

# Fondamentaux de L'Apprentissage Automatique

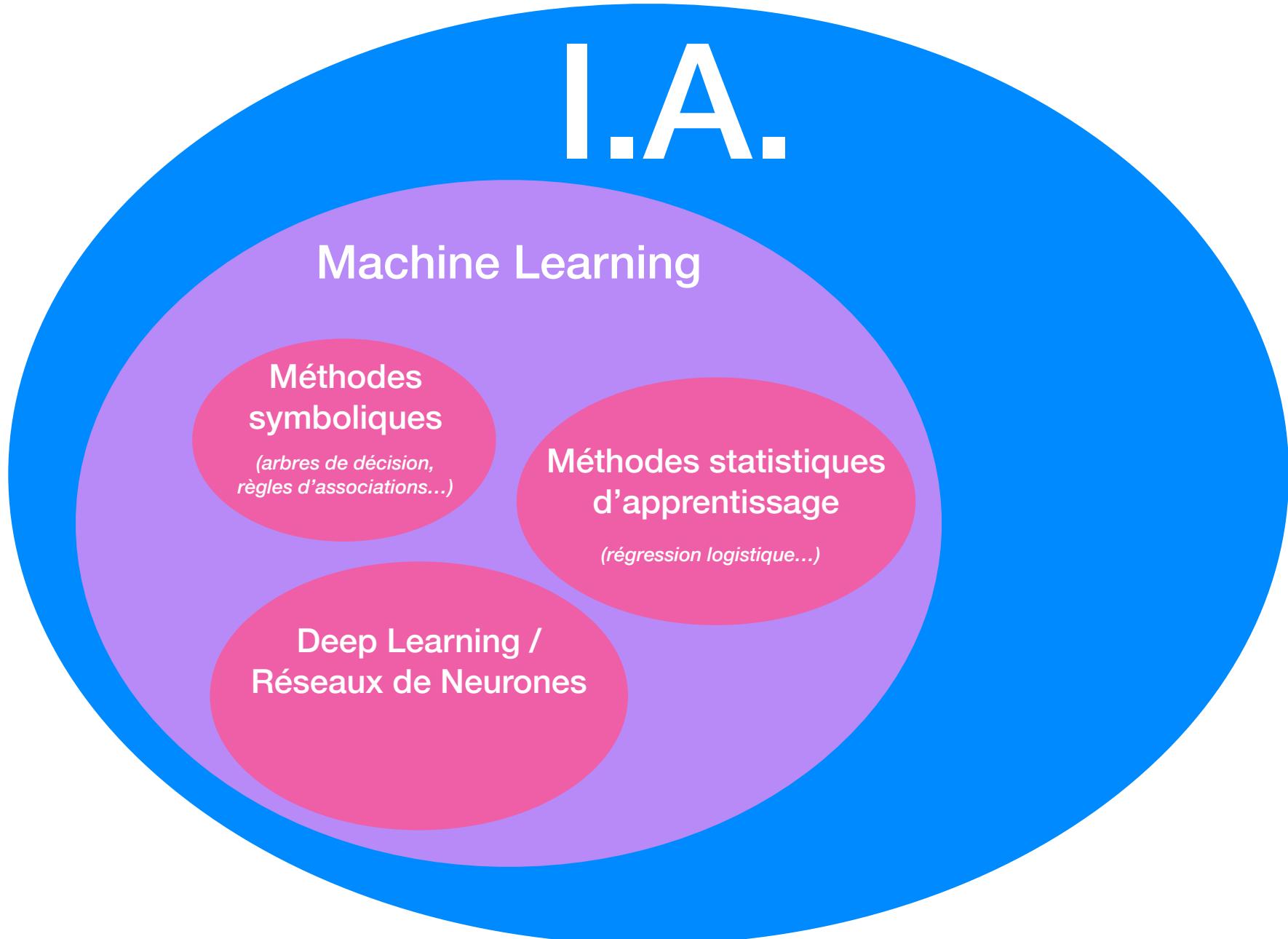
Yann Chevaleyre

# Plan CM1 - Introduction to ML and Statistical Learning Theory

- General Introduction
  - What is ML ?
  - Various Machine Learning problems (supervised, unsupervised,...)
- Supervised Learning: A Formal Setting, definition of the risk
- Warmup: Bounding the risk in a simple case
- Bad News about the risk (No Free Lunches)
- Chosing the right classifier space:  
decomposition of the risk
- Studying the risk of k-Nearest Neighbors

# Introduction générale

# Les Méthodes de ML en IA



# Quelques repères historiques sur les méthodes d'Apprentissage Supervisé

- XIX siècle - Régression linéaire [Legendre, Gauss]
- 1936 - Linear Discriminant Analysis [Fisher]
- 1943 - Modèle mathématique du neurone. Pas d'apprentissage [McCulloch & Pitts]
- 1949 - Algorithme d'apprentissage non supervisé pour neurone [Hebb]
- 1958 - Algorithme du Perceptron [Rosenblatt]
- 1951 - Algorithme de la descente de gradient stochastique [Robbins, Monro]
- 1963 - Arbres de décision [Morgan, Sonquist]  
Classificateurs à vaste marge [Vapnik]
- 1971 - Réseaux de neurones à 8 couches [Ivakhnenko & Lapa]
- 1974 - Algorithme de rétro-propagation du gradient [Werbos]
- 1979 - Convolutional neural networks (*Neocognitron*, pas de back-prop. [Fukushima])
- 1988 - Naïve Bayes [Ohmann]
- 1989 - Réseau de neurone LeNet [Y. Lecun]
- 1990 - Méthodes de Boosting
- 1992 - Méthodes à Noyaux [V. Vapnik]
- 2012 - AlexNet [A. Krizhevsky]
- 2015 - ResNets

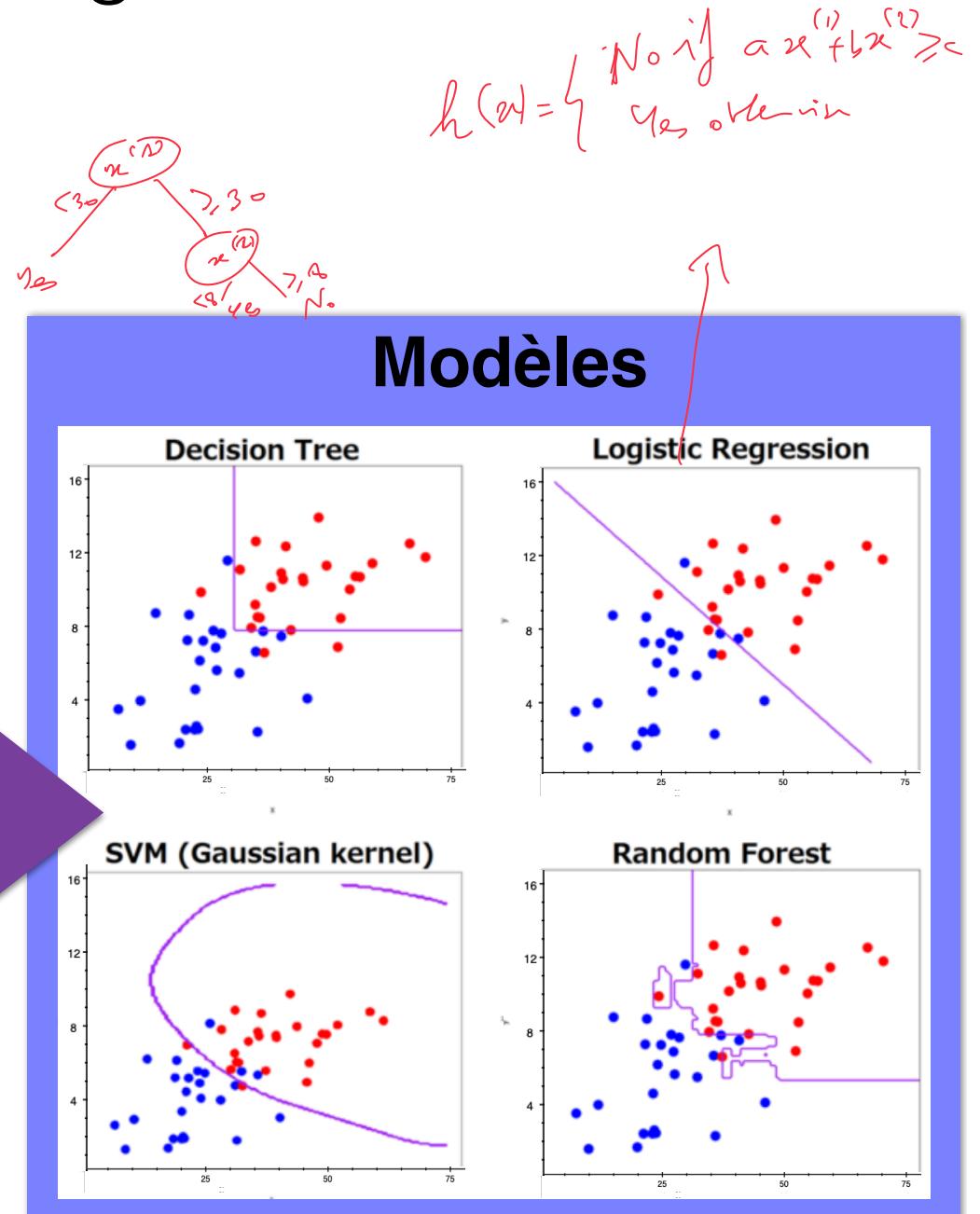
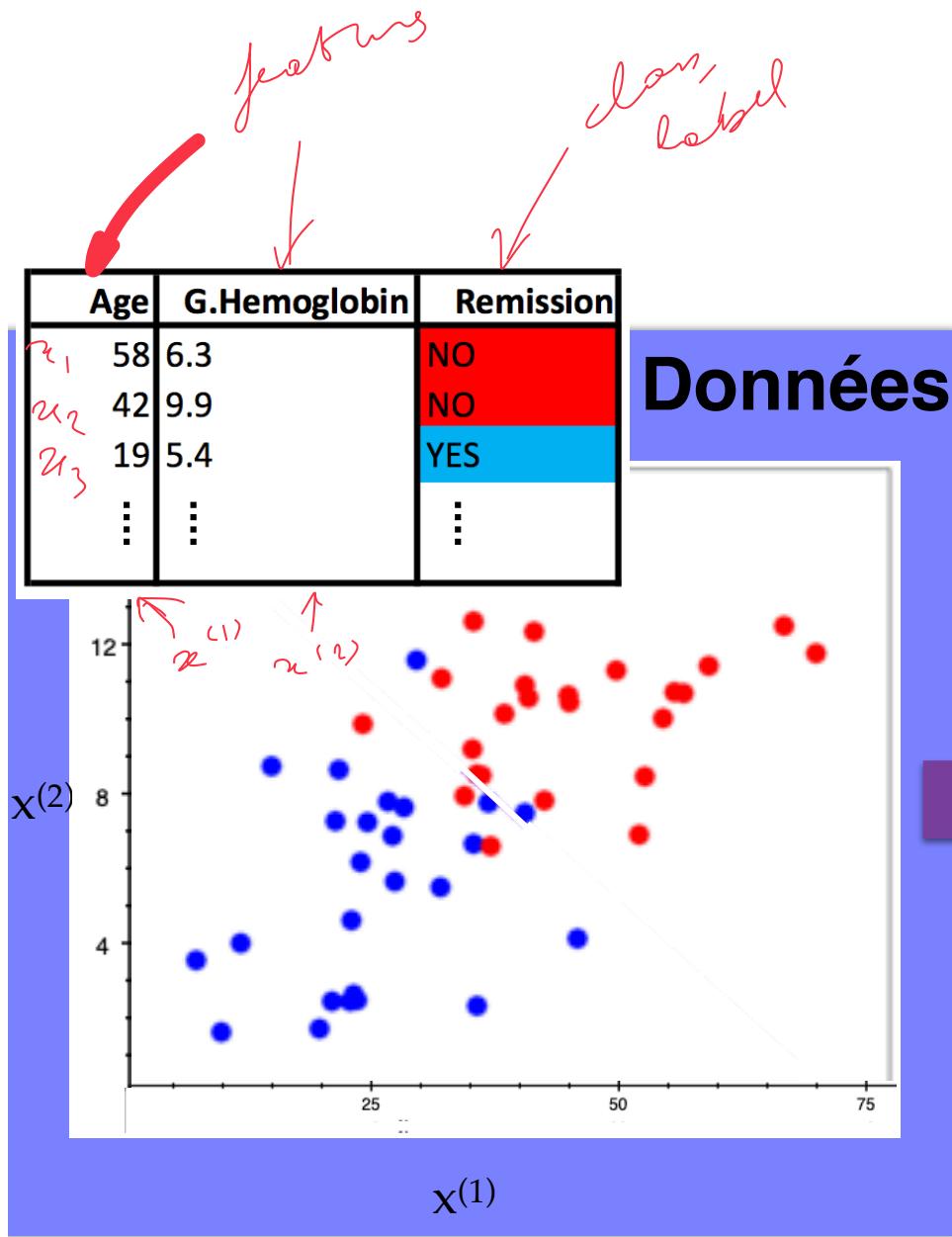
# Typologie des problèmes d'apprentissage

- **Apprentissage Supervisé**
  - Régression, classification, analyse de série temporelle  
filtrage collaboratif (recommandation)
- **Apprentissage non-supervisé**
  - Clustering, réduction de dimension, estimation de densité
- **Apprentissage par renforcement**

# Supervised ML: Learning a Model from Data

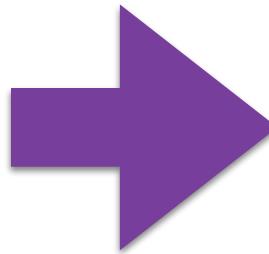
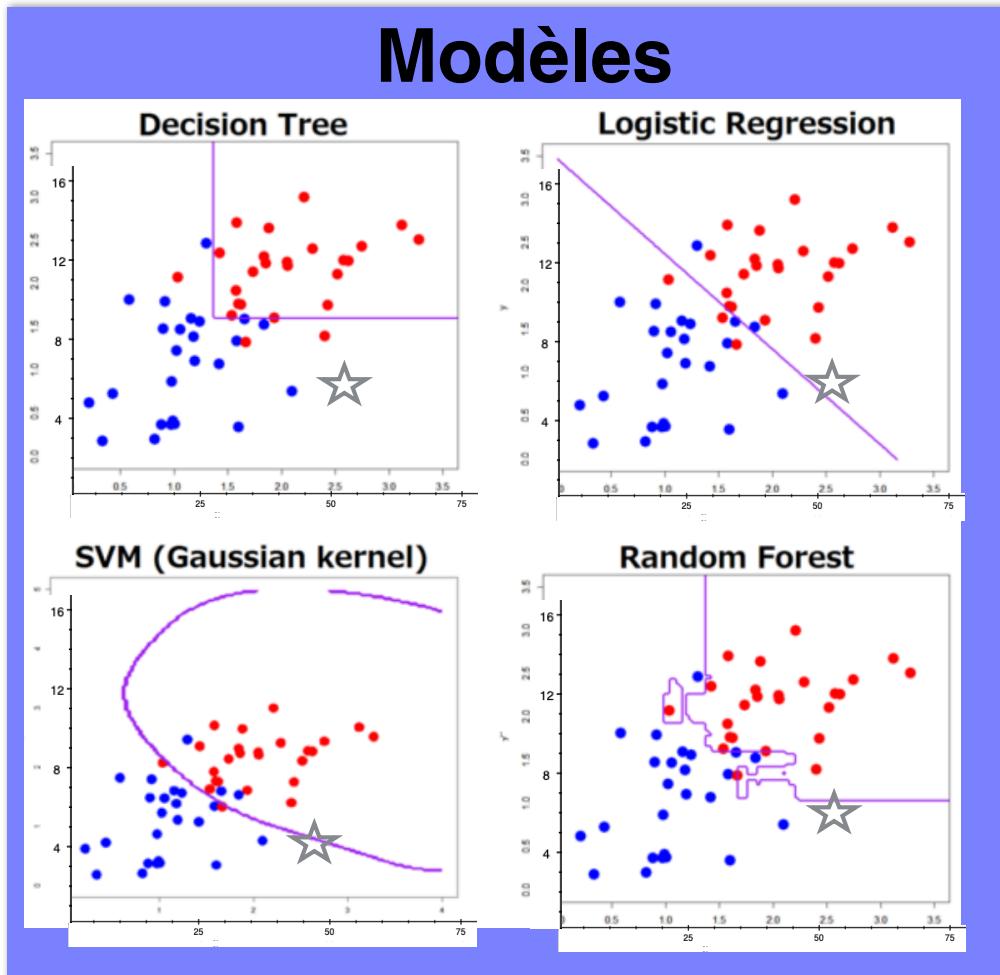


# Supervised ML: Learning a Model from Data



source: <http://tjo-en.hatenablog.com>

# Supervised ML: Learning a Model from Data Using this Model for Prediction



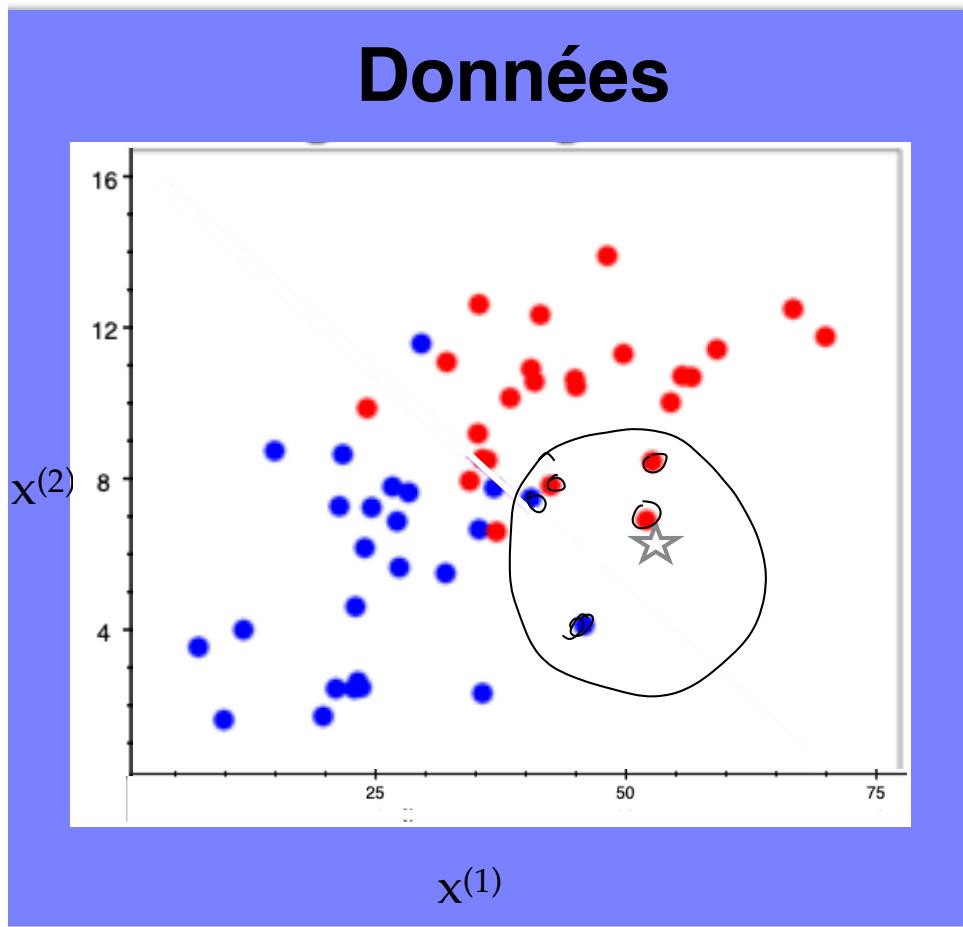
Age	G.Hemoglobin	Remission
55	6.1	?

D <: No  
LR: No  
SVM: No  
RF: No

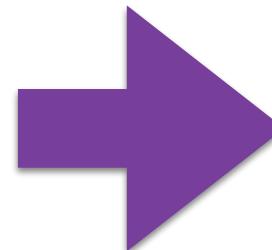
# Supervised ML: Lazy methods

Ex: k-nearest neighbors

$$k = 5$$



$h(x) = \begin{cases} \text{the majority label} \\ \text{over the } k \text{ nearest} \\ \text{neighbors of } x \end{cases}$



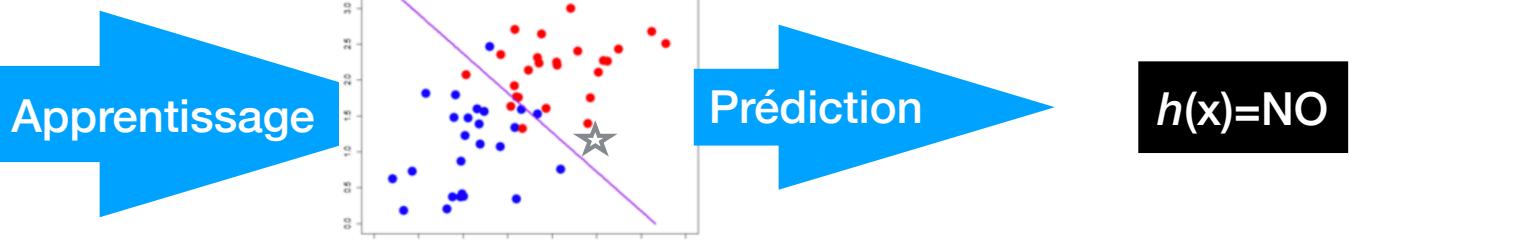
Age	G.Hemoglobin	Remission
55	6.1	?

5NN: No

# Apprentissage Supervisé: Classification vs Régression

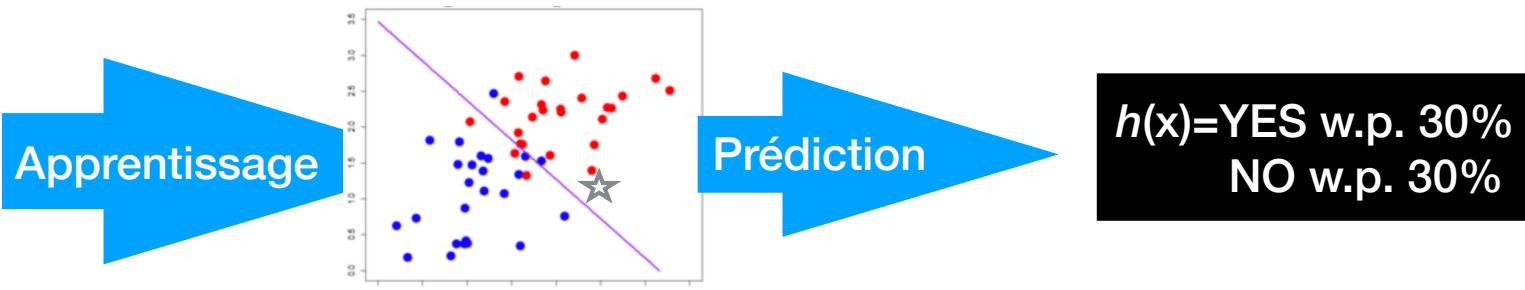
- Si les prédictions sont des catégories (Yes/No, Colors), c'est de la classification

Age	G.Hemoglobin	Remission
58	6.3	NO
42	9.9	NO
19	5.4	YES
⋮	⋮	⋮

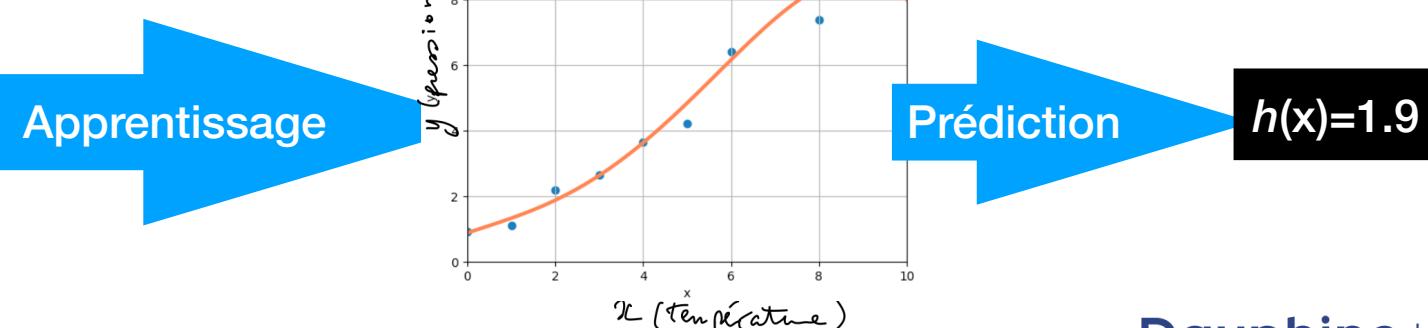


- Si les prédictions sont des nombres, c'est de la régression

Age	G.Hemoglobin	Remission
58	6.3	NO
42	9.9	NO
19	5.4	YES
⋮	⋮	⋮



Température	Pression
2	2.1
1.1	0.9
6	6.2
⋮	⋮



# Application de l'apprentissage supervisé et Types de Données

- **Banque/Assurance/Commerce (tabular data):**

- Prédire la solvabilité d'un individu
- Prédire combien un client potentiel peut rapporter
- Détection de fraude

exemple n°1 →

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Variables, attributs explicatifs      étiquette, classe, Variable à prédire

categorical      categorical      continuous      class

# Application de l'apprentissage supervisé et Types de Données

- **Vision**

- Reconnaissance d'images, d'objet dans les images
- Transformer une image

x	y
	Chien
	Pandas
	Bus

x	y

- **Chimie**

- Prédiction de mutagénécité

- **Texte**

- Traduction de texte d'une langue vers une autre
- Texte->parole
- parole->texte
- « Analyse de sentiment »

- **Robotique:**

- Voiture autonome
- Robotique humanoïde

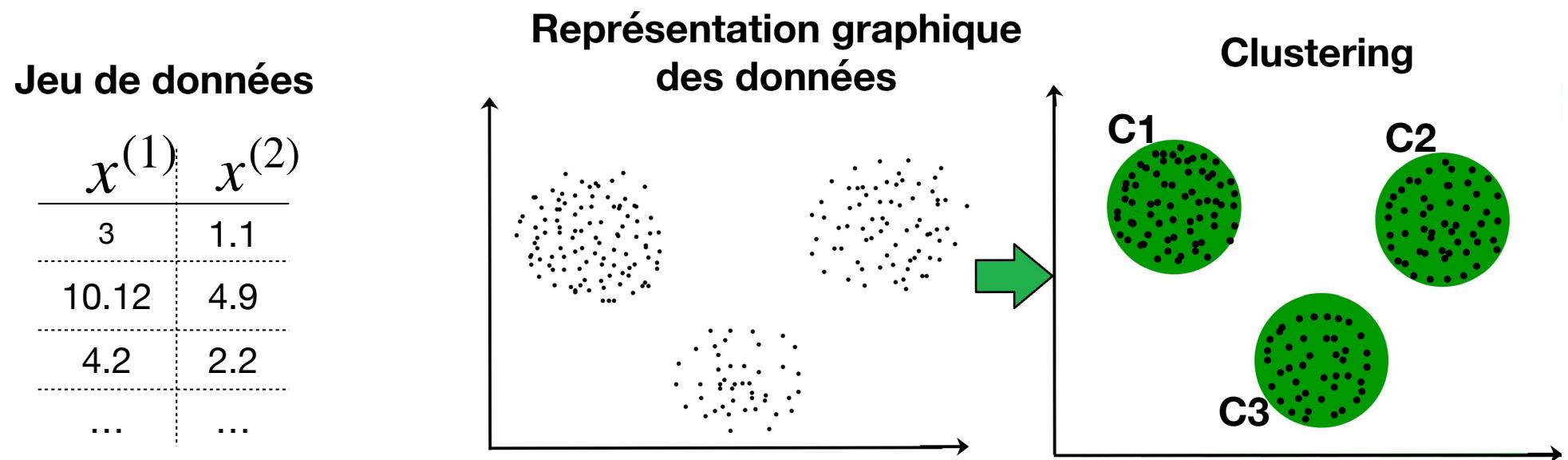
x	y
<chem>c1ccc2c(c1)oc(=O)c2</chem>	Active
<chem>c1ccc2c(c1)oc(=O)c2</chem>	Inactive
<chem>c1ccc2c(c1)oc(=O)c2</chem>	Active

x	y
My tailor is rich	Mon tailleur est riche
The cat eats the mouse	Le chat mange la souris
...	...

# Apprentissage *non* supervisé

- **Clustering**

Pas de supervision (juste  $x_1 \dots x_N$ , pas d'étiquette y)



- *Application:* prédire des profils types de clients

# Apprentissage *non* supervisé

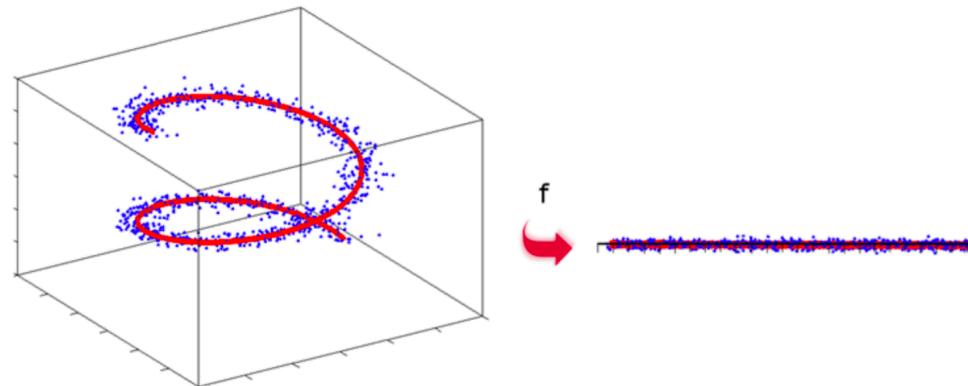
- **Réduction de dimension**

Pas de supervision (juste  $x_1 \dots x_N$ , pas d'étiquette y)

Jeu de données

	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$
3	1.1	2.1	
10.1	4.9	5.2	
...			
4.2	2.2	23	
...	...		

Représentation graphique des données



Représentation graphique des données en basse dimension

Jeu de données en basse dimension

	$z^{(1)}$
21	
4.3	
4.9	
...	

- *Application:* visualiser, préparer des données, compresser

# Apprentissage *non* supervisé

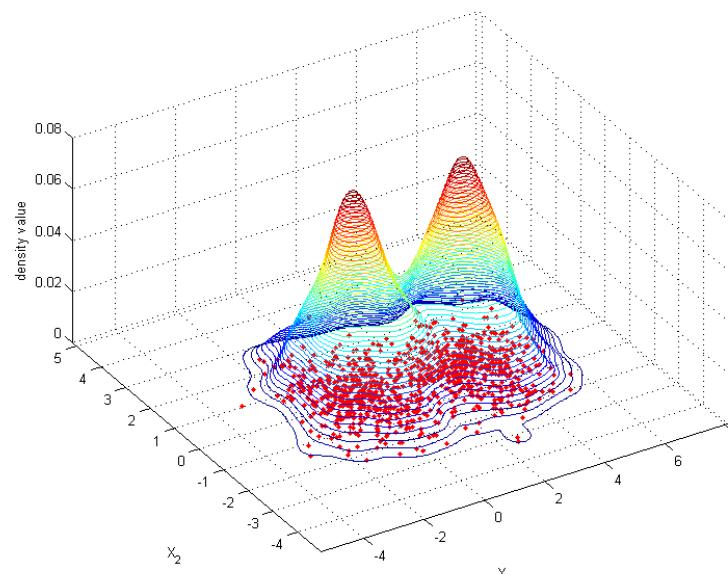
- **Apprentissage d'une Distribution**

Pas de supervision (juste  $x_1 \dots x_N$ , pas d'étiquette y)

Jeu de données

	$x^{(1)}$	$x^{(2)}$
3	1.1	
10.12	4.9	
4.2	2.2	
...	...	

Apprentissage



Génération

Estimation  
de densité

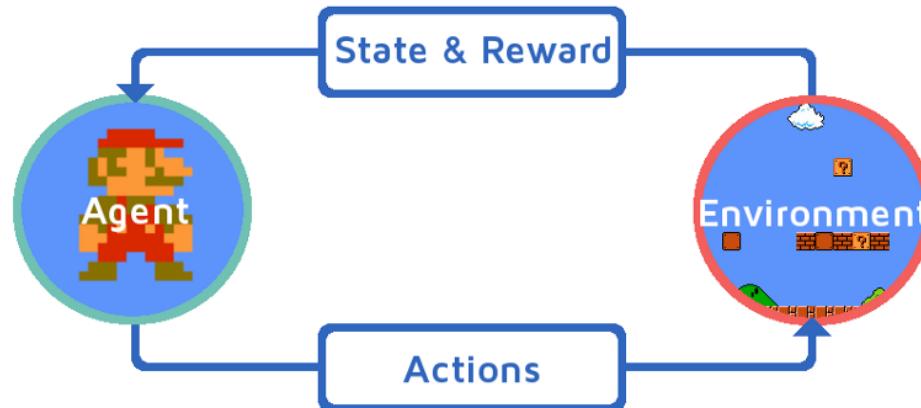
$$p(x^{(1)}, x^{(2)})$$

- *Applications:* détecter des intrusions dans des réseaux

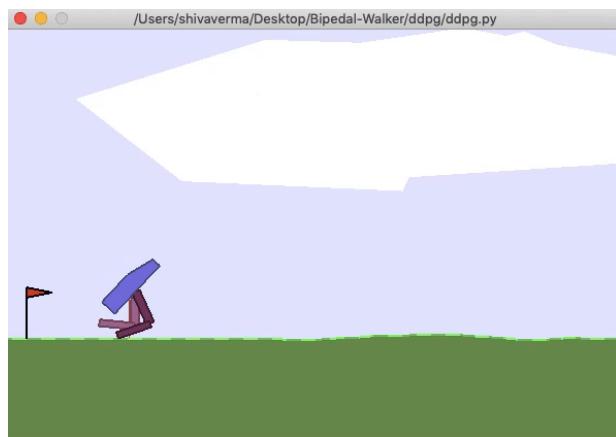
[https://commons.wikimedia.org/wiki/File:Bivariate\\_example.png](https://commons.wikimedia.org/wiki/File:Bivariate_example.png)

# Apprentissage par Renforcement

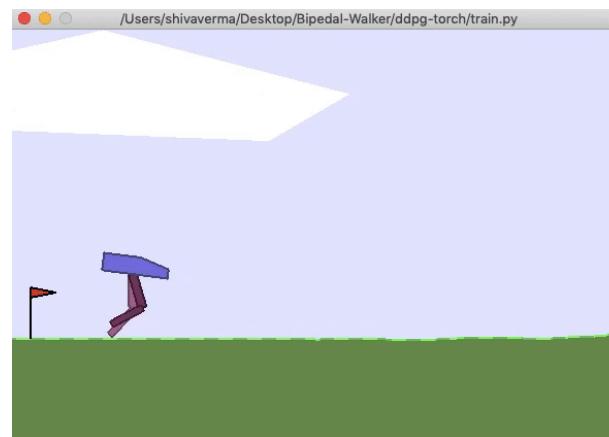
- Apprentissage par essais/erreurs



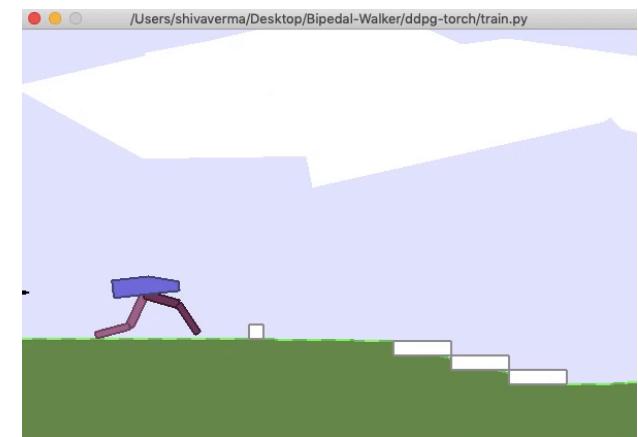
- Ex: apprendre à marcher, à gagner à des jeux vidéos, à exécuter une tâche dans une usine, à conduire une voiture automatiquement



1ière itération



500 itérations

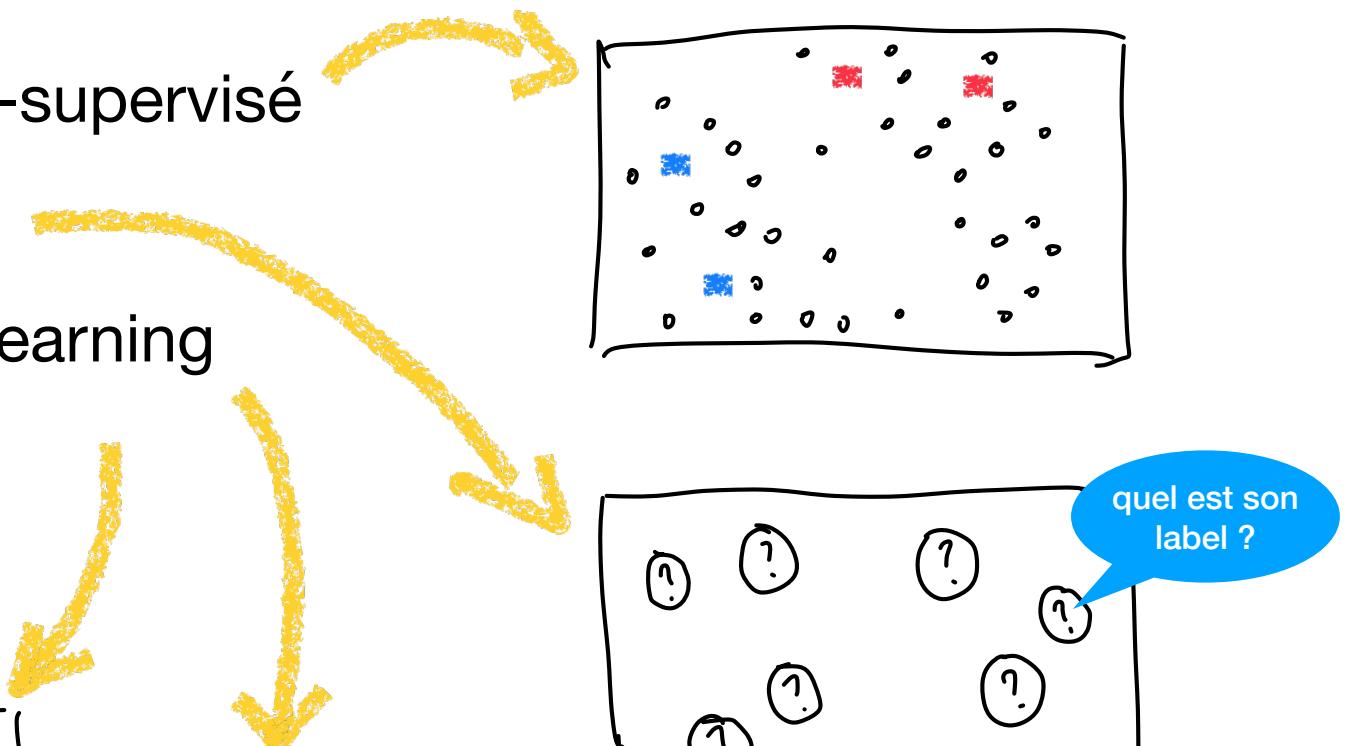
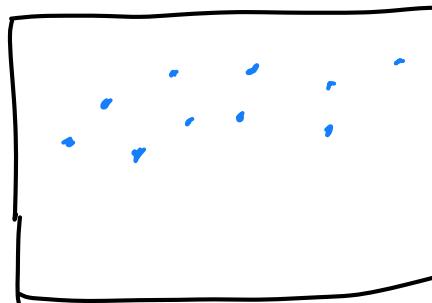
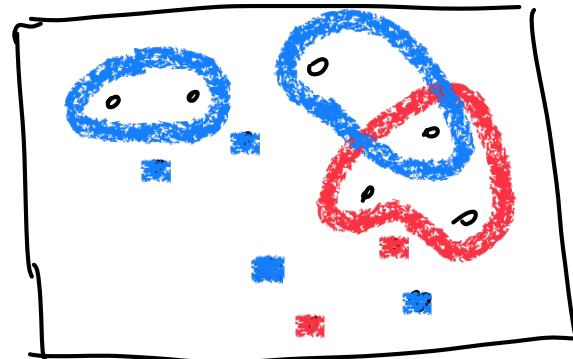


+1000 itérations

<https://www.freecodecamp.org/news/a-brief-introduction-to-reinforcement-learning-7799af5840db/>  
<https://towardsdatascience.com/teach-your-ai-how-to-walk-5ad55fce8bca>

# Nombreux autres paradigmes d'apprentissage

- Apprentissage semi-supervisé
- Apprentissage actif
- Weakly supervised learning
- few-shot learning
- ...



# Références

- S. Schwartz,  
*Understanding machine Learning*
- T. Hastie,  
*The Elements of Statistical Learning*
- C. Bishop,  
*Pattern recognition and Machine Learning*
- Mehryar Mohri  
*Foundations of Machine Learning*
- G. Saporta  
*Probabilités, Analyse de données et Statistiques*
- A. Cornuéjols  
*Apprentissage Artificiel, concepts et algorithmes*
- C.A. Azencott  
*Introduction au Machine Learning*

