

# Machine Learning

---

Constance Morel

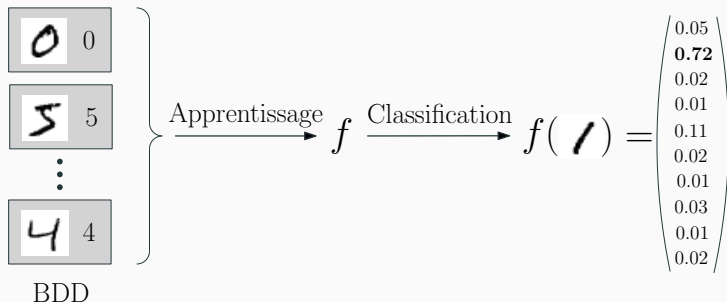
# Introduction

---

# Types d'apprentissage

- Apprentissage supervisé
- Apprentissage non supervisé
- Apprentissage par renforcement
- Système de recommandations

# Apprentissage supervisé



## Applications :

- Imagerie : reconnaissance d'écriture manuscrite, d'objets
- Voix : reconnaissance de la parole
- Texte : traduction, filtrage du courrier

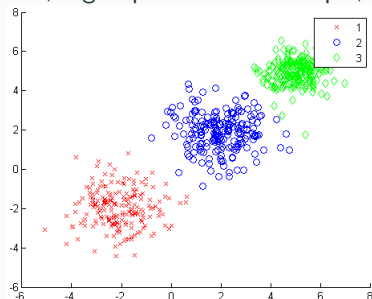
**Algorithmes :** SVM, réseaux de neurones, algorithmes génétiques, arbres de décision, ...

# Apprentissage non supervisé

## Partitionnement de données :

- Analyse de l'ADN
- Analyse marketing (des clients)
- Analyse des réseaux sociaux

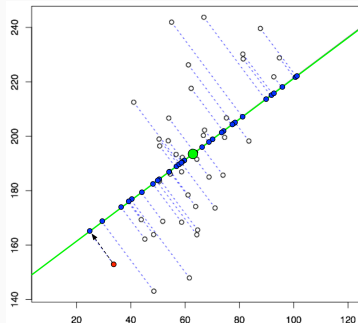
**Algorithmes :** k-means, algorithmes EM, regroupement hiérarchique, ...



## Réduction de dimension :

- Compression de données
- Extraction de caractéristiques

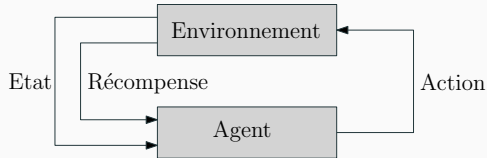
**Algorithmes :** PCA, LDA, GDA, ...



## Applications :

- Apprentissage des robots
- Jeux : backgammon, dames, go

**Algorithmes :** Q-learning, Hidden Markov Model, ...



# Système de recommandations



						
	2			4	5	
	5		4			1
			5		2	
		1		5		4
			4			2
	4	5		1		

## Applications :

- Recommandation de films et livres
- Marketing personnalisé
- Publicité

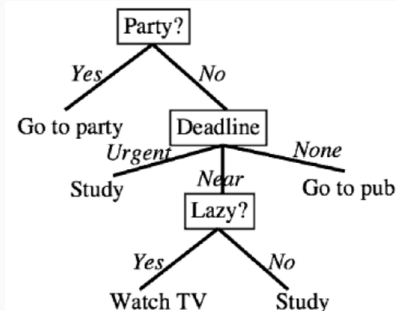
**Algorithmes :** collaborative filtering, content-based filtering, hybrid recommender system

# Arbre de décision ID3

---



Deadline	Soirée	Flemmard	Activité
Urgent	Oui	Oui	Soirée
Urgent	Non	Oui	Etude
Proche	Oui	Oui	Soirée
Non	Oui	Non	Soirée
Non	Non	Oui	Pub
Non	Oui	Non	Soirée
Proche	Non	Non	Etude
Proche	Non	Oui	TV
Proche	Oui	Oui	Soirée
Urgent	Non	Non	Etude



# Entropie

Entropie : quantité d'impureté :  $H(p) = \sum_i -p_i \log p_i$

Act1	Act2	Act3	
Soirée	Soirée	Soirée	$H(Act1) = -\frac{5}{10} \log \frac{5}{10} - \frac{3}{10} \log \frac{3}{10} - 2 \times \frac{1}{10} \log \frac{1}{10}$ $= 1.68$
Etude	Soirée	Soirée	
Soirée	Soirée	Soirée	
Soirée	Soirée	Etude	$H(Act2) = -1 \log(1) - 0 \log(0) - 0 \log(0) - 0 \log(0)$ $= 0$
Pub	Soirée	Etude	
Soirée	Soirée	Pub	$H(Act3) = -2 \times \frac{3}{10} \log \frac{3}{10} - 2 \times \frac{2}{10} \log \frac{2}{10}$ $= 1.97$
Etude	Soirée	Pub	
TV	Soirée	Pub	
Soirée	Soirée	TV	
Etude	Soirée	TV	

# Gain d'information

Gain d'information : évaluer la chute d'entropie si nous sélectionnons une caractéristique  $F$  :  $G(S, F) = H(S) - \sum_{f \in F} \frac{|S_f|}{|S|} H(S_f)$

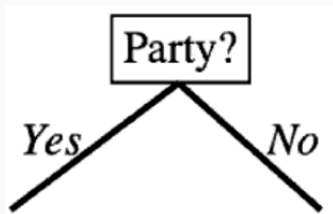
Deadline	Soirée	Flemmard	Act1	
Urgent	Oui	Oui	Soirée	$H(Urgent) = -\frac{1}{3} \log \frac{1}{3} - \frac{2}{3} \log \frac{2}{3}$ $= 0.92$
Urgent	Non	Oui	Etude	
Proche	Oui	Oui	Soirée	
Non	Oui	Non	Soirée	$H(Non) = -\frac{1}{3} \log \frac{1}{3} - \frac{2}{3} \log \frac{2}{3}$ $= 0.92$
Non	Non	Oui	Pub	
Non	Oui	Non	Soirée	
Proche	Non	Non	Etude	$H(Proche) = -\frac{2}{4} \log \frac{2}{4} - 2 \times \frac{1}{4} \log \frac{1}{4}$ $= 1.5$
Proche	Non	Oui	TV	
Proche	Oui	Oui	Soirée	
Urgent	Non	Non	Etude	

$$Gain(S, Deadline) = 1.68 - \frac{3}{10} \times 0.92 - \frac{3}{10} \times 0.92 - \frac{4}{10} \times 1.5 = 0.53$$

$$Gain(S, Soiree) = 1.0$$

$$Gain(S, Flemmard) = 0.21$$

# Début de l'arbre



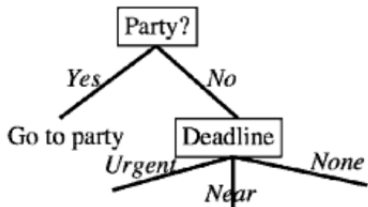
Deadline	Flemmard	Activité
Urgent	Oui	Soirée
Proche	Oui	Soirée
Non	Non	Soirée
Non	Non	Soirée
Proche	Oui	Soirée

Deadline	Flemmard	Activité
Urgent	Oui	Etude
Non	Oui	Pub
Proche	Non	Etude
Proche	Oui	TV
Urgent	Non	Etude

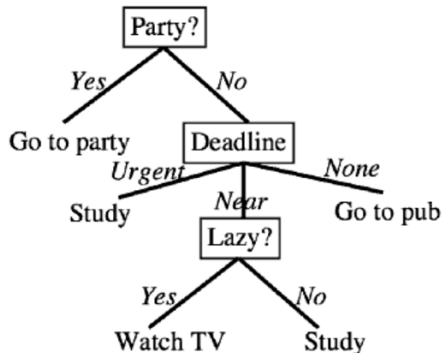
$$G(S, \text{Deadline}) = 0.97$$

$$G(S, \text{Flemmard}) = 0.17$$

# Arbre ID3



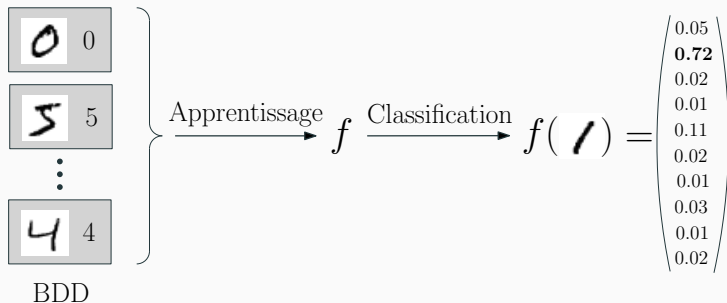
Deadline	Flemmard	Activité
Urgent	Oui	Etude
Urgent	Non	Etude
Non	Oui	Pub
Proche	Non	Etude
Proche	Oui	TV



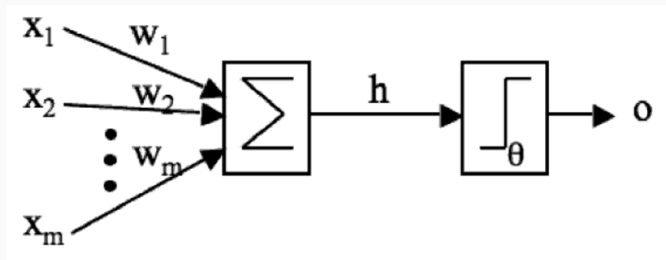
# Réseau de neurones

---

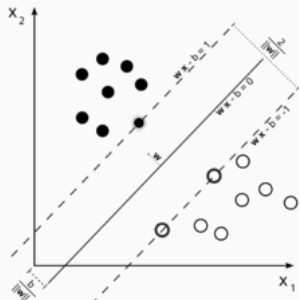
# Problème à résoudre



# Perceptron (1957)



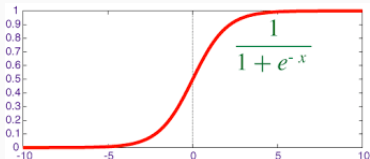
$$h_{W,b}(x) = w_1x_1 + w_2x_2 + b$$





# Evaluation de l'erreur

$$h_{W,b}(x) = w_1x_1 + w_2x_2 + b$$



$$\text{sigmoid} \left( h_{W,b} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) =$$
$$\text{sigmoid}(0.2) = 0.55$$

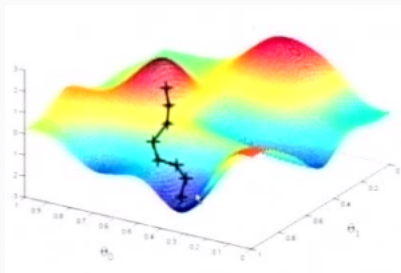
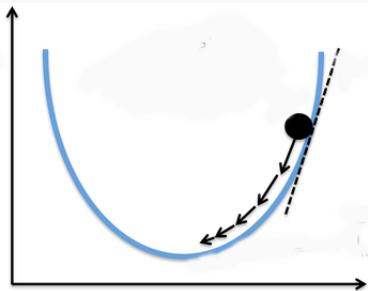
Résultat de la fonction :  $f(X) = \begin{pmatrix} 0.55 \\ 0.45 \end{pmatrix}$

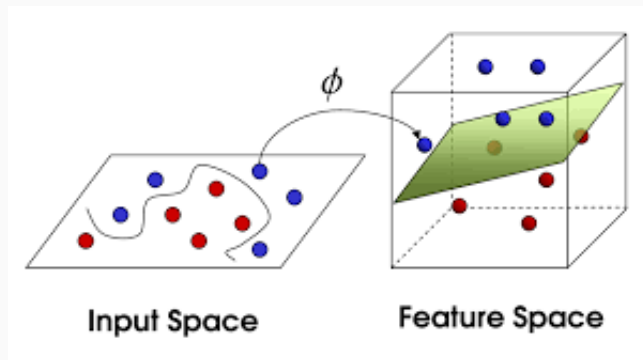
Résultat souhaité :  $y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Erreur sur une donnée :  $ED(f(X), y)^2$

Erreur sur la base d'apprentissage :  $\frac{1}{m} \sum_i ED(f(X_i), y_i)^2$

# Descente de gradient





# Multi-Layer Perceptron MLP

