

# Introduction au Machine Learning

De quoi s'agit-il ?

**Emmanuel Dellandréa**

**[emmanuel.dellandrea@ec-lyon.fr](mailto:emmanuel.dellandrea@ec-lyon.fr)**

Version du 06/01/2025



# Plan du cours

- ➔ ■ Introduction
- Classification non supervisée
- Modèles linéaires pour la régression
- Régression logistique pour la classification
- Réseaux de neurones : representation et apprentissage
- Conseils pratiques pour l'apprentissage automatique
- Apprentissage par renforcement

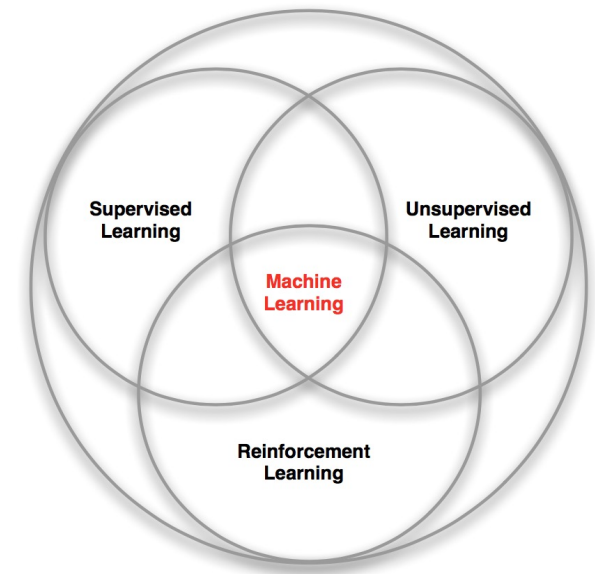
# Apprentissage automatique : définitions

- Arthur Samuel (1959) : « *Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed* ».
- Tom Mitchell (1998) : « *Well-posed Learning Problem: A computer program is said to learn from experience  $E$  with respect to some task  $T$  and some performance measure  $P$ , if its performance on  $T$ , as measured by  $P$ , improves with experience  $E$ .* »

# Apprentissage automatique

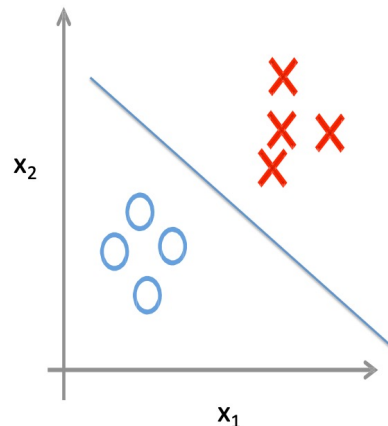
→ Elaboration d'algorithmes permettant d'automatiser la construction d'une fonction de prédiction

- 3 grandes familles d'algorithmes :
  - Apprentissage supervisé
  - Apprentissage non supervisé
  - Apprentissage par renforcement



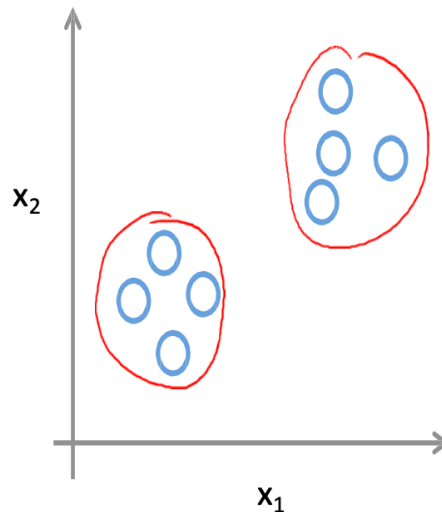
# Apprentissage supervisé

- Suppose que l'on dispose d'un ensemble d'exemples caractérisés par des variables prédictives, et dont on connaît les valeurs pour la variable cible
- L'objectif de l'apprentissage est de construire une fonction de prédiction permettant l'association entre les variables prédictives et la variable cible



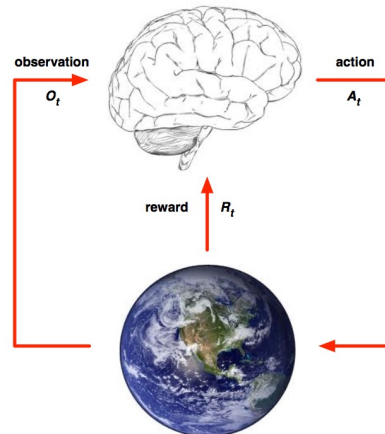
# Apprentissage non supervisé

- On ne dispose pas de valeurs pour la variable cible
- L'objectif est de réaliser le regroupement en catégories des exemples fournis en exemple



# Apprentissage par renforcement

- On ne dispose pas de valeurs pour la variable cible
- Un agent a la capacité d'agir
- Chaque action influence l'état futur de l'agent
- Le succès est mesuré par une valeur de récompense
- Objectif : sélectionner les actions pour maximiser la récompense future
- Apprentissage par essais/erreurs

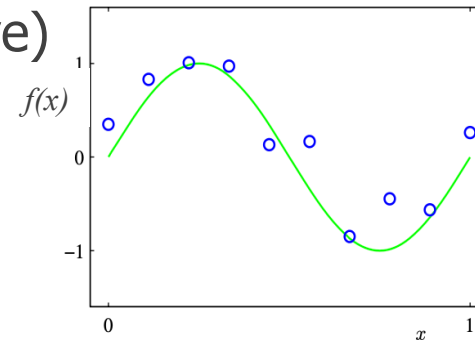


# Apprentissage supervisé

## → Principal paradigme étudié dans le cadre de ce cours

- Trouver une fonction de prédiction telle que :  $y = f(x)$

- $y$  : variable cible
- $x$  : variable prédictive (ou explicative)
- $f$  : fonction de prédiction



→ Régression

- Régression vs classification

- Régression :
  - La variable cible est quantitative
- Classification :
  - La variable cible est qualitative

$f(\text{apple}) = \text{"apple"}$

$f(\text{tomato}) = \text{"tomato"}$

→ Classification

$f(\text{cow}) = \text{"cow"}$



# Apprentissage supervisé

$$y = f(x)$$

- **Apprentissage :**

- ➔ Utiliser un ensemble de données labellisées  $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots (x^{(N)}, y^{(N)})\}$  afin d'estimer la fonction de prédiction  $f$  par minimisation de l'erreur de prédiction

- **Test :**

- ➔ Appliquer la fonction  $f$  sur des données  $x$  non utilisées pendant l'apprentissage pour obtenir les valeurs prédites

- **Erreur de prédiction :**

- Évalue la différence entre les valeurs prédites  $f(x^{(i)})$  et les vraies valeurs  $y^{(i)}$
  - **L'erreur calculée sur l'ensemble de test permet d'évaluer la performance du modèle**

# Evaluation de la performance

- Matrice de confusion

		Prédiction du modèle	
		Classe	Non classe
Réalité	Classe	VP	FN
	Non classe	FP	VN

- VP / VN : vrais positifs / vrais négatifs
- FP / FN : faux positifs / faux négatifs

# Evaluation de la performance

- Taux de classification : proportion d'éléments bien classés  
→  $(VP + VN) / (VP + VN + FP + FN)$
- Rappel (sensibilité) : proportion d'éléments bien classés par rapport au nombre d'éléments de la classe à prédire  
→  $VP / (VP + FN)$
- Précision : proportion d'éléments bien classés pour une classe donnée  
→  $VP / (VP + FP)$
- F-mesure : mesure de compromis entre précision et rappel  
→  $2 * (Rappel * Précision) / (Rappel + Précision)$

		Prédiction du modèle	
		Classe	Non classe
Réalité	Classe	VP	FN
	Non classe	FP	VN

# Evaluation de la performance

- Pour un modèle de régression :
  - Erreur quadratique moyenne
    - ➔ Moyenne des différences au carré entre les vraies valeurs et les valeurs prédites
  - Coefficient de corrélation de Pearson
    - ➔ Mesure la relation linéaire ("proportionnalité") entre les valeurs réelles et les valeurs prédites

# Exemples d'application

- La conduite autonome
- La génération de texte, d'images, de vidéos
- La reconnaissance automatique de la parole
- Le traitement des langues naturelles
- Les systèmes de recommandation
- La vision par ordinateurs
- La robotique
- La médecine
- La finance
- Les assurances
- ...



ÉCOLE  
**CENTRALE** LYON

36, avenue Guy de Collongue 69130 Écully - France  
+33 (0)4 72 18 60 00

[www.ec-lyon.fr](http://www.ec-lyon.fr)