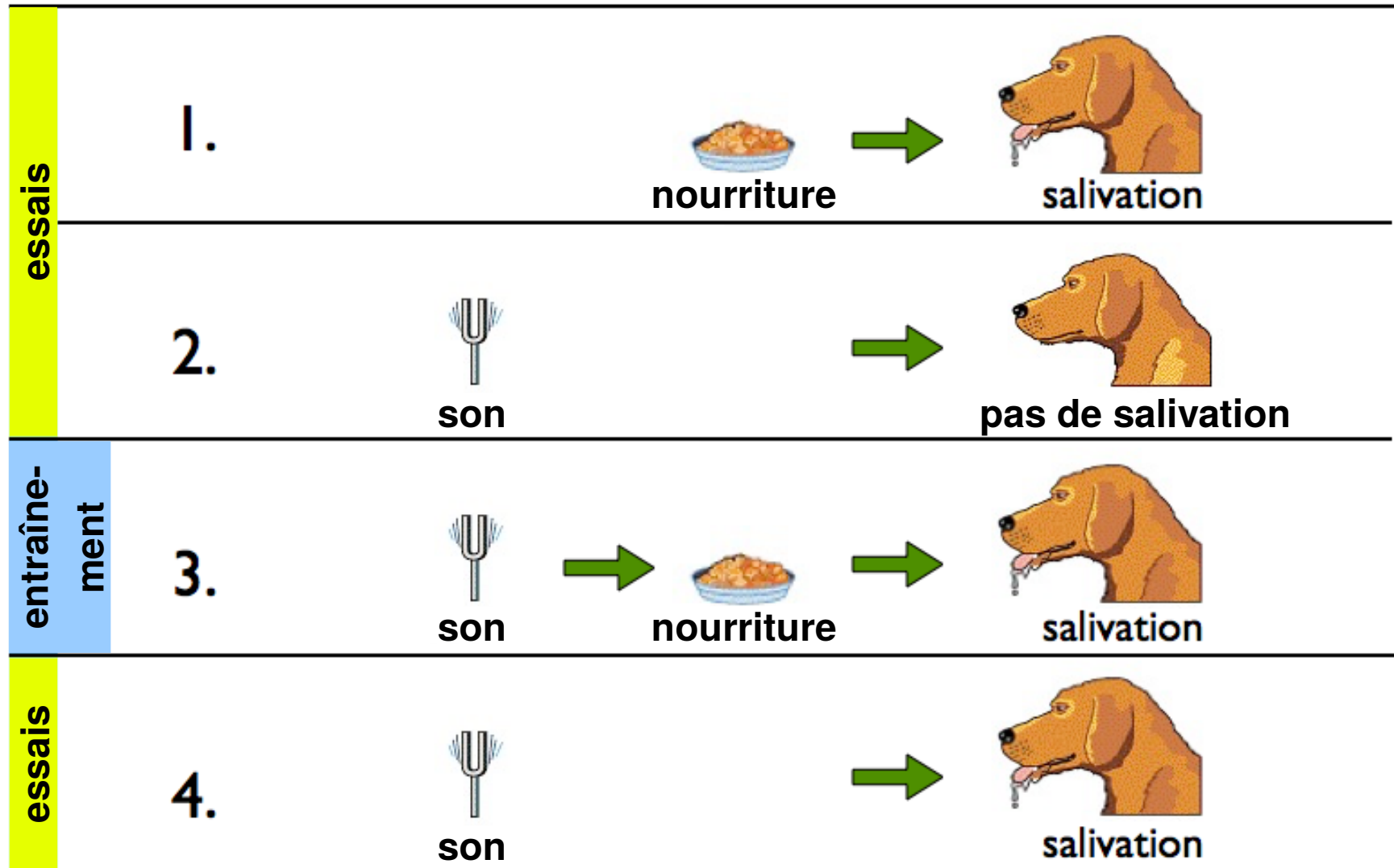
- **Apprentissage supervisé :**

Ensemble de données qui comprend les valeurs cibles (les valeurs que nous voulons prédire). Apprendre une fonction qui prédit correctement les valeurs cibles, qui peut ensuite être utilisée pour faire des prédictions sur d'autres exemples.

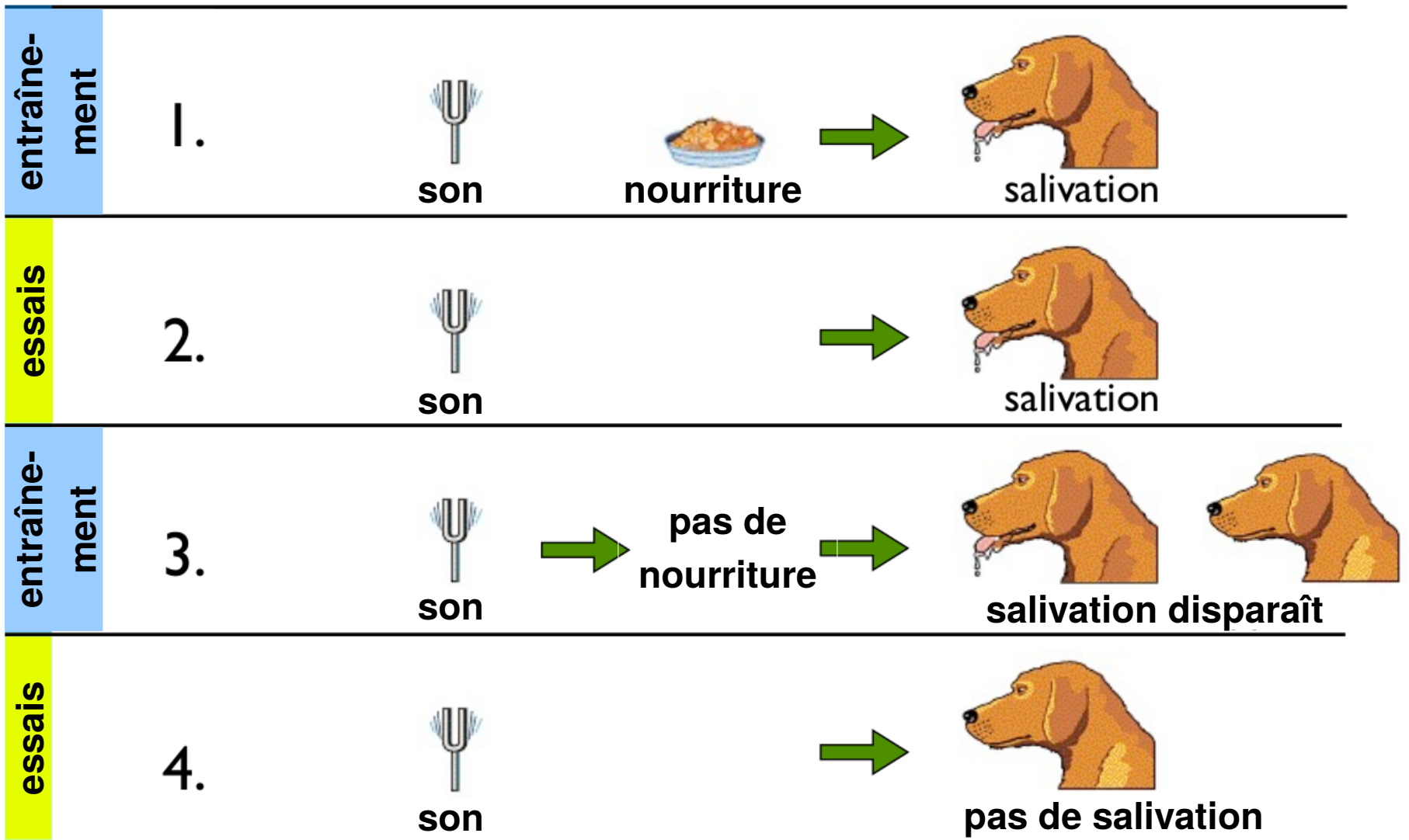
- Exemple : classification, régression, ...

- Le perceptron

Apprentissage par renforcement : conditionnement classique



Apprentissage par renforcement : extinction



Interprétation : conditionnement classique et extinction

Pourquoi est-ce que ça arrive?

- comportement involontaire
- apprentissage des associations
- proximité temporelle des stimuli



Ivan Pavlov
(1849-1936)
travaux: 1890s



John B. Watson
(1878-1958)
travaux: 1920s

Pourquoi est-ce que ça arrive?

- apprendre à prévoir

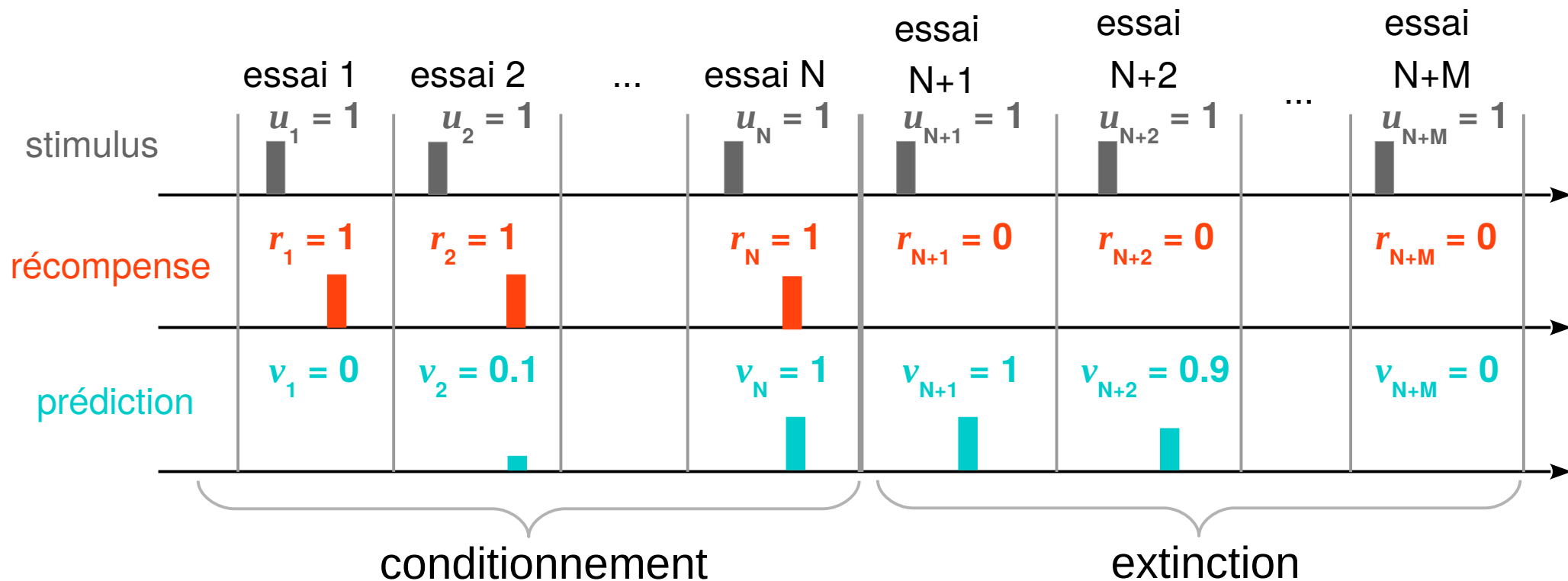


Robert Rescorla



Allan R. Wagner

Modèle Rescorla-Wagner

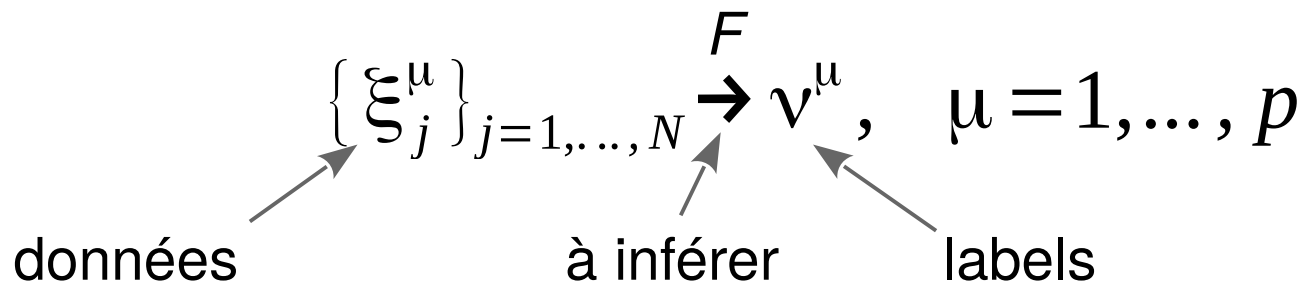


Hypothèse : Le chien veut pouvoir prédire la récompense.
→ veut créer une association entre stimulus et récompense

Modèle Rescorla-Wagner : Peut expliquer le conditionnement et l'extinction.

Apprentissage supervisé

- **But** : inférer une fonction F à partir d'un ensemble
- **Ensemble d'apprentissage** :



- **Intérêt**:
On peut ensuite utiliser cette fonction sur des données ne faisant pas partie de l'ensemble d'apprentissage.
- **Condition** :
Pour inférer la fonction F « correctement », il faut un ensemble d'apprentissage important.

Apprentissage supervisé : exemple

Ensemble d'apprentissage :

$$\left\{ \begin{matrix} \xi^\mu \\ \xi_j \end{matrix} \right\}_{j=1, \dots, N} \rightarrow v^\mu, \quad \mu = 1, \dots, p$$

données

labels ($\mu=2$)



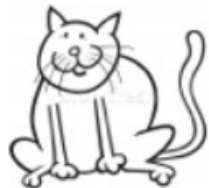
→ chat



→ chien



→ chien



→ chat

...

à inférer

Test :

une fois la fonction apprise



→ ?

Apprentissage supervisé : Le Perceptron

- **Le perceptron (1958) :**

The New York Times

NEW NAVY DEVICE LEARNS BY DOING

July 8, 1958

“The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence... Dr. Frank Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said Perceptrons might be fired to the planets as mechanical space explorers”



Rosenblatt 1958

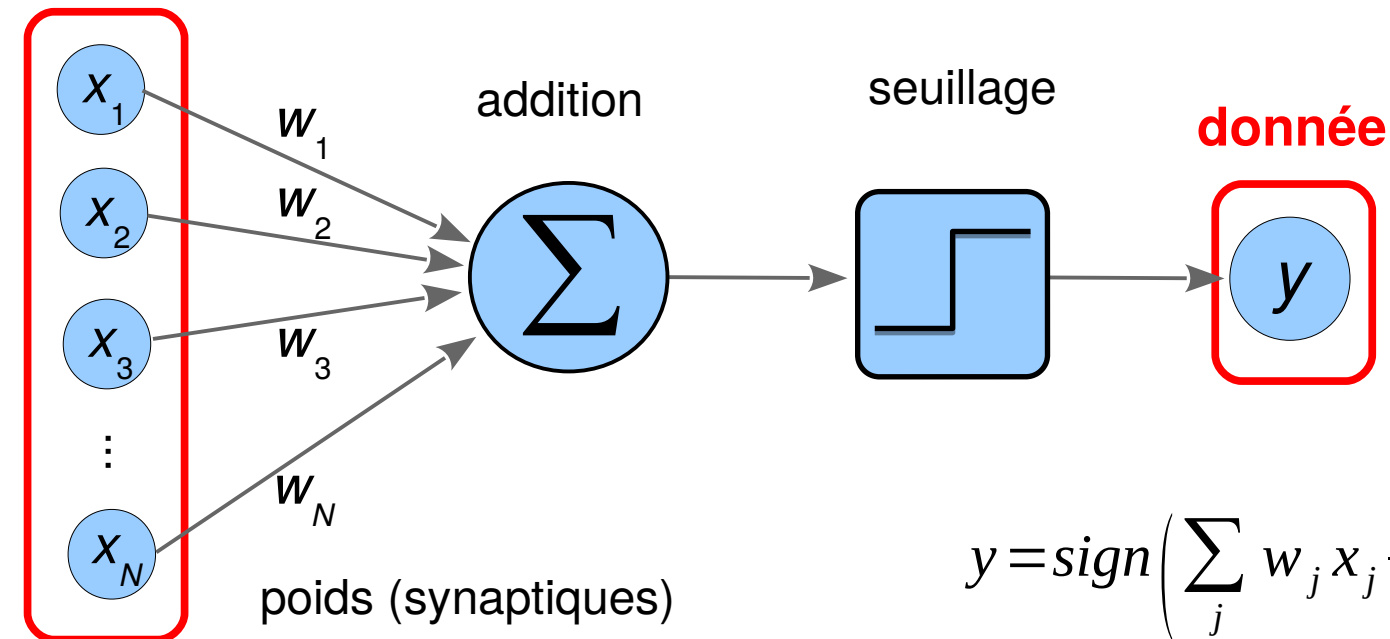
Apprentissage supervisé : Le Perceptron

- **Le perceptron :**

Une machine qui permet de faire des classifications binaires

Couche d'entrée

Couche de sortie



donnée

ajuster/apprendre

$$y = \text{sign} \left(\sum_j w_j x_j - \theta \right)$$

$$\text{sign}(s) = \begin{cases} 1 & \text{si } s \geq 0 \\ -1 & \text{si } s < 0 \end{cases}$$

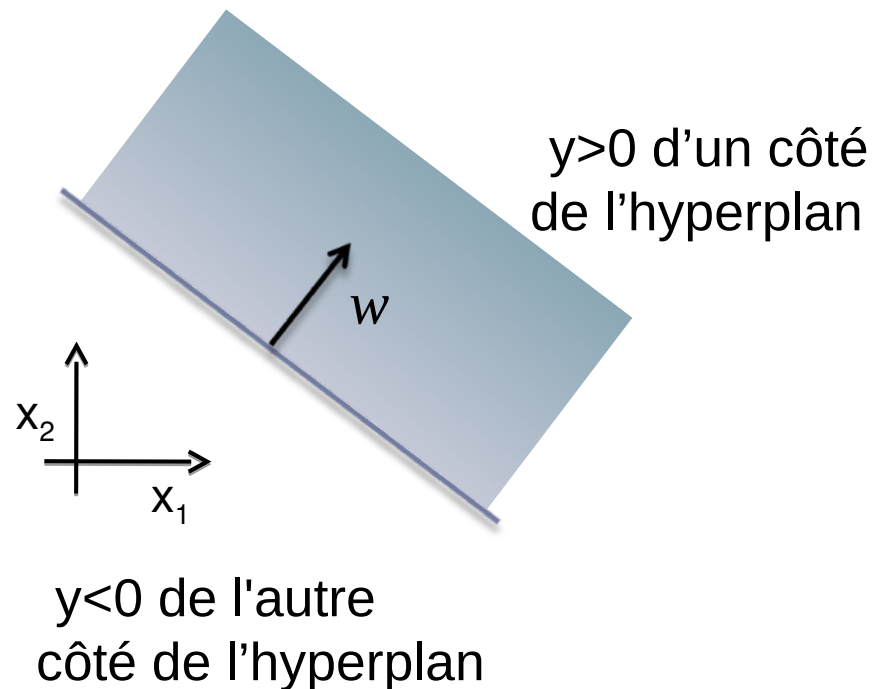
Apprentissage supervisé : Le Perceptron

- **Le perceptron :**

- Une machine qui permet de faire des classifications binaires
- Les perceptrons à une seule couche sont seulement capables d'apprendre des motifs linéaires séparables.
- le réseau neuronal feedforward à deux ou plusieurs couches (également appelé un perceptron multicouche) a une puissance de traitement beaucoup plus grande que les perceptrons à une seule couche
- Outil pour étudier les problèmes de capacité de stockage.

Le Perceptron : Interprétation géométrique

- Équation du perceptron : $y = \text{sign}\left(\sum_j w_j x_j - \theta\right)$ souvent : $y = \text{sign}\left(\sum_{j=0} w_j x_j\right)$
avec $x_0 = 1$
- Équation d'un hyperplan : $\vec{w} \cdot \vec{x} + b = 0$

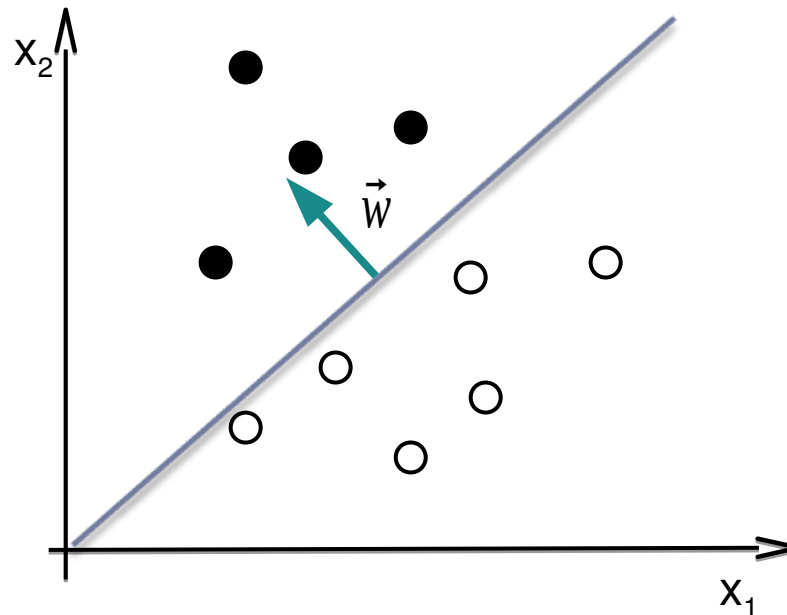


Classification par un perceptron

- Quelles tâches de classification peut effectuer un perceptron?

Tache :

Peut-on trouver des poids (w_1, w_2) tels que $y=1$ pour les points noirs et $y=0$ pour les points blancs?



oui:

$$w_1 = -1$$

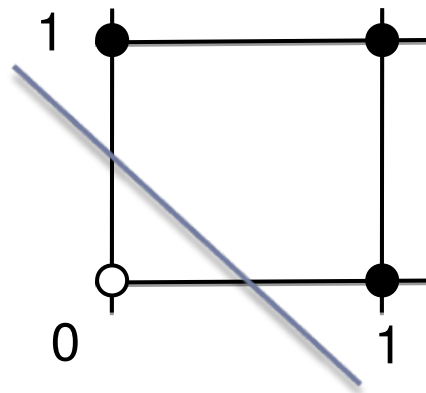
$$w_2 = 1$$

$$b = 0$$

Exemples

- On regarde le cas binaire: $x_i = 0, 1$

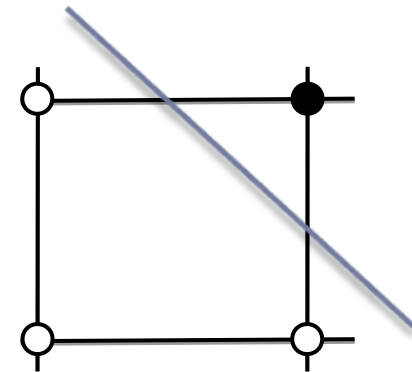
OU(x_1, x_2)



○ = Faux
● = Vrai

$$\begin{aligned}w_1 &= 1 \\w_2 &= 1 \\b &= -1/2\end{aligned}$$

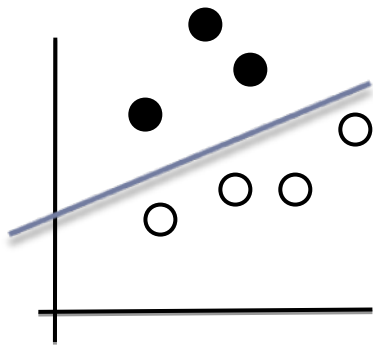
ET(x_1, x_2)



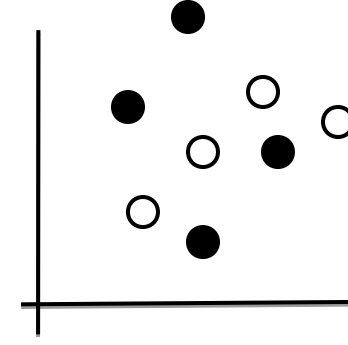
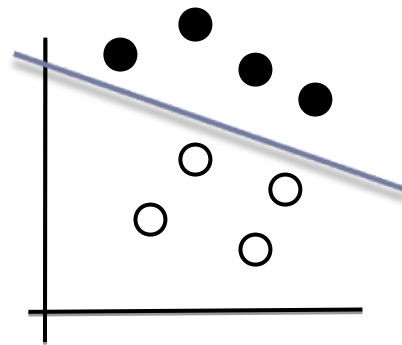
$$\begin{aligned}w_1 &= 1 \\w_2 &= 1 \\b &= -3/2\end{aligned}$$

Séparabilité linéaire

- Deux ensembles de points linéairement séparables = sont de deux côtés d'un hyperplan



linéairement séparable

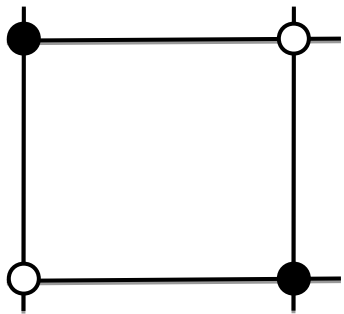


non linéairement
séparable

- Le perceptron peut distinguer des ensembles de points linéairement séparables.

Le ou exclusif

- $XOR(x_1, x_2)$: $XOR(0,1) = XOR(1,0) = 1$
 $XOR(0,0) = XOR(1,1) = 0$



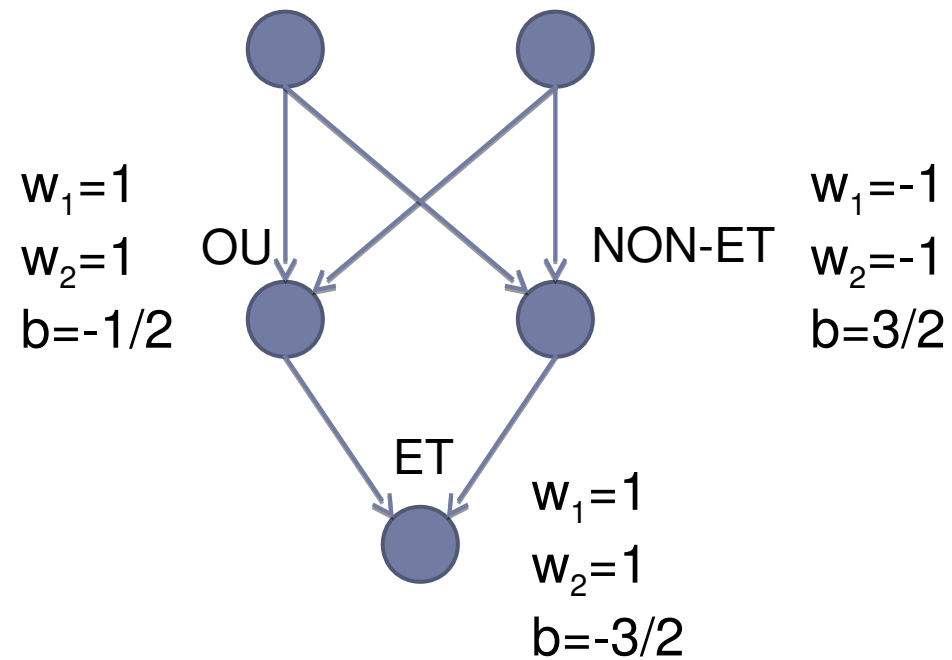
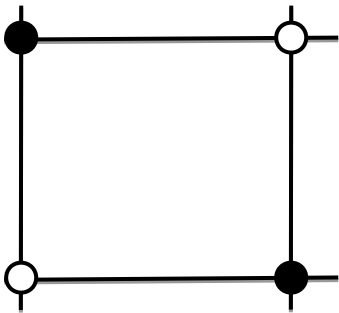
○ = $XOR^{-1}(0)$
● = $XOR^{-1}(1)$

non linéairement séparable

- Un perceptron ne peut pas exécuter un ou exclusif.

Le ou exclusif : possible avec plusieurs couches

- $XOR(x_1, x_2)$: $XOR(0,1) = XOR(1,0) = 1$
 $XOR(0,0) = XOR(1,1) = 0$



Algorithme d'apprentissage Perceptron

Sortie désirée

$$y(n) = \begin{cases} 1 & \text{si } x(n) \in A \\ -1 & \text{si } x(n) \in B \end{cases}$$

1. Initialiser poids et seuils : mettre w_i et b à petits nombre alleatoire dans l'intervalle $[-1, +1]$.
2. Sélectionner un échantillon aléatoire à partir de l'ensemble d'entraînement comme entrée.
3. Si la classification est correcte, ne rien faire
4. Si la classification est incorrecte, modifier le vecteur de poids w en utilisant :

$$w_i = w_i + \eta y(n) x_i(n)$$

Répétez cette procédure jusqu' à ce que l'ensemble du entrainement soit correctement classé.

Perceptron : exemple d'apprentissage

l'ensemble du entrainement :

classe A (étoiles)
sortie désirée +1

classe B (cercles)
sortie désirée -1

$$x(1) = (-2, 1)$$

$$x(4) = (-2, -1)$$

$$x(2) = (1, 1)$$

$$x(5) = (-1, -1.5)$$

$$x(3) = (1.5, -0.5)$$

$$x(6) = (2, -2)$$

valeurs initiales :

$$\eta = 0.2$$

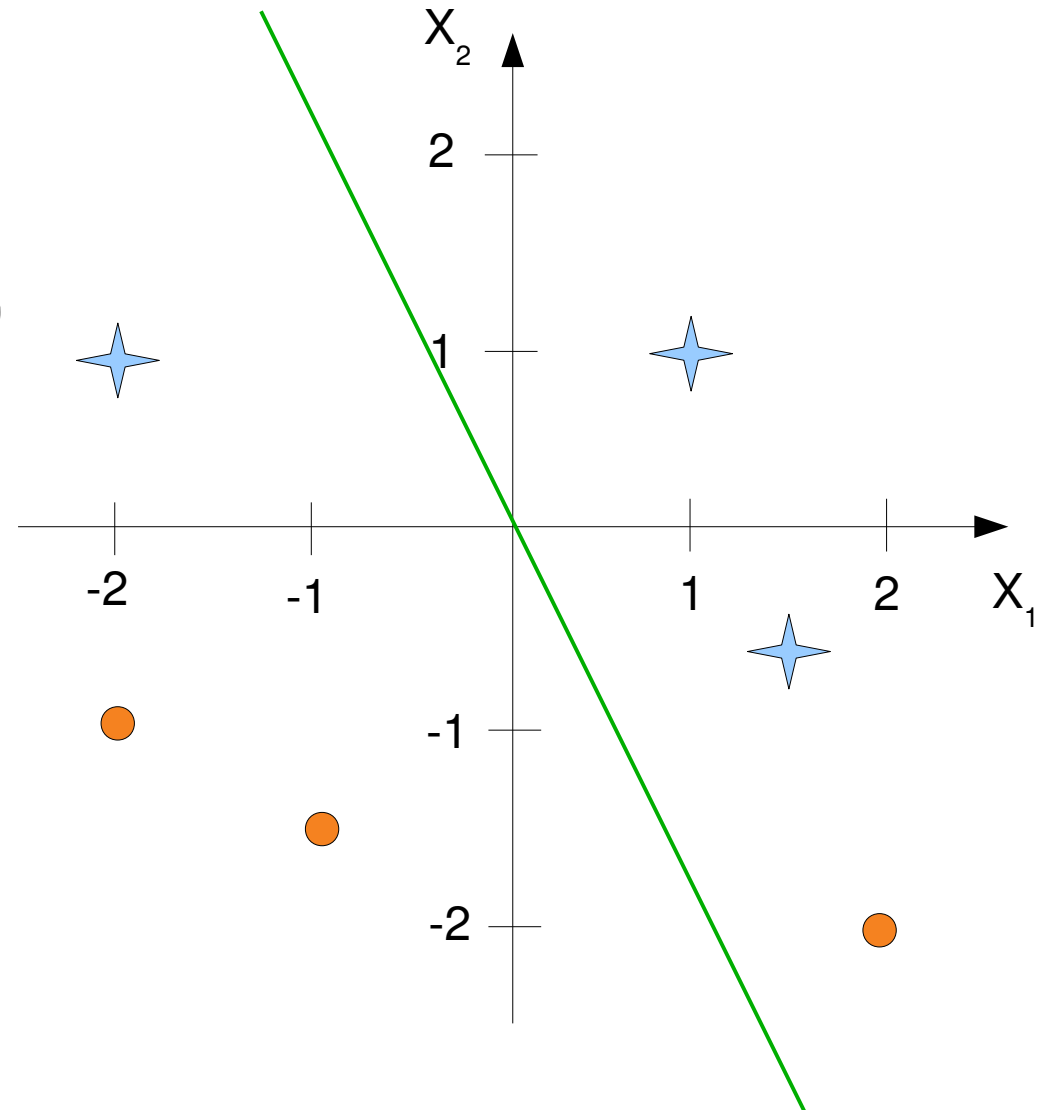
$$w = \begin{pmatrix} 0 \\ 1 \\ 0.5 \end{pmatrix}$$

l'hyperplan :

$$0 = w_0 + w_1 x_1 + w_2 x_2$$

$$0 = 0 + x_1 + 0.5 x_2$$

$$\Rightarrow x_2 = -2 x_1$$



Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} 0 \\ 1 \\ 0.5 \end{pmatrix}$$

$$x_1 = 1, x_2 = 1$$

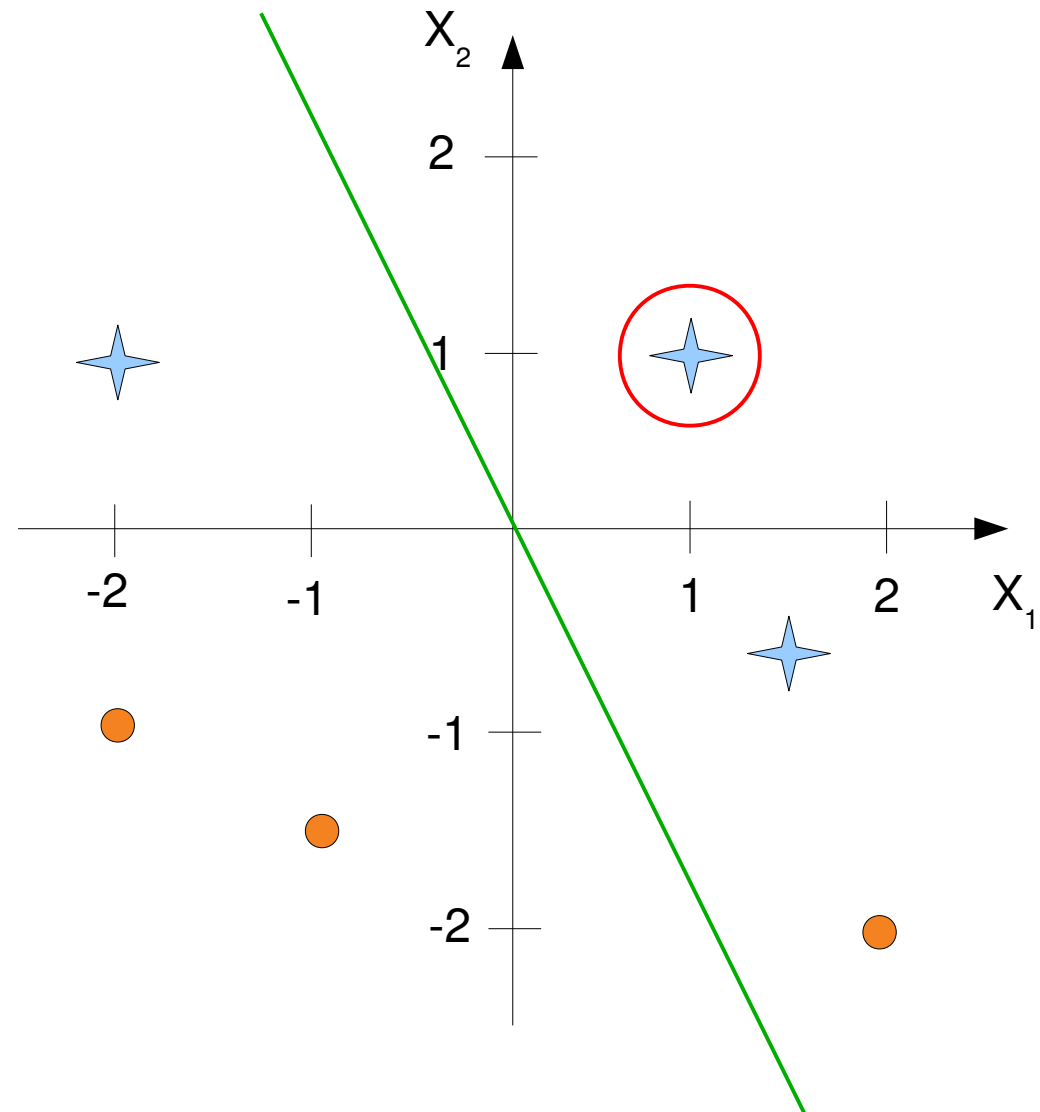
$$w^T x = ???$$

$$y = ???$$

$$w^T x = 1.5$$

$$y = +1$$

Classification correcte,
aucune mesure



Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} 0 \\ 1 \\ 0.5 \end{pmatrix}$$

$$x_1 = 2, x_2 = -2$$

$$w^T x = \mathcal{P}??$$

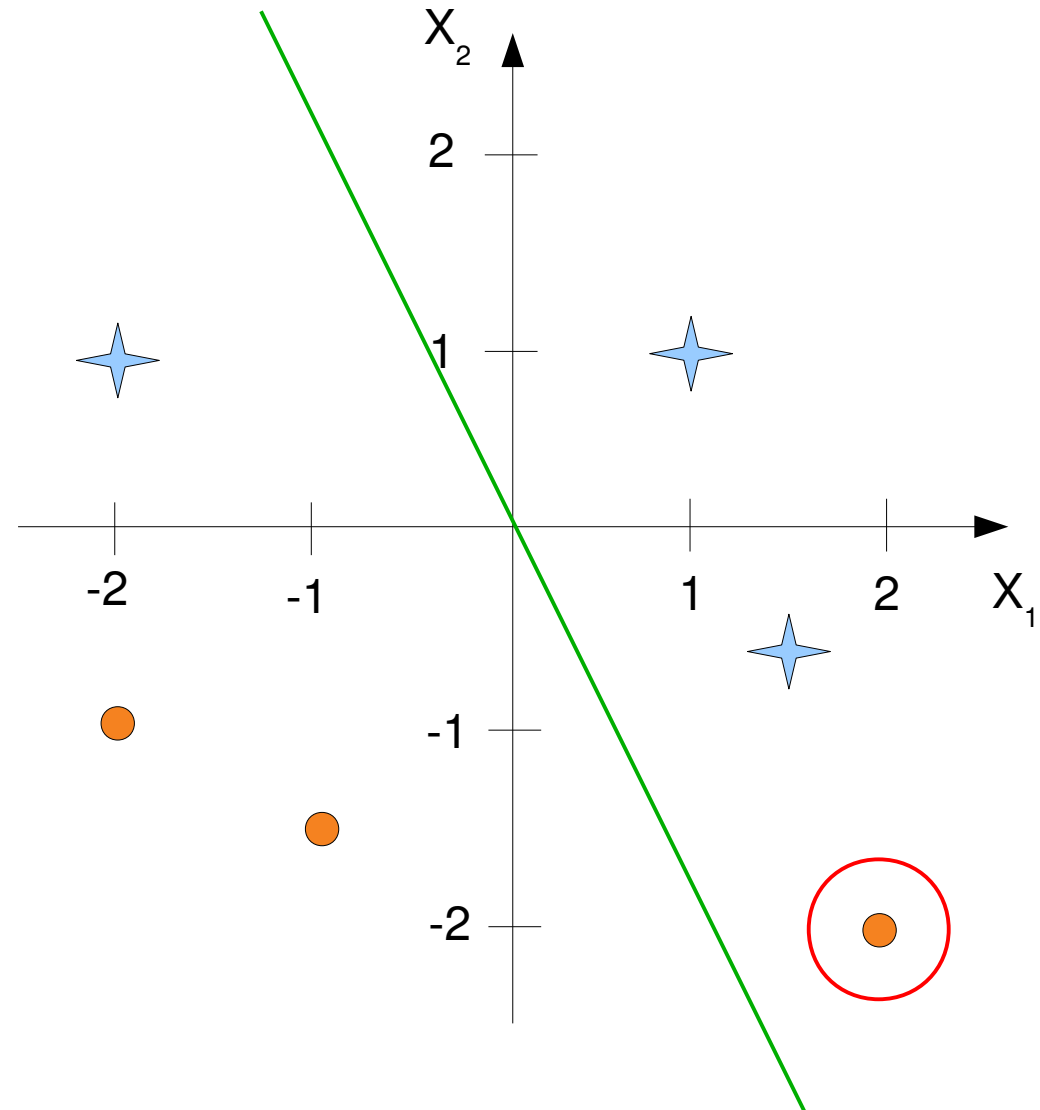
$$y = +1??$$

Classification incorrecte :

$$w_0 = w_0 - 0.2 * 1$$

$$w_1 = w_1 - 0.2 * 2$$

$$w_2 = w_2 - 0.2 * (-2)$$

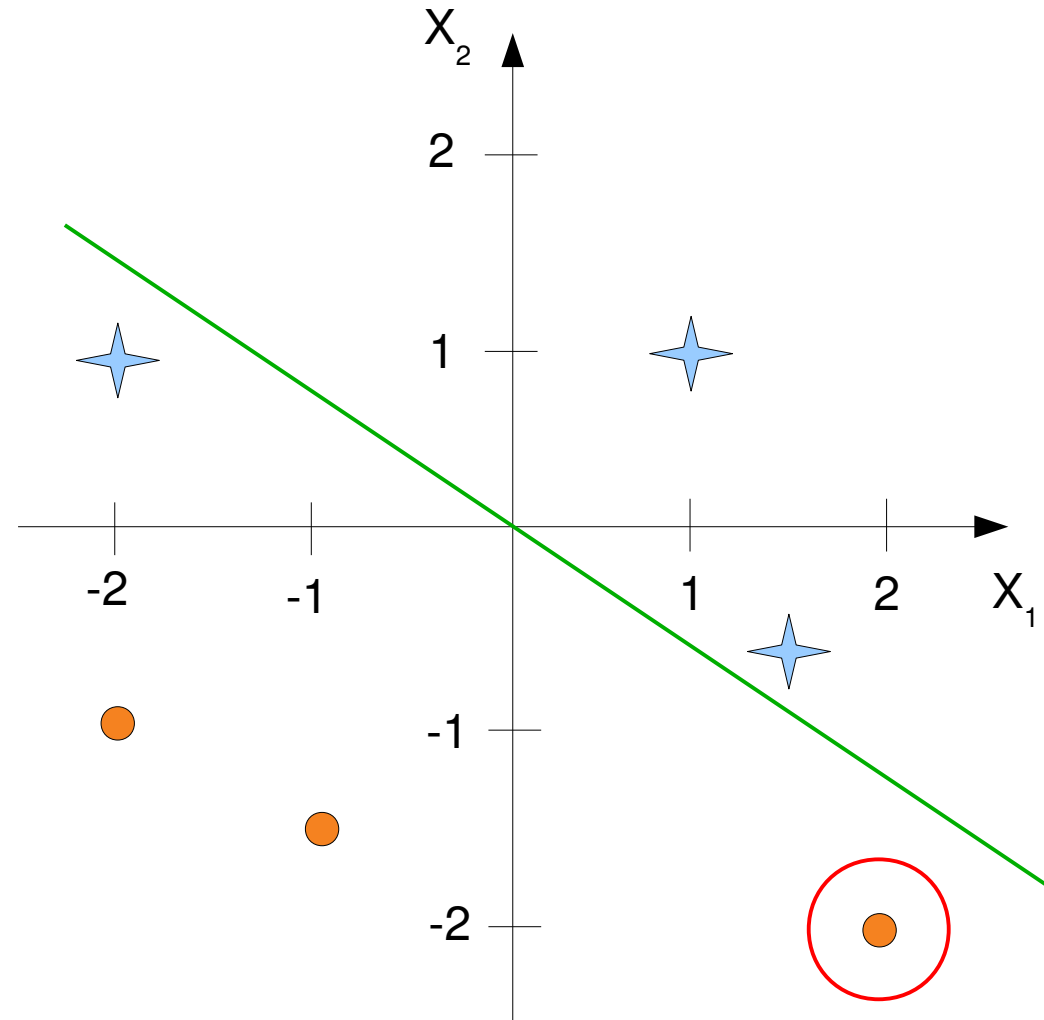


Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} -0.2 \\ 0.6 \\ 0.9 \end{pmatrix}$$

$$x_1 = 2, x_2 = -2$$



Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} -0.2 \\ 0.6 \\ 0.9 \end{pmatrix}$$

$$x_1 = -1, x_2 = -1.5$$

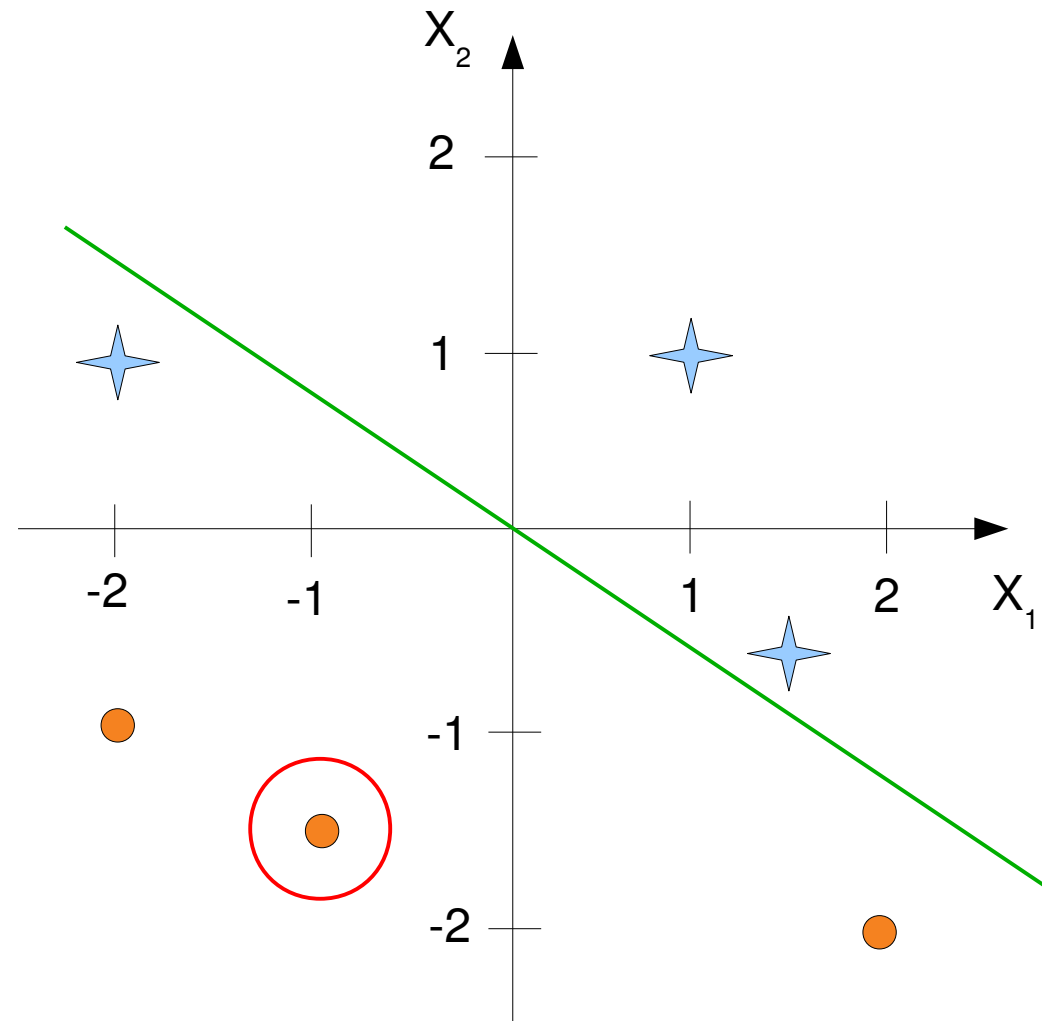
$$w^T x = ???$$

$$y = ???$$

$$w^T x = -2.15$$

$$y = -1$$

Classification correcte,
aucune mesure



Perceptron : exemple d'apprentissage

$$\eta=0.2$$

$$w = \begin{pmatrix} -0.2 \\ 0.6 \\ 0.9 \end{pmatrix}$$

$$x_1 = -2, x_2 = -1$$

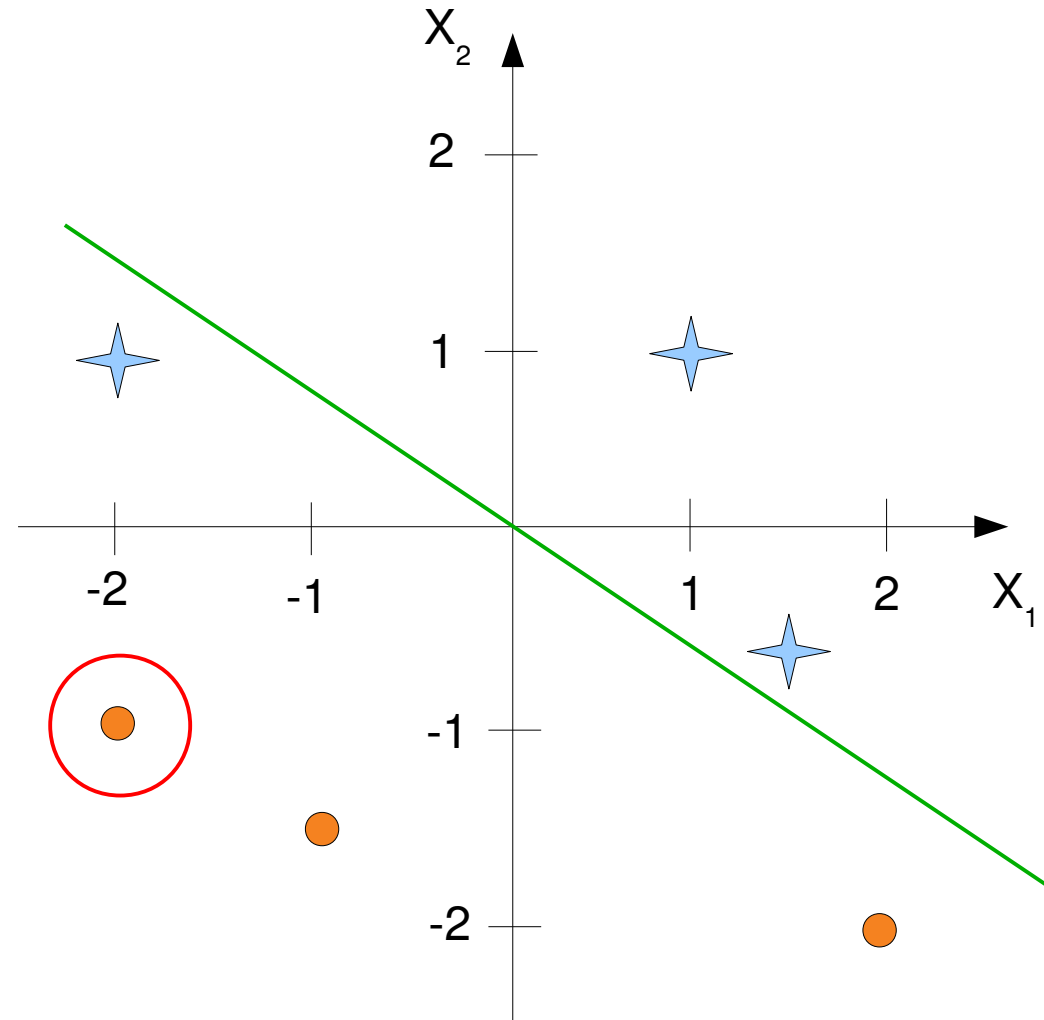
$$w^T x = ???$$

$$y = ???$$

$$w^T x = -2.3$$

$$y = -1$$

Classification correcte,
aucune mesure



Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} -0.2 \\ 0.6 \\ 0.9 \end{pmatrix}$$

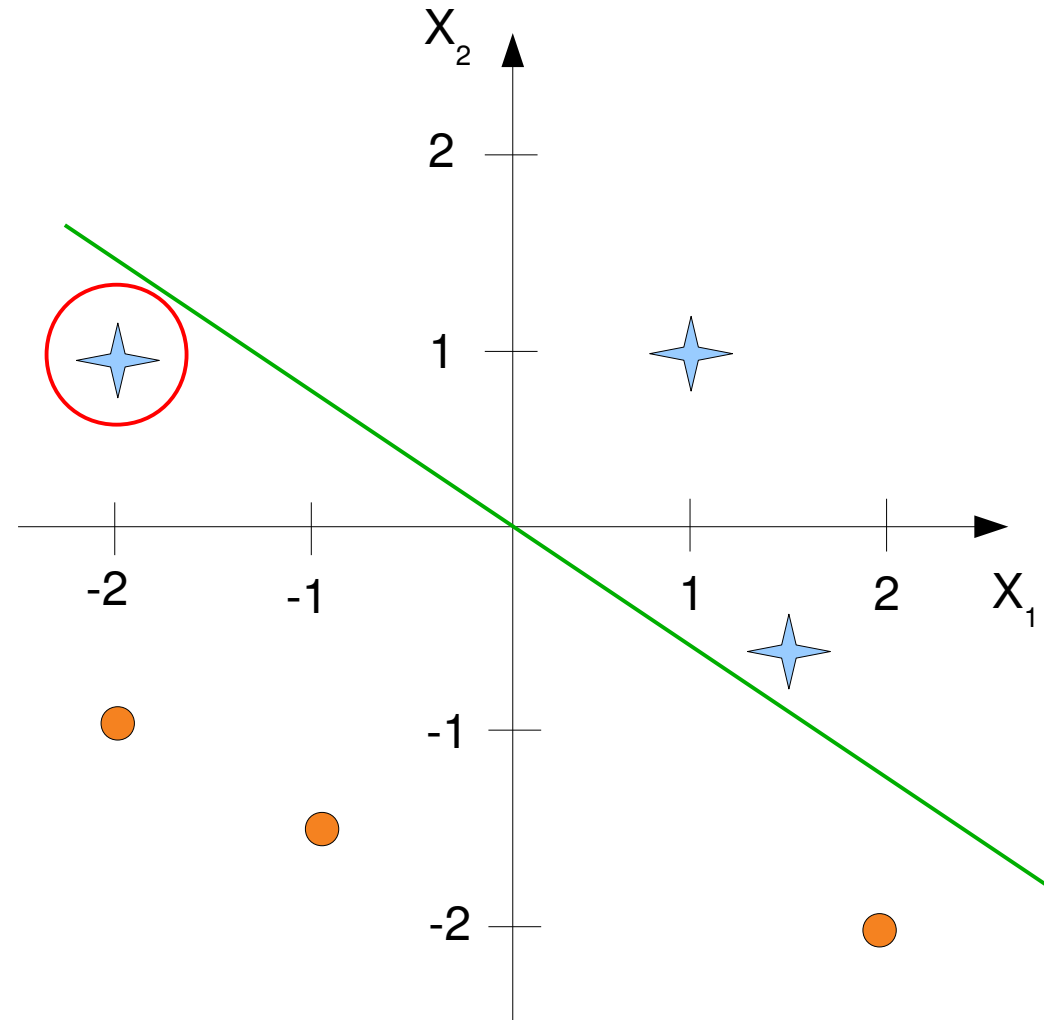
$$x_1 = -2, x_2 = 1$$

Classification incorrecte :

$$w_0 = w_0 + 0.2 * 1$$

$$w_1 = w_1 + 0.2 * (-2)$$

$$w_2 = w_2 + 0.2 * 1$$



Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} 0 \\ 0.2 \\ 1.1 \end{pmatrix}$$

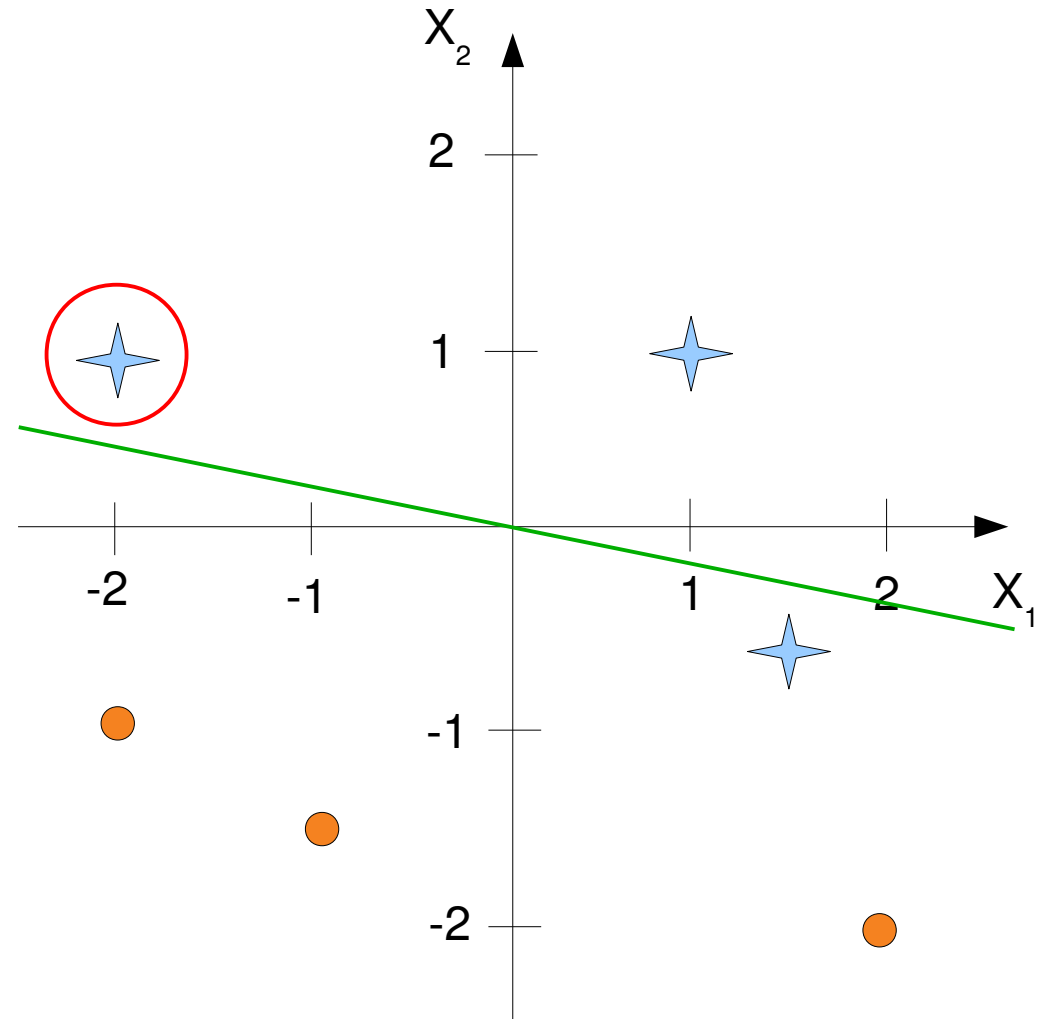
$$x_1 = -2, x_2 = 1$$

Classification incorrecte :

$$w_0 = w_0 + 0.2 * 1$$

$$w_1 = w_1 + 0.2 * (-2)$$

$$w_2 = w_2 + 0.2 * 1$$



Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} 0 \\ 0.2 \\ 1.1 \end{pmatrix}$$

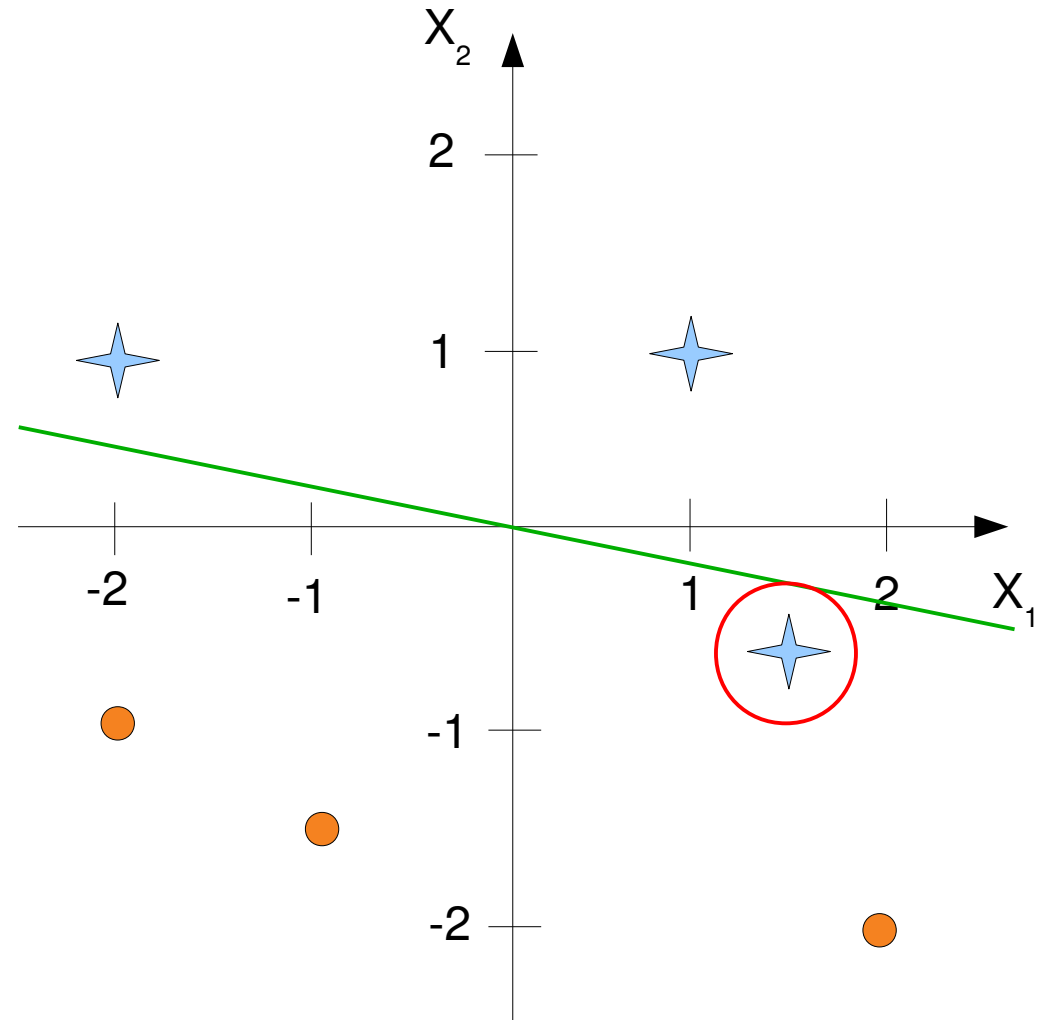
$$x_1 = 1.5, x_2 = -0.5$$

Classification incorrecte :

$$w_0 = w_0 + 0.2 * 1$$

$$w_1 = w_1 + 0.2 * 1.5$$

$$w_2 = w_2 + 0.2 * (-0.5)$$



Perceptron : exemple d'apprentissage

$$\eta = 0.2$$

$$w = \begin{pmatrix} 0.2 \\ 0.5 \\ 1 \end{pmatrix}$$

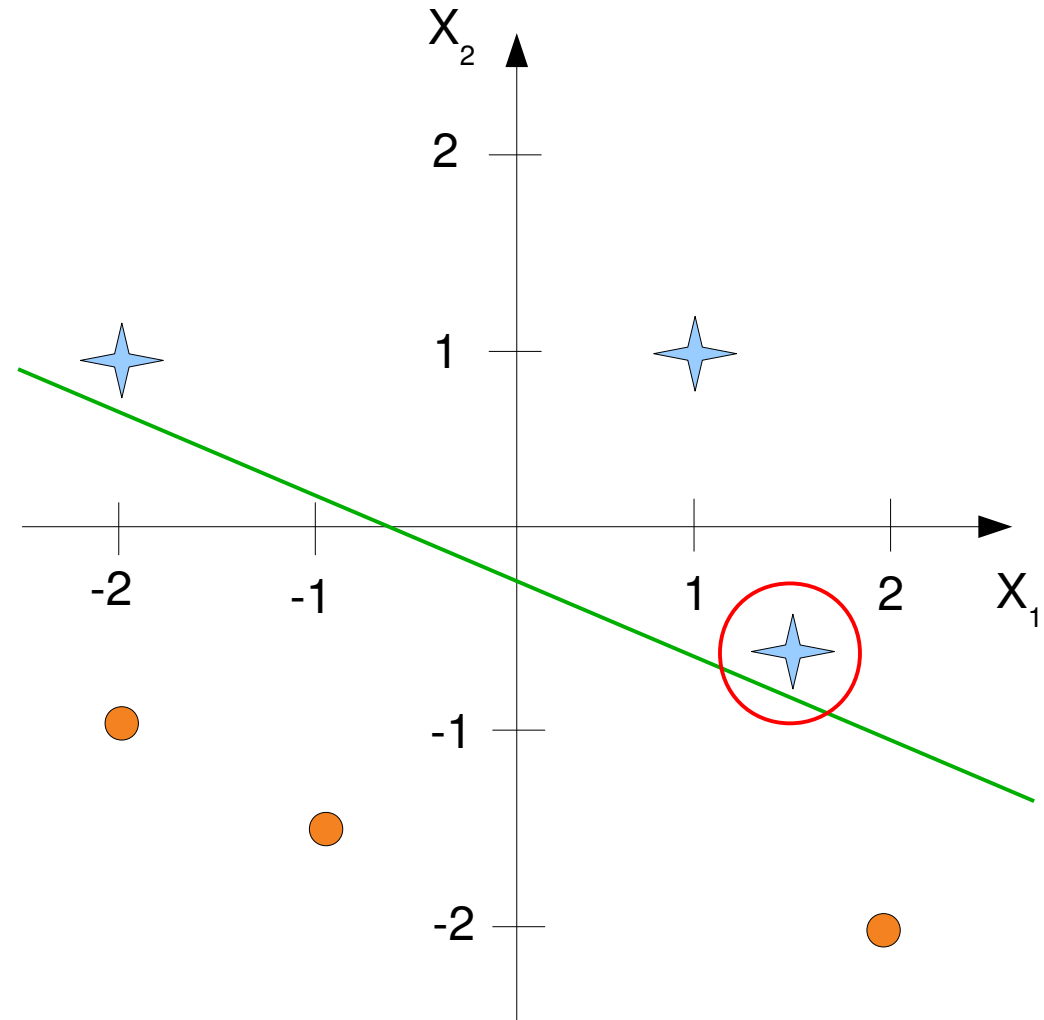
$$x_1 = 1.5, x_2 = -0.5$$

Classification incorrecte :

$$w_0 = w_0 + 0.2 * 1$$

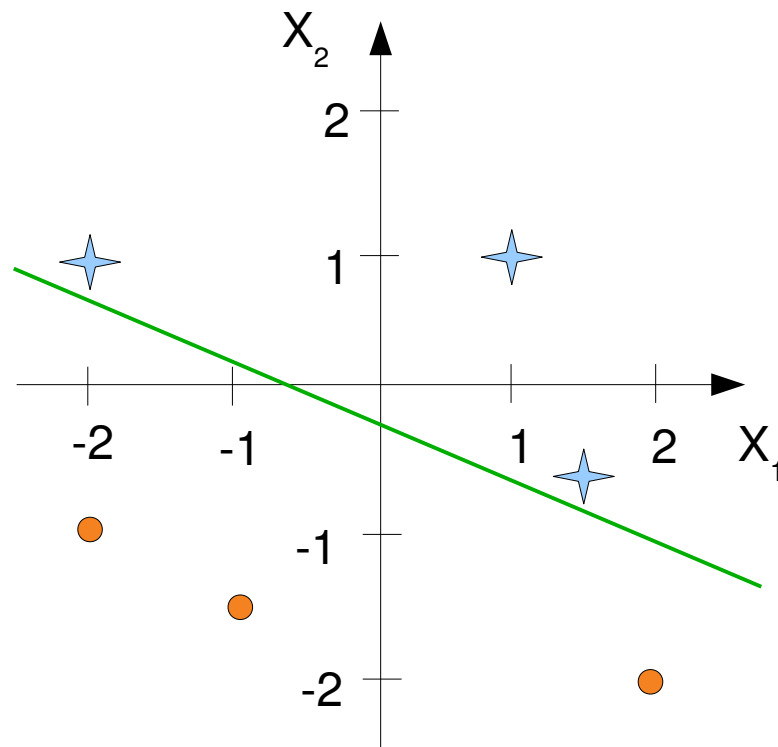
$$w_1 = w_1 + 0.2 * 1.5$$

$$w_2 = w_2 + 0.2 * (-0.5)$$



Théorème de convergence Perceptron

Le théorème stipule que pour tout ensemble de données linéairement séparable, la règle d'apprentissage du perceptron est garantie de trouver une solution dans un nombre fini d'itérations.



Perceptron : exemple

Un Perceptron qui apprend à reconnaître les chiffres écrits à la main.



Projet de fin de cours : les entrées

La base de données MNIST : <http://yann.lecun.com/exdb/mnist/>

60,000 chiffres pour l'entrainement

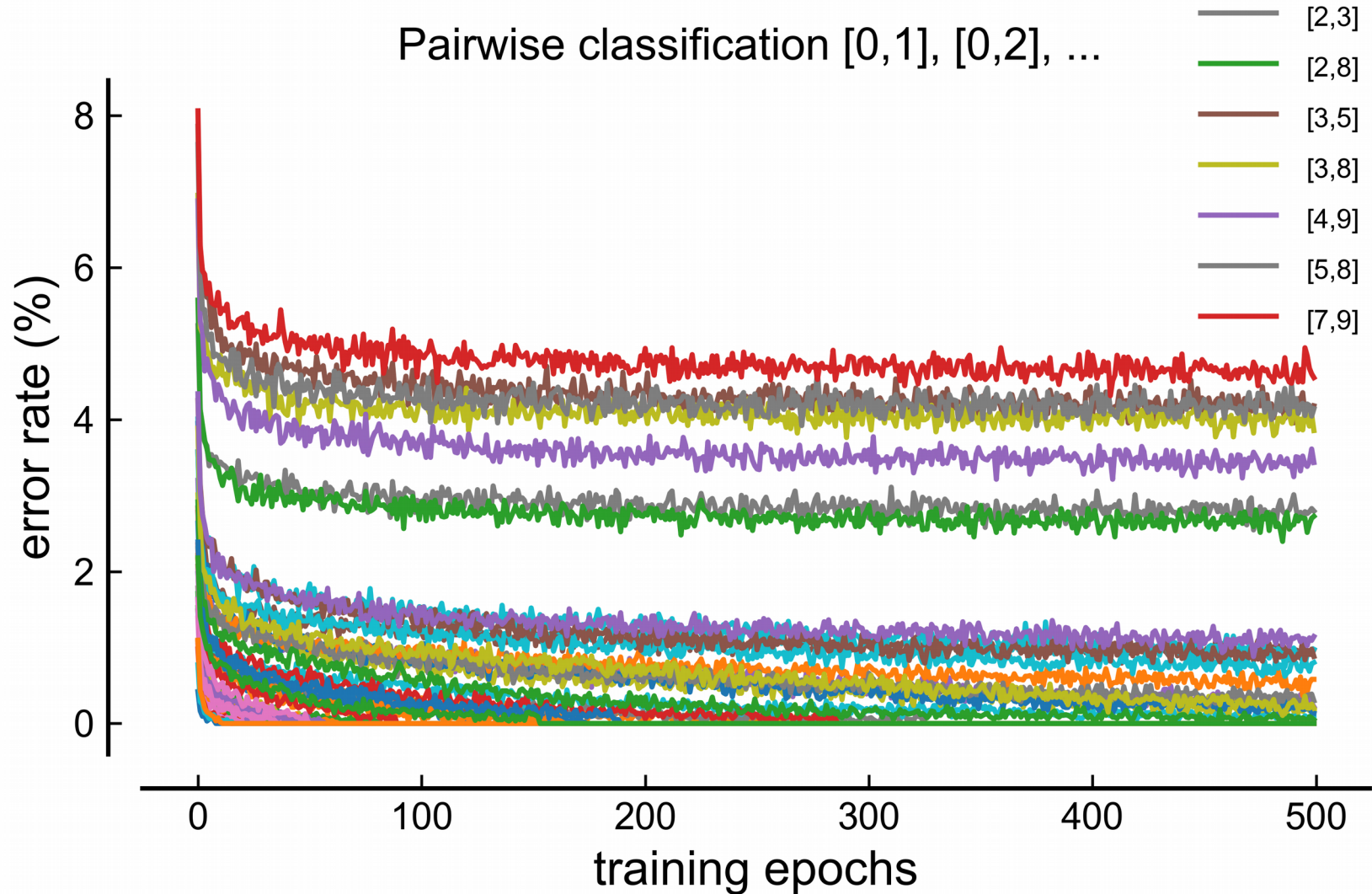
10,000 chiffres pour le test



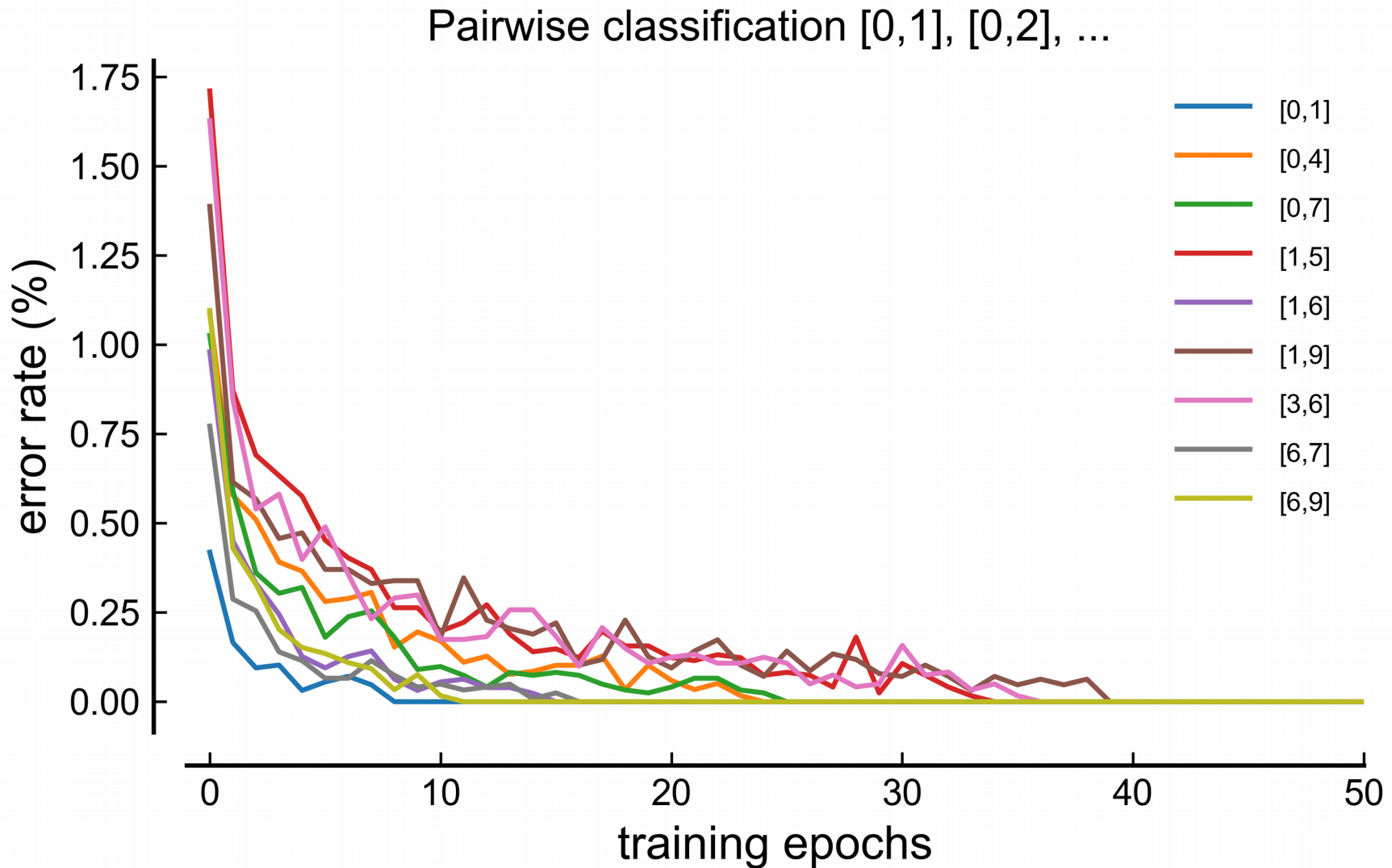
Projet de fin de cours : conseils

- Le but est d'entraîner le Perceptron avec les 60,000 chiffres et ensuite de tester la reconnaissance des 10,000 chiffres.
- Pour ce faire, entraînez le perceptron à différencier d'abord deux chiffres.
- Utilisez un Perceptron d'une seule couche où l'image est traitée comme entrée et l'unité de sortie doit décider si l'un ou l'autre chiffre a été présenté.
- Suivez la performance de la classification pendant l'entraînement et déterminez la performance avec les chiffres de test.
- Important : L'entraînement nécessitera plusieurs boucles de présentation des 60,000 chiffres.
- Réfléchir à la façon de généraliser le perceptron pour reconnaître tous les nombres et pas seulement deux.

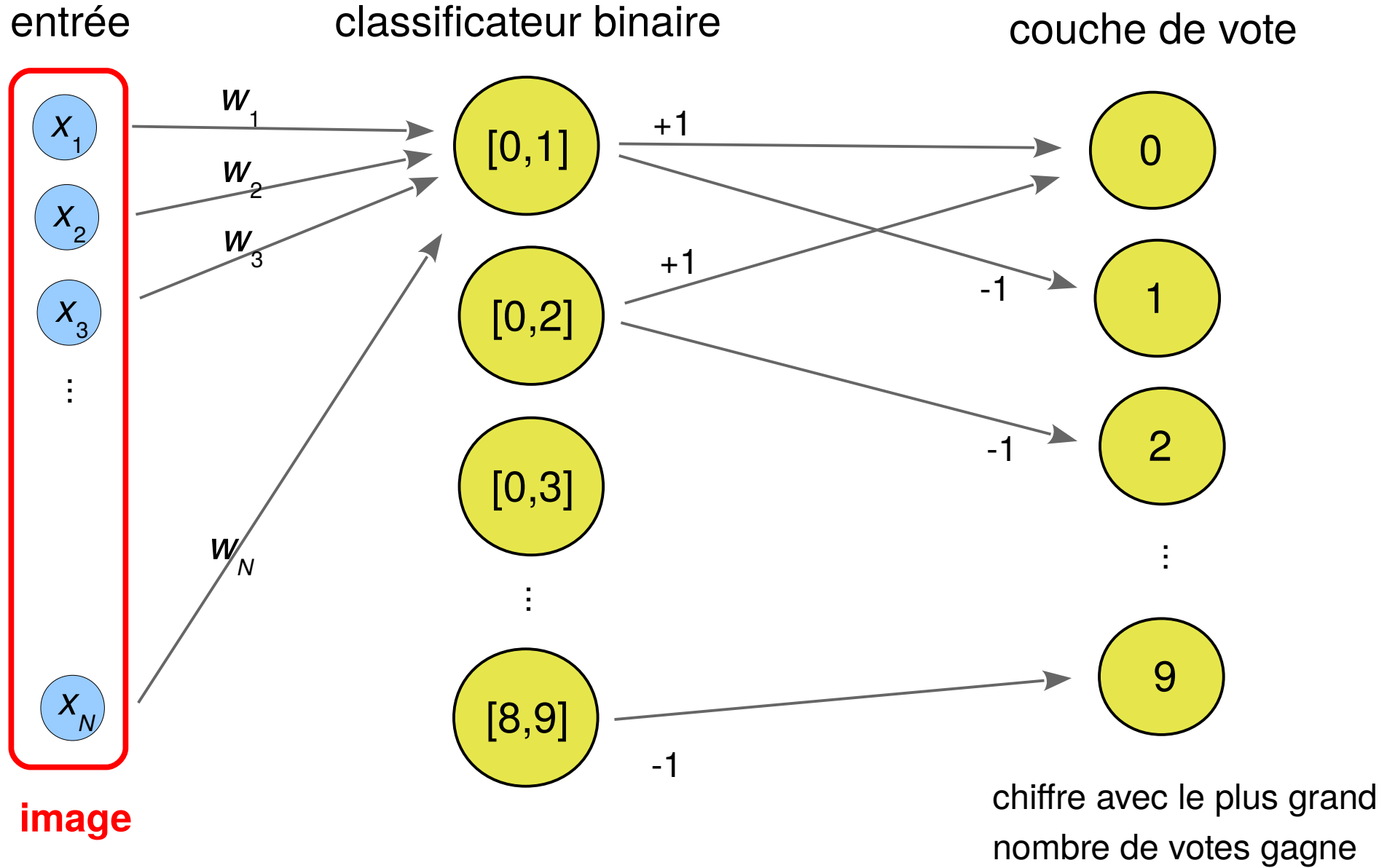
Perceptron : performance de classification binaire



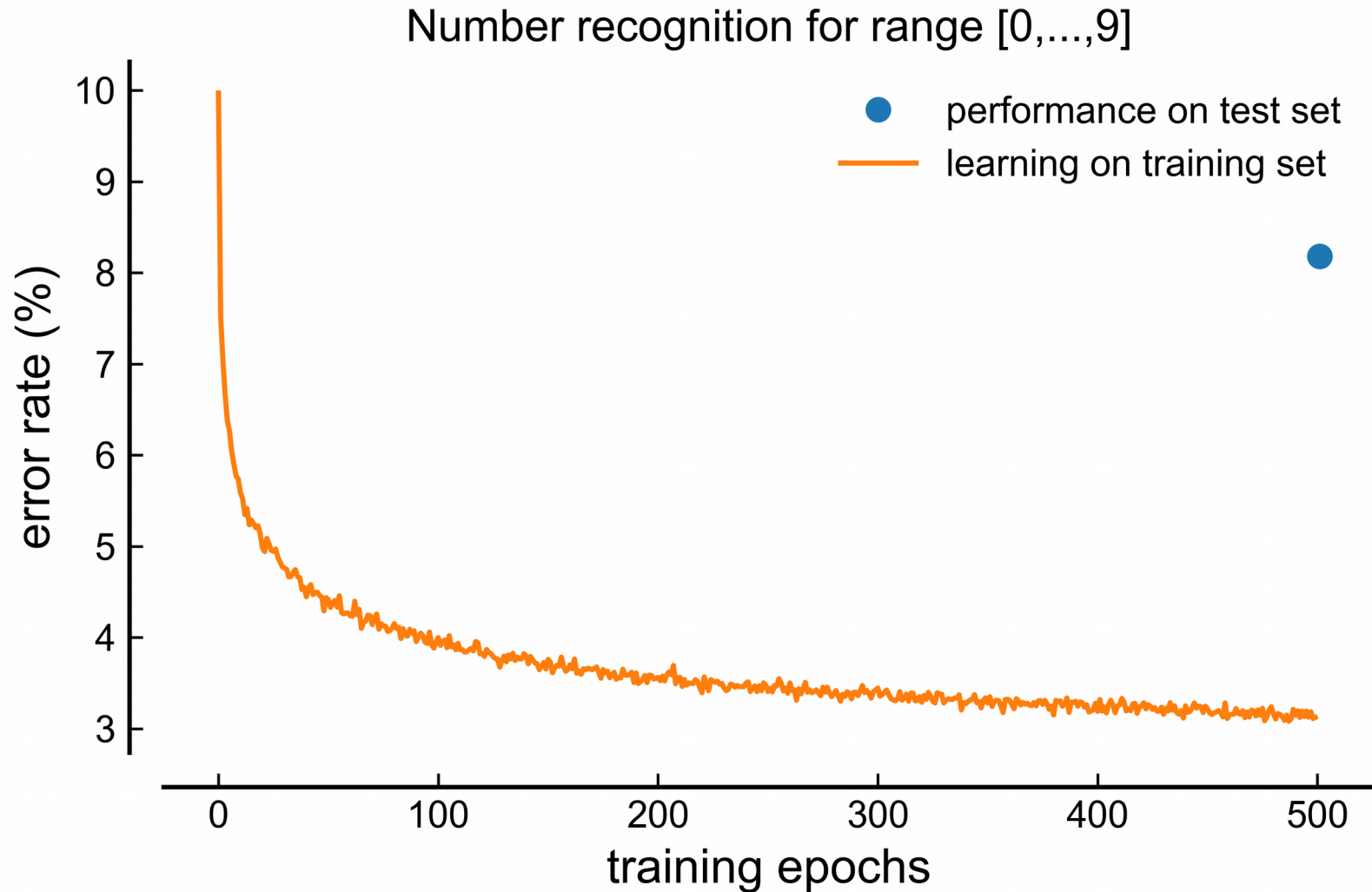
Perceptron : performance de classification binaire



Perceptron : reconnaître 10 chiffres



Perceptron : reconnaître 10 chiffres



Perceptron : reconnaître 10 chiffres

