



# Machine Learning

**Prof. Meryam ZERYOUH**

[zeryouh.meryam@gmail.com](mailto:zeryouh.meryam@gmail.com)

2024/2025

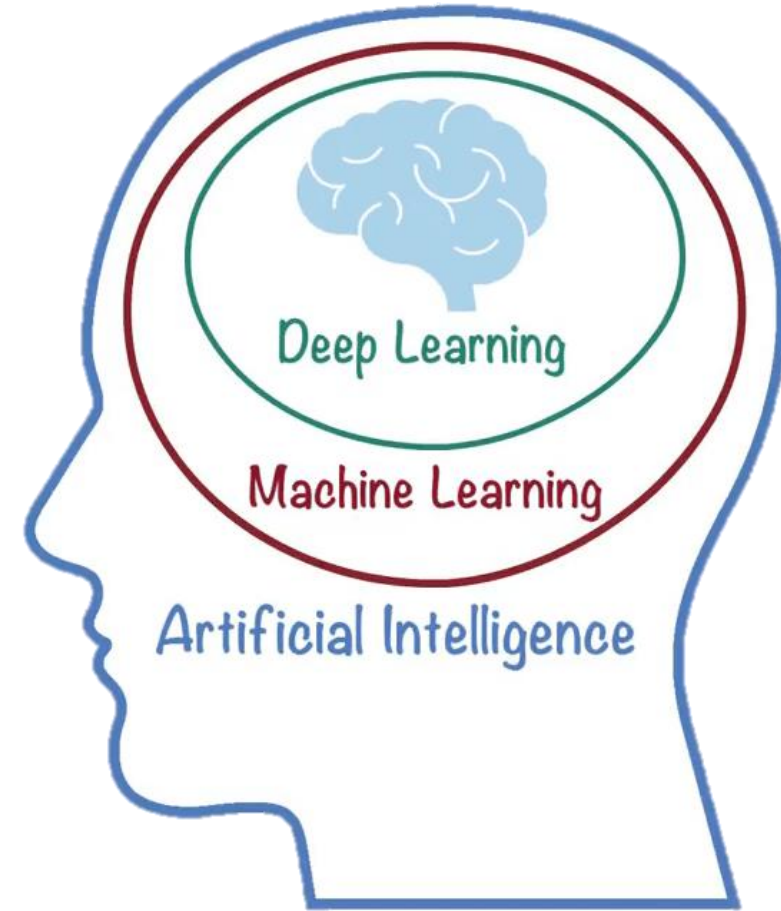
# Acquis d'apprentissage du cours

## ❑ Etre capable :

- ✓ Expliquer les principes du machine learning, du deep learning et de l'intelligence artificielle.
- ✓ Transformer un problème métier à un problème de **machine learning**.
- ✓ Mettre en oeuvre les techniques du **machine learning** et de **deep learning**.

# CHAPITRE 1 :

INTRODUCTION AU CONCEPTS D'INTELLIGENCE  
ARTIFICIELLE - ML



# Artificial Intelligence

## Comment l'IA est utilisée dans notre vie quotidienne ?

- ✓ Déterminer l'itinéraire le plus court selon le trafic,
- ✓ Estimer l'heure d'arrivée,
- ✓ Evaluer le trafic en temps réel,
- ✓ Repérer d'éventuelles zones de danger (accidents, travaux, radars),
- ✓ ...



# Artificial Intelligence

## Comment l'IA est utilisée dans notre vie quotidienne ?

- ✓ Réaliser des actions personnalisées selon les commandes de l'utilisateur
- ✓ Contrôler les objets connectés dans son environnement.
- ✓ ...



# Artificial Intelligence

## Comment l'IA est utilisée dans notre vie quotidienne ?

- ✓ Contrôler le flux,
- ✓ Adapter et transmettre les informations les plus pertinentes,
- ✓ Affiner les sélections et proposer un contenu optimal à l'internaute,
- ✓ ...



# Artificial Intelligence

Comment l'IA est utilisée dans notre vie quotidienne ?

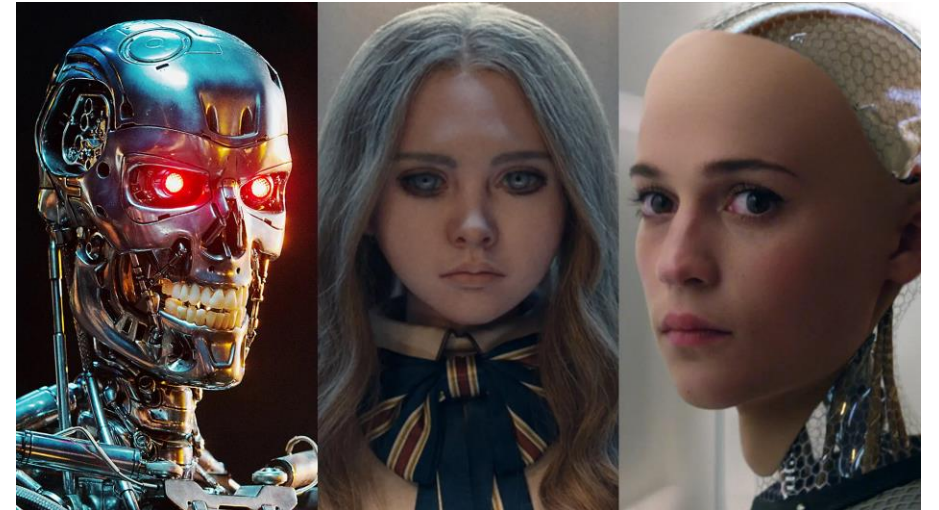
- ✓ Recommandation personnalisée de contenu
- ✓ La société exploite l'IA pour créer des teaser, des bande-annonce, etc.
- ✓ ...

The Netflix logo, consisting of the word "NETFLIX" in a bold, red, sans-serif font.

# Artificial Intelligence

## Comment l'IA est utilisée dans notre vie quotidienne ?

- ✓ Planifier et trouver les lieux de tournage
- ✓ Prendre des décisions de casting afin de déterminer le meilleur acteur pour un rôle donné
- ✓ Utiliser pour vieillir des personnages et créer des arrière-plans numériques et augmenter l'éclairage
- ✓ Composer la musique
- ✓ ...

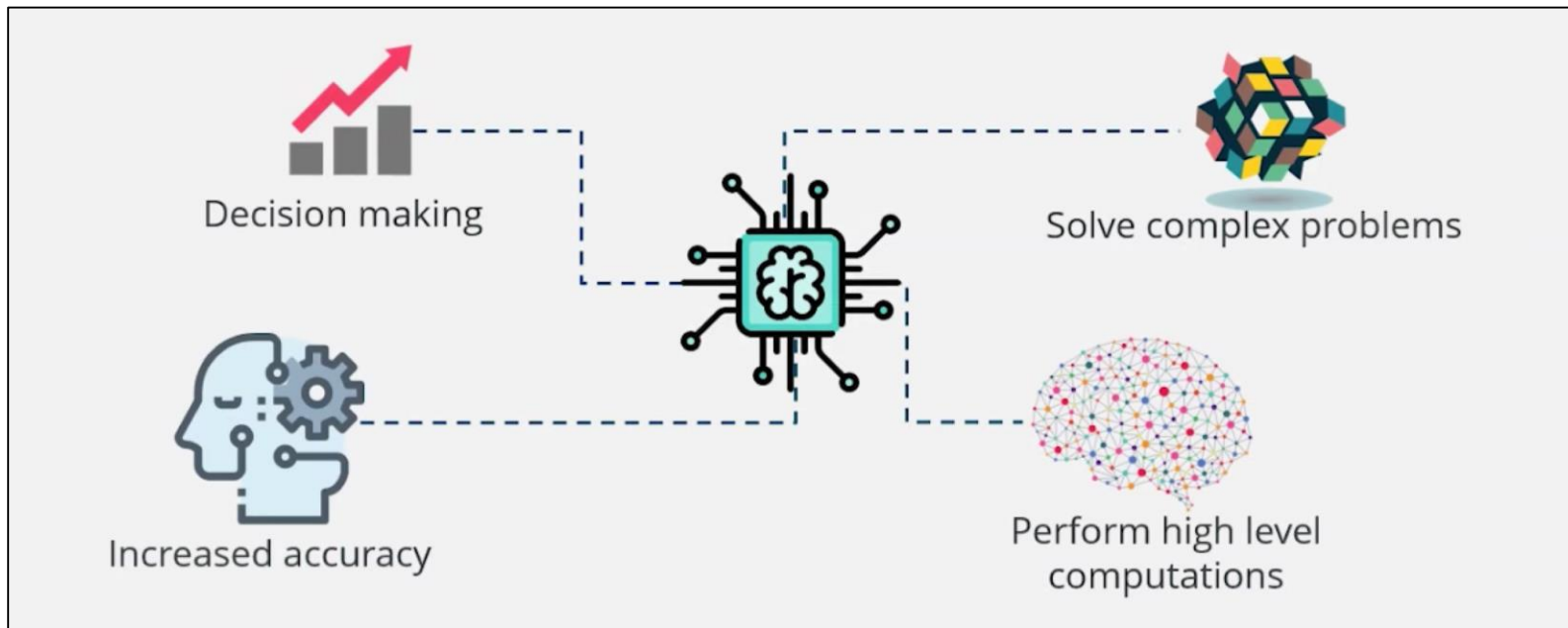




# Artificial Intelligence

## De quoi s'agit-il ?

Le développement de systèmes informatiques capables d'effectuer des tâches **qui nécessitent normalement l'intelligence humaine**, comme la prise de décision, la détection d'objets, la résolution de problèmes complexes, etc.

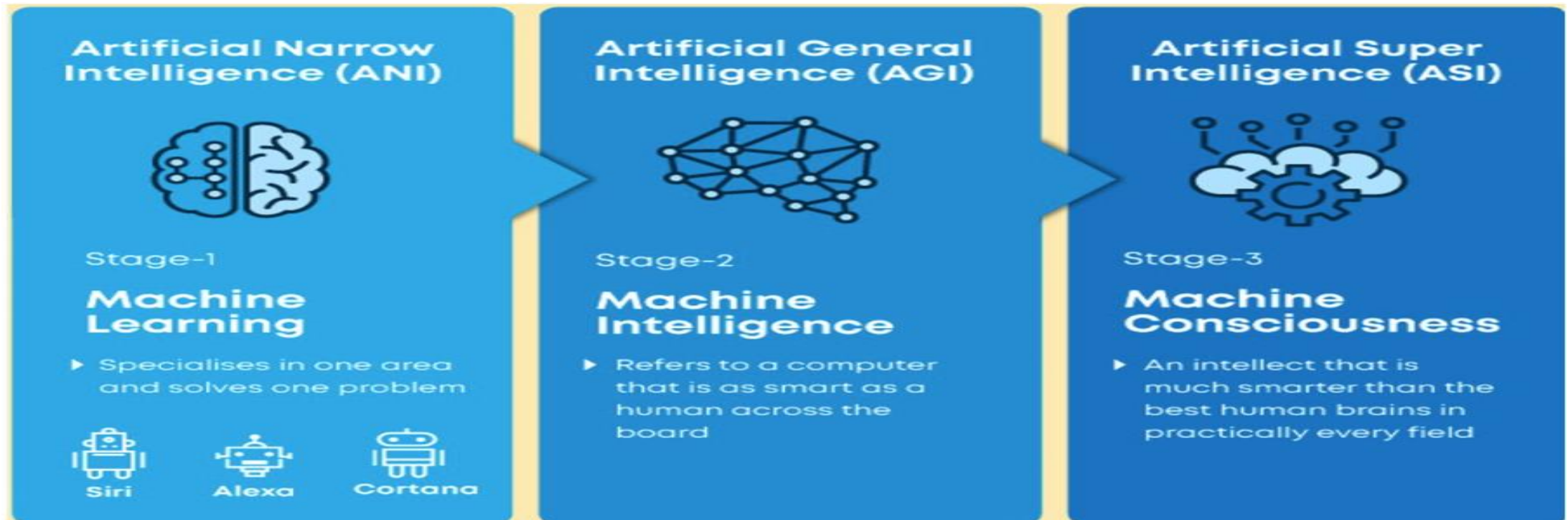


# Artificial Intelligence Stages

## 1. Intelligence artificielle étroite

## 2. Intelligence générale artificielle

## 3. Super intelligence artificielle



# Artificial Intelligence Stages

## 1. Intelligence artificielle étroite

- L'IA faible
- **Le seul type d'intelligence artificielle que nous ayons réussi à réaliser à ce jour.**
- Orientée vers un but précis, conçue pour accomplir des tâches uniques – par exemple, la reconnaissance faciale, la reconnaissance vocale/les assistants vocaux, la conduite d'une voiture ou la recherche sur Internet.
- **Siri n'est pas une machine consciente qui répond à nos requêtes.**
- Au contraire, ce que Siri est capable de faire - ce pour quoi il a été conçu - c'est de traiter le langage humain, de le saisir dans un moteur de recherche (Google) et de nous renvoyer des résultats.

# Artificial Intelligence Stages

## 1. Intelligence artificielle étroite

### ❑ Exemples :

- ✓ Siri d'Apple, Alexa d'Amazon, Cortana de Microsoft et autres assistants virtuels.
- ✓ Logiciels de reconnaissance d'images et de visages
- ✓ Outils de cartographie et de prédiction des maladies
- ✓ Robots de fabrication et drones
- ✓ Filtres anti-spam pour les courriels / outils de surveillance des médias sociaux pour les contenus dangereux
- ✓ Recommandations de contenu de divertissement ou de marketing basées sur le comportement de visionnage, d'écoute ou d'achat.
- ✓ Voitures à conduite autonome

# Artificial Intelligence Stages

## 2. Intelligence générale artificielle

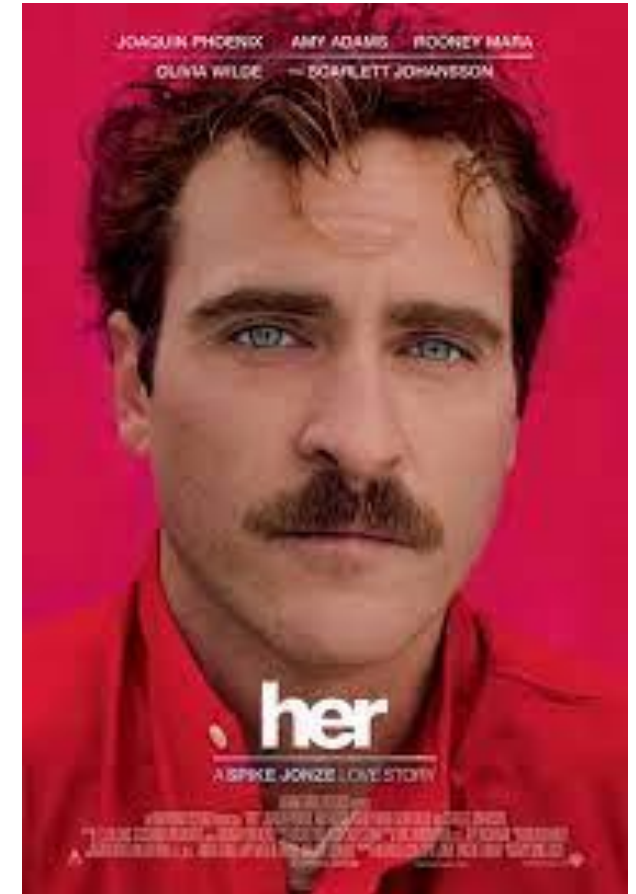
- IA forte ou IA profonde
- Stade de l'IA où les machines auront **la capacité de penser et de prendre des décisions** comme les humains.
- le concept d'une machine dotée d'une intelligence générale qui imite l'intelligence et/ou les comportements humains, avec la capacité d'apprendre et d'appliquer son intelligence pour résoudre tout problème
- L'IA peut penser, comprendre et agir d'une manière indiscernable de celle d'un être humain dans une situation donnée.
- Les chercheurs et les scientifiques spécialisés dans l'IA n'ont pas encore atteint une IA forte. Pour y parvenir, ils devraient **trouver un moyen de rendre les machines conscientes**, en programmant un ensemble complet de capacités cognitives.

# Artificial Intelligence Stages

## 2. Intelligence générale artificielle

➤ L'IA devrait être capable de raisonner, de résoudre des problèmes, de juger dans l'incertitude, de planifier, d'apprendre, d'intégrer des connaissances antérieures dans la prise de décision et d'être innovante, imaginative et créative.

➤ **Conscience.**

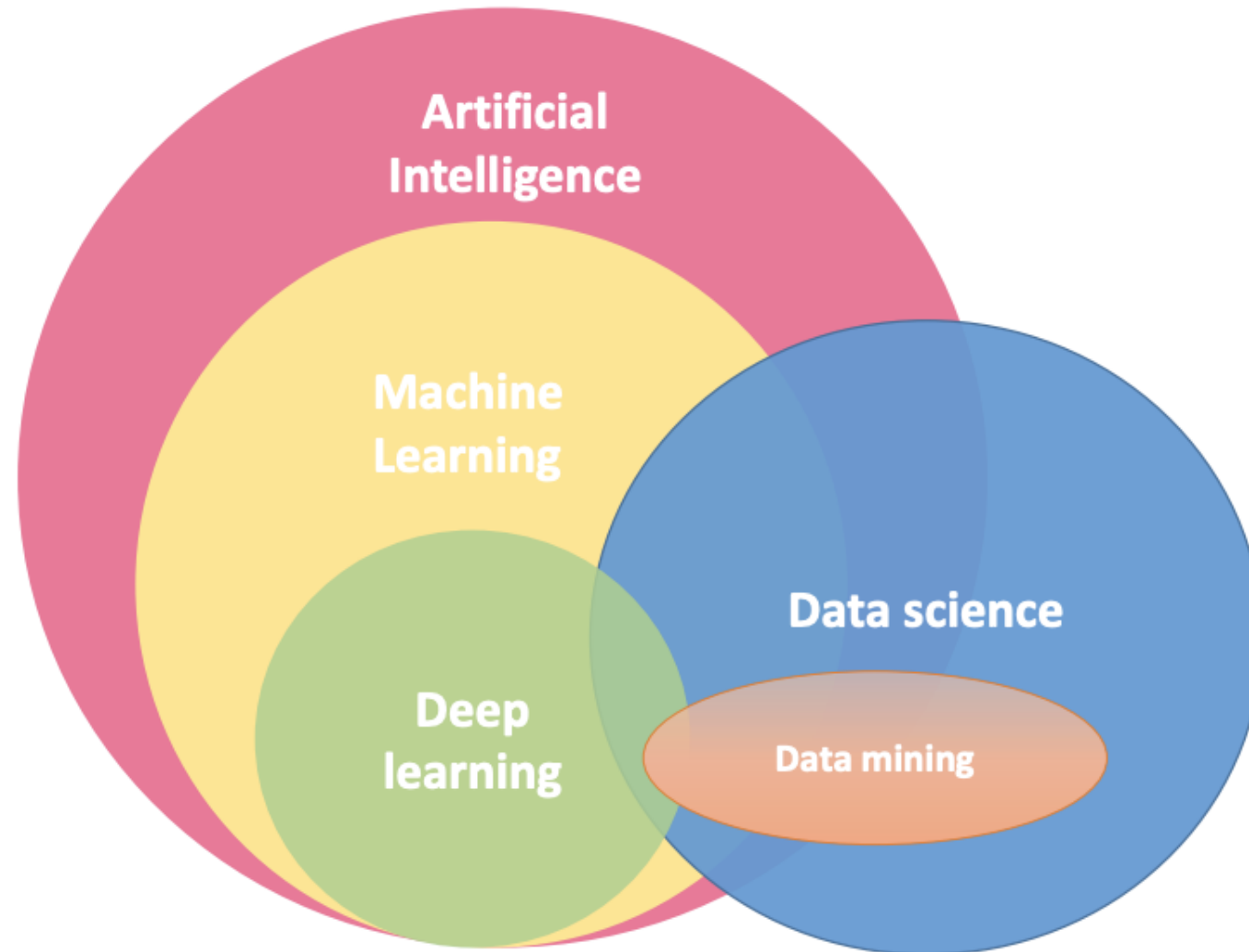


# Artificial Intelligence Stages

## 3. Super Intelligence artificielle

- Stade de l'IA où les capacités des ordinateurs **dépasseront** celles des êtres humains.
- (ASI) ne se contente pas d'imiter ou de comprendre l'intelligence et le comportement humains. l'ASI est l'endroit où les machines deviennent conscientes d'elles-mêmes et dépassent la capacité de l'intelligence et des aptitudes humaines.
- L'ASI aurait une meilleure mémoire et une capacité plus rapide à traiter et à analyser les données et les stimuli. Par conséquent, les capacités de prise de décision et de résolution de problèmes des êtres super intelligents seraient bien supérieures à celles des êtres humains.

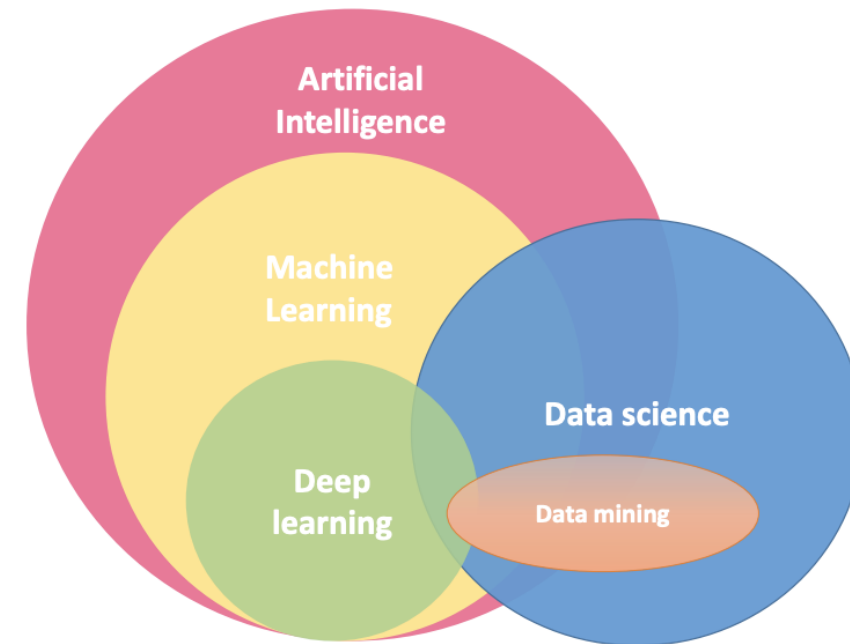
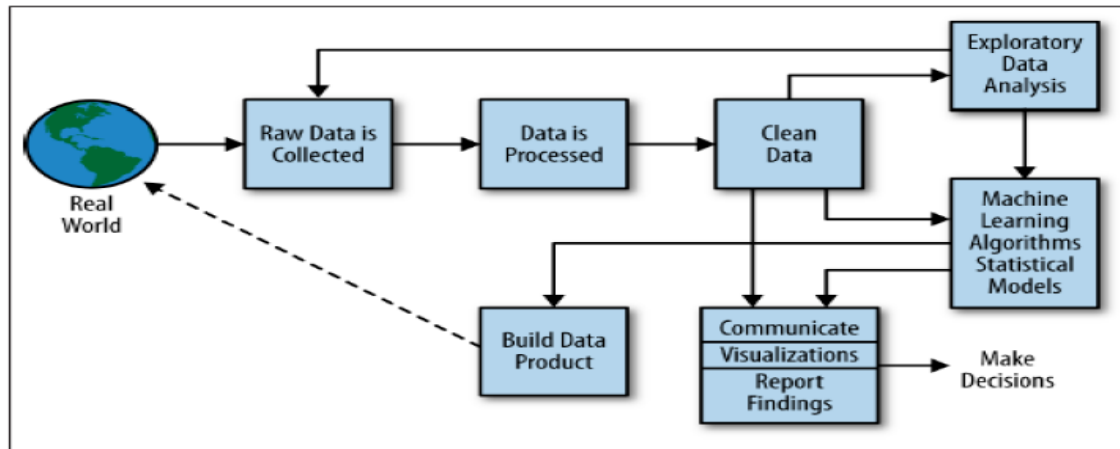
# AI concepts





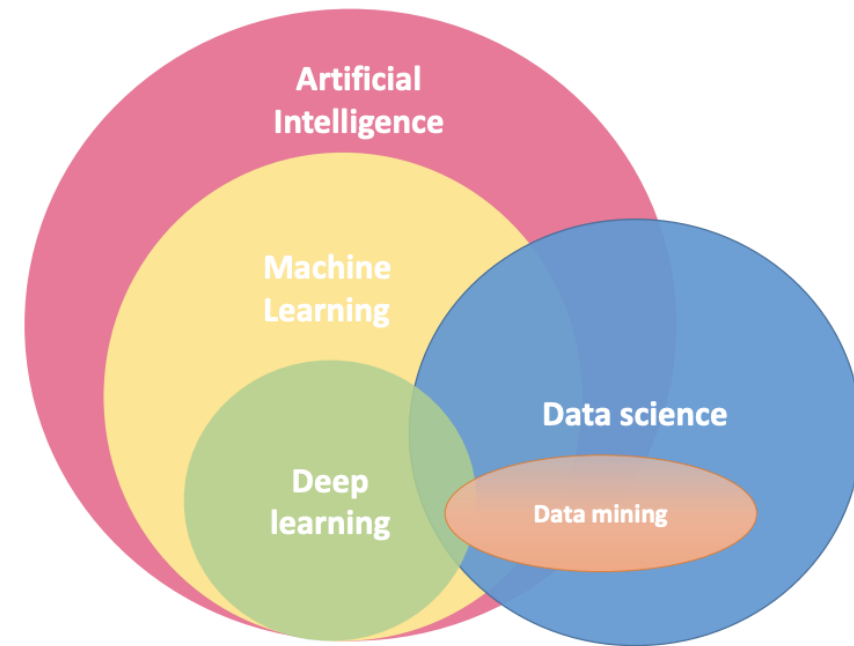
# AI concepts : Data science

- **Data science** est un domaine **interdisciplinaire** qui utilise des méthodes, des processus et des algorithmes pour extraire des connaissances et des idées à partir de nombreuses **données structurées et non structurées**.
- Les différentes étapes et procédures de la science des données impliquent l'extraction, la manipulation, la visualisation et la maintenance des données pour prévoir l'occurrence d'événements futurs.



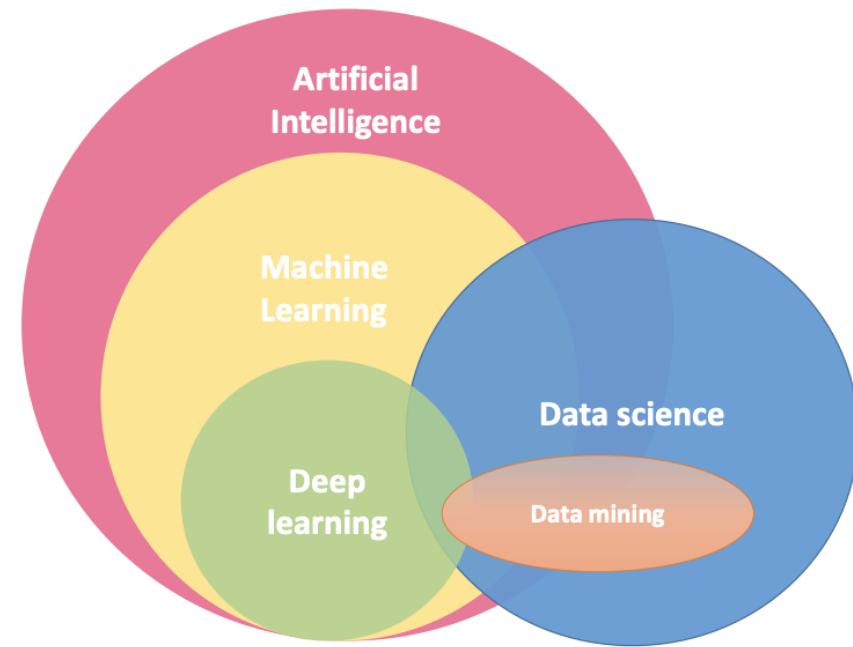
# AI concepts : Data mining

- **Data mining ou exploration de données** est considérée comme le processus d'extraction d'informations utiles à partir d'une grande quantité de données. Il est utilisé pour découvrir des modèles nouveaux, précis et utiles dans les données, en recherchant le sens et les informations pertinentes pour l'organisation ou l'individu qui en a besoin. Il s'agit d'un outil utilisé par les humains.
- L'exploration de données traite principalement des **données structurées** pour trouver des informations utiles dans un ensemble de données et à utiliser ces informations pour **découvrir des modèles cachés**.



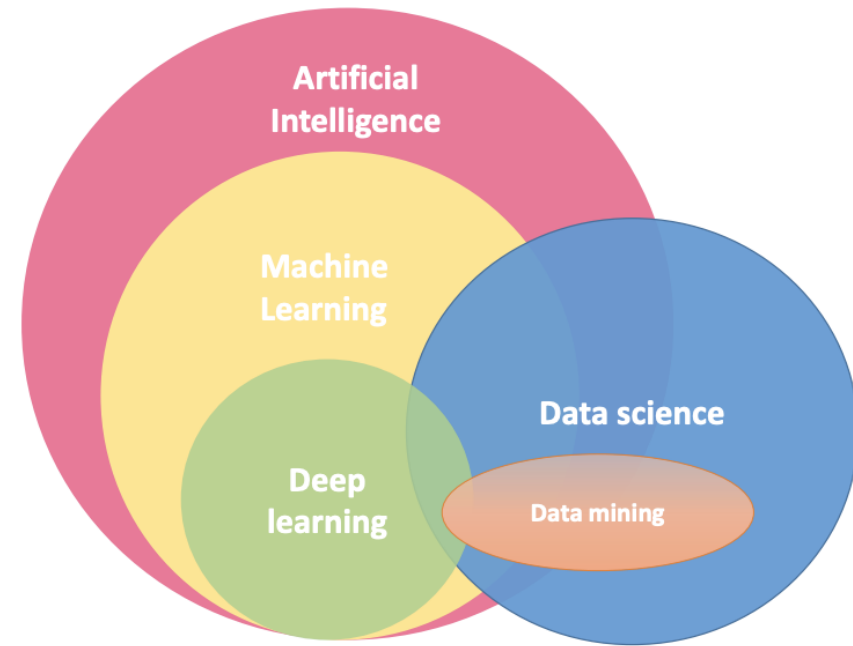
# AI concepts : Machine learning

- **Le Machine learning** est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches statistiques pour donner aux ordinateurs la capacité **d' « apprendre » à partir de données.**
- Apprendre aux machines en leur fournissant des données et en les laissant apprendre par elles-mêmes sans aucune intervention humaine.
- L'apprentissage automatique utilise **des méthodes d'exploration de données** et des algorithmes pour construire des modèles sur la logique derrière les données qui prédisent les résultats futurs. Les algorithmes sont basés sur les mathématiques et les langages de programmation.



# AI concepts : Deep learning

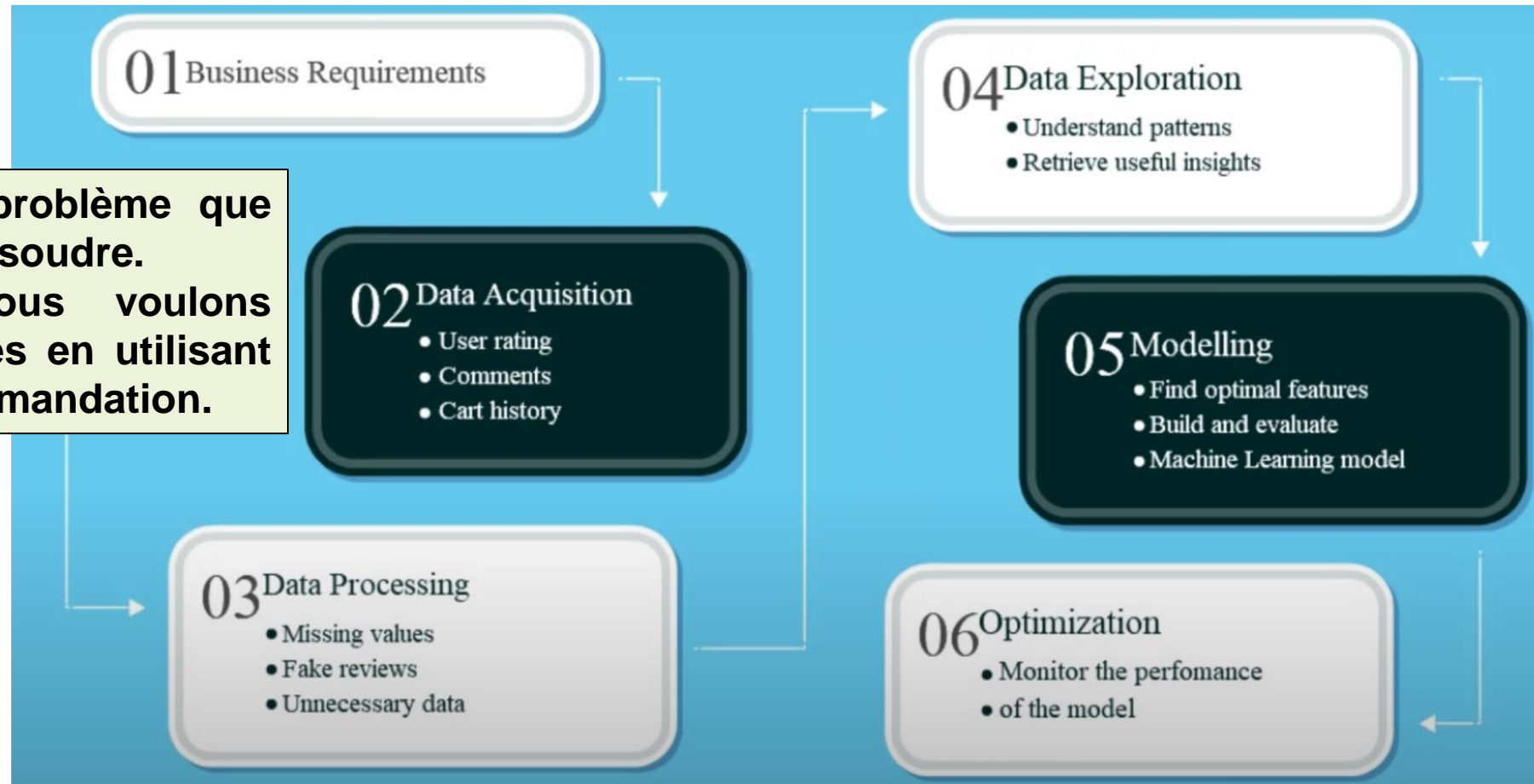
- **Le Deep learning** est un sous domaine du ML qui utilise des algorithmes conçus pour **fonctionner de façon similaire au réseau de neurones d'un cerveau humain**.
- Inspiré par le cerveau humain, qui se compose de milliards de neurones.
- Méthode la plus proche de l'intelligence humaine



# Machine learning VS Data science

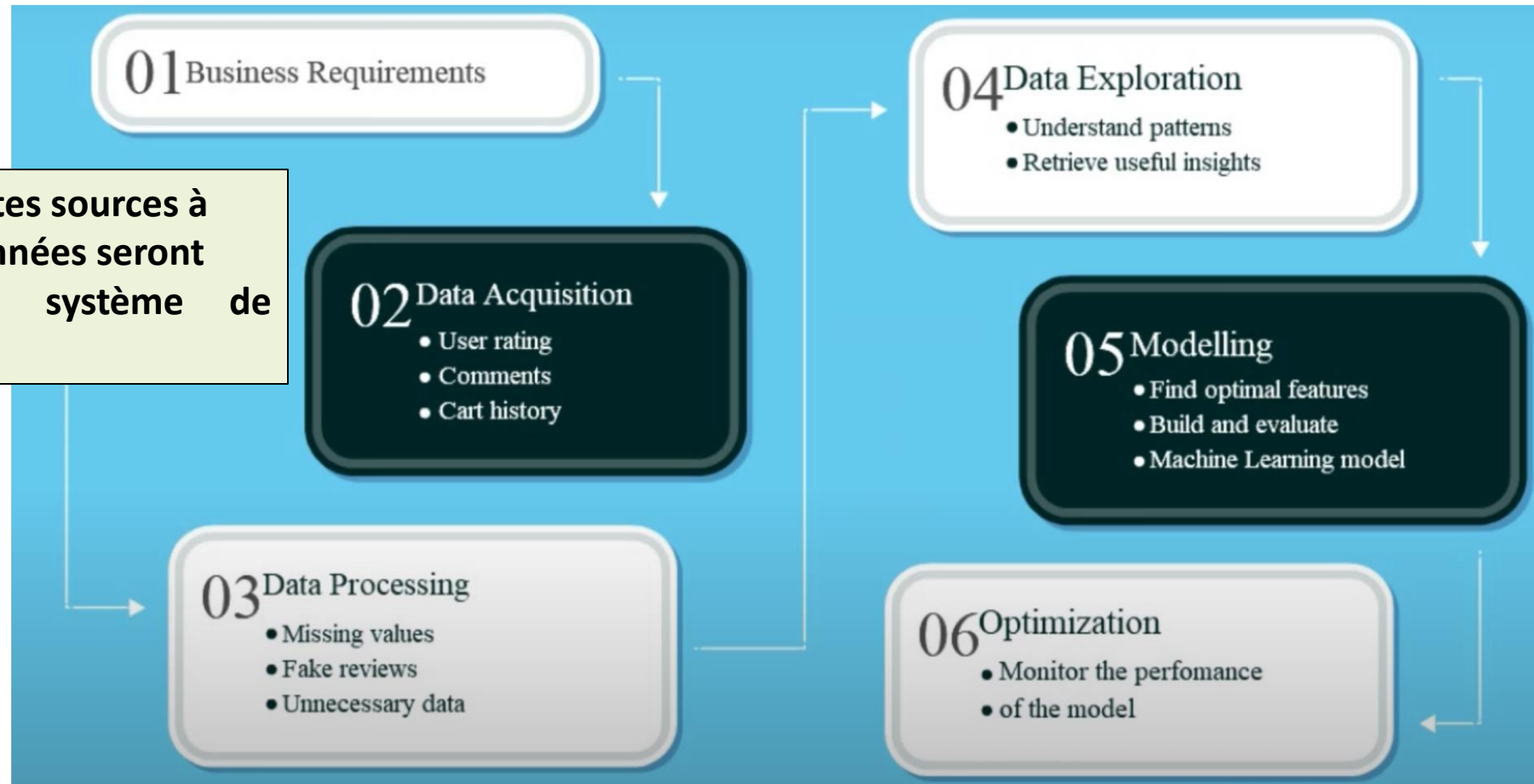
## Comment le Machine Learning est t-il utilisé en Data science ?

1- Comprendre le problème que nous essayons de résoudre.  
Dans ce cas, nous voulons augmenter les ventes en utilisant le système de recommandation.



# Machine learning VS Data science

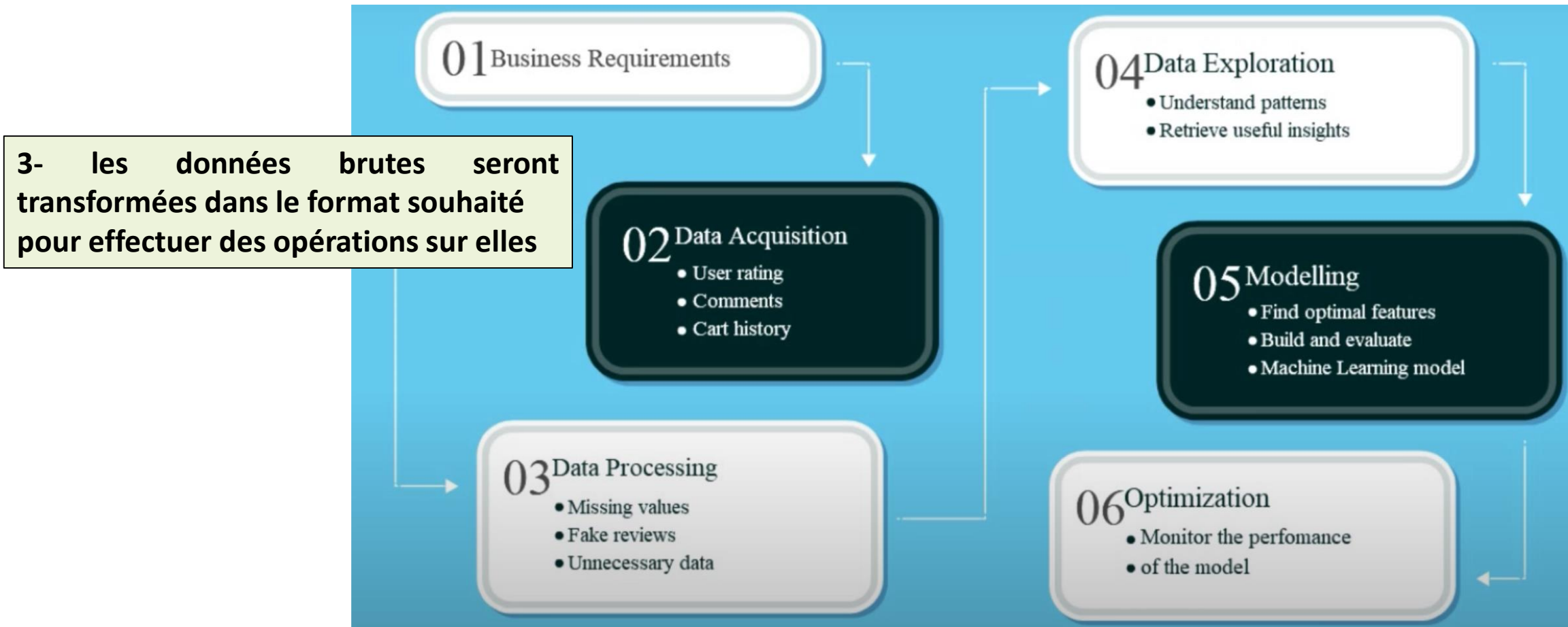
## Comment le Machine Learning est t-il utilisé en Data science ?



**2- Identifier les différentes sources à partir desquelles les données seront acquises pour notre système de recommandation.**

# Machine learning VS Data science

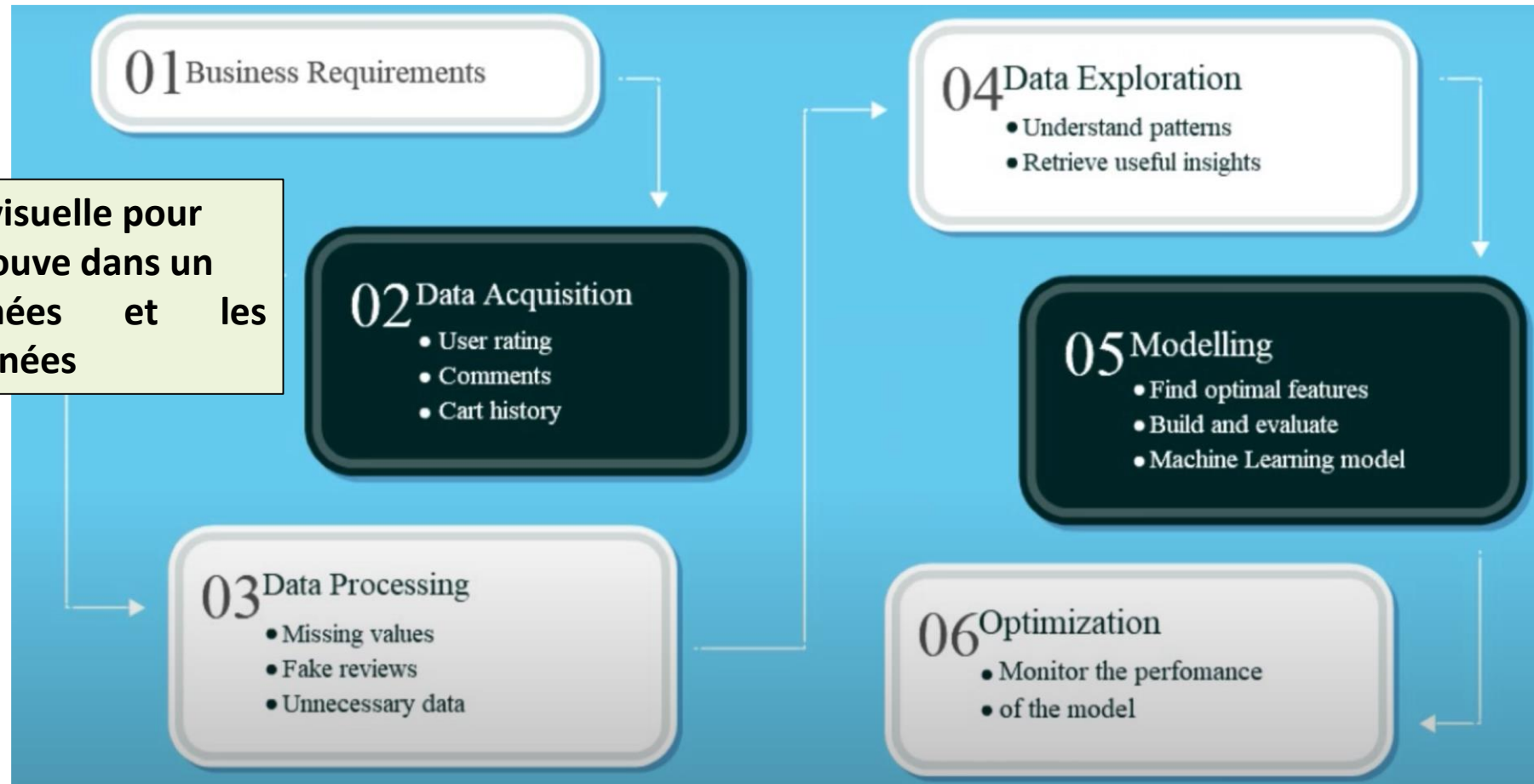
## Comment le Machine Learning est t-il utilisé en Data science ?



# Machine learning VS Data science

## Comment le Machine Learning est t-il utilisé en Data science ?

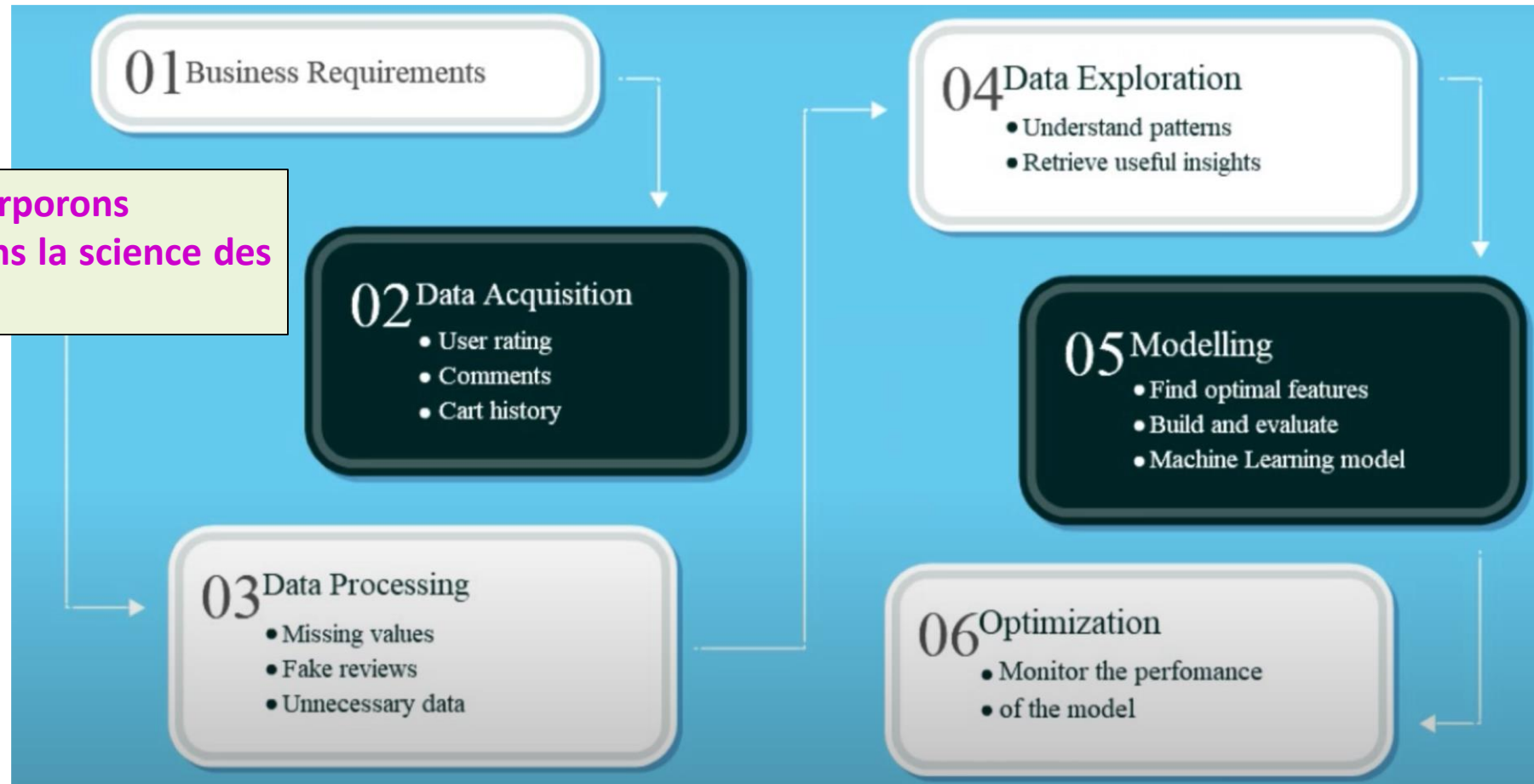
4- Utiliser l'exploration visuelle pour comprendre ce qui se trouve dans un ensemble de données et les caractéristiques des données





# Machine learning VS Data science

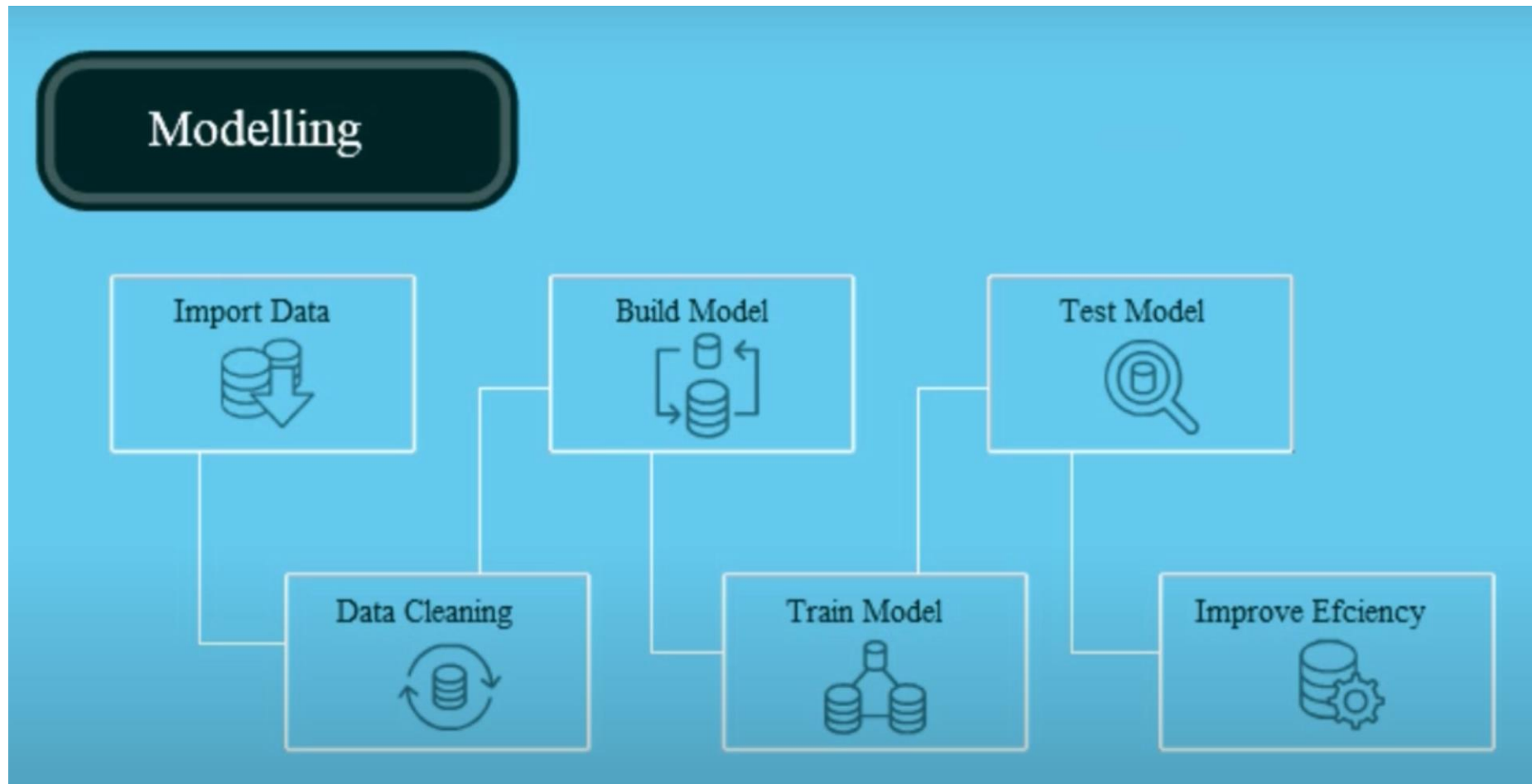
## Comment le Machine Learning est t-il utilisé en Data science ?



5- c'est là que nous incorporons  
Le machine learning dans la science des  
données

# Machine learning VS Data science

Comment le Machine Learning est t-il utilisé en Data science ?

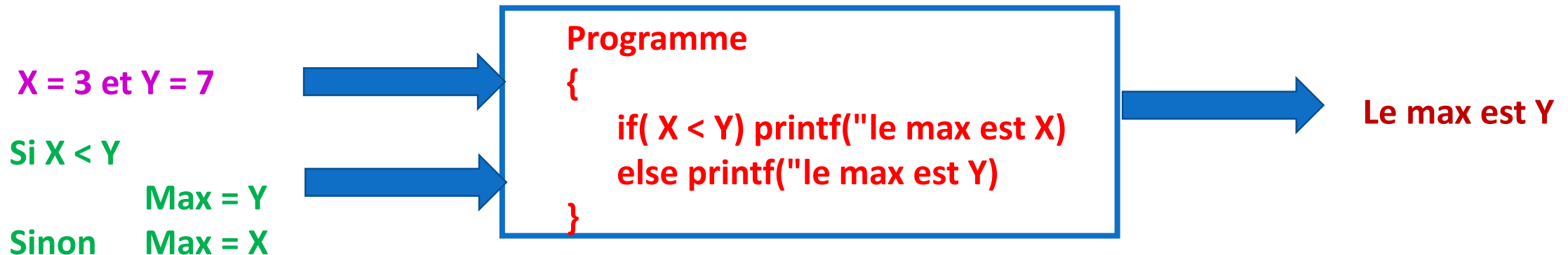


# Programmation traditionnelle vs ML

Qu'est-ce qui fait la différence entre la programmation classique et le Machine Learning (ML) ?



❑ **Exemple** : afficher la valeur maximale



# Programmation traditionnelle vs ML

Qu'est-ce qui fait la différence entre la programmation classique et le Machine Learning (ML) ?

Data

Règles

PROGRAMMATION CLASSIQUE

Réponses



Programme

{

$F(\text{photo}) = \text{????}$

}

Règles ???

"Chat"

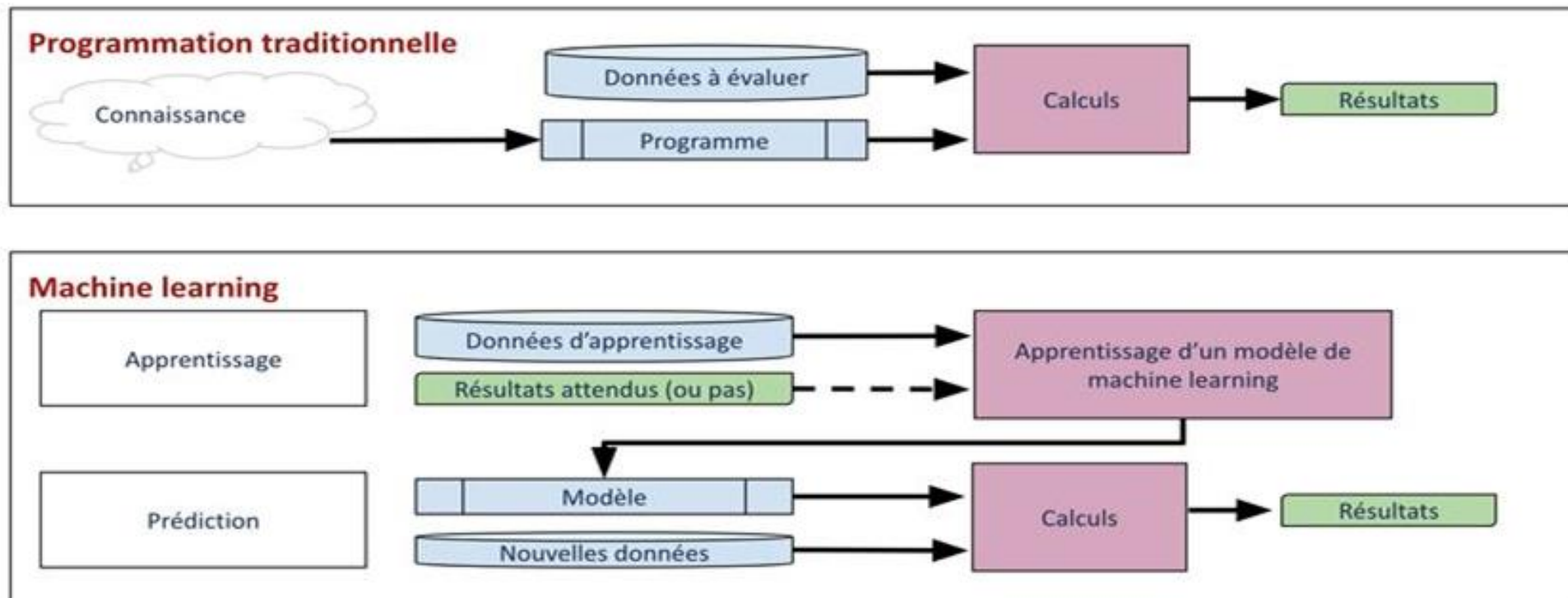
# Programmation traditionnelle vs ML

Qu'est-ce qui fait la différence entre la programmation classique et le Machine Learning (ML) ?



# Programmation traditionnelle vs ML

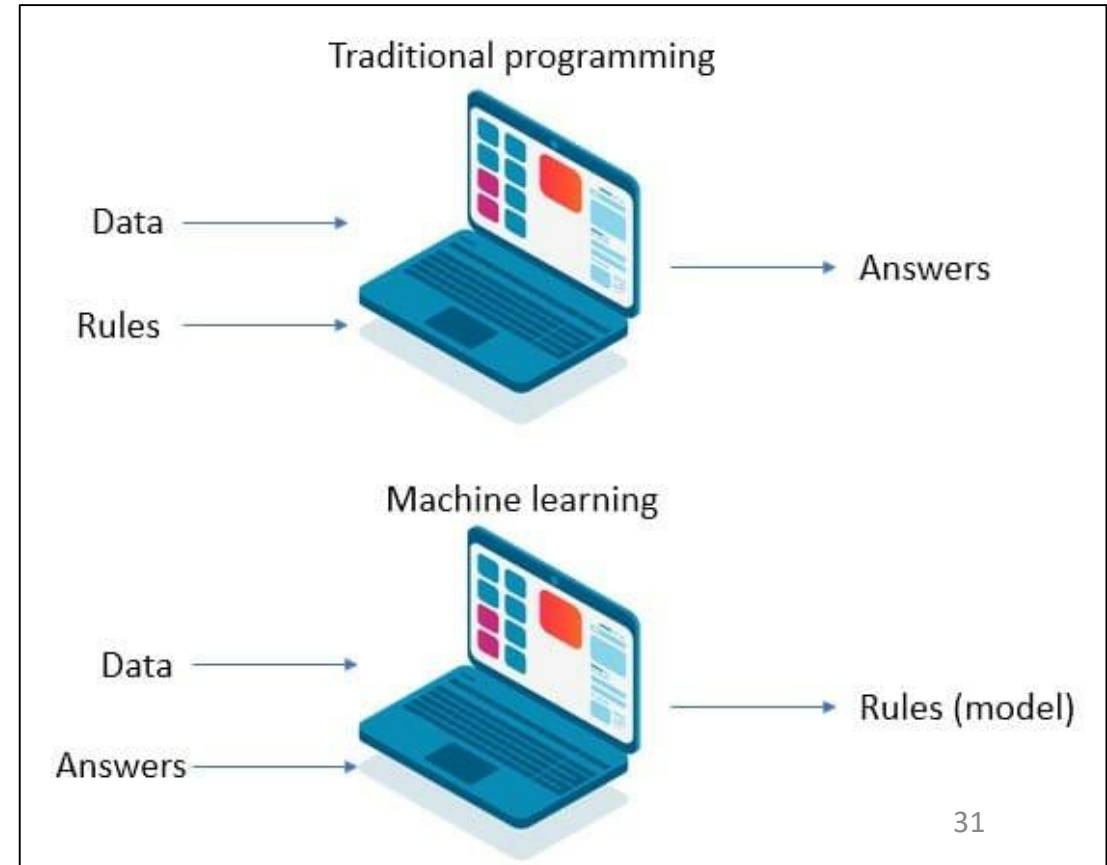
Qu'est-ce qui fait la différence entre la programmation classique et le Machine Learning (ML) ?



# Programmation traditionnelle vs ML

## Qu'est-ce qui fait la différence entre la programmation classique et le Machine Learning (ML) ?

- ✓ Dans la **programmation classique**, on fait la combinaison des inputs avec un ensemble défini de règles pour avoir le résultat.
- ✓ Dans le **Machine Learning**, on fournit les entrées et les sorties tandis que les règles sont déduites dans la phase d'entraînement (training) des modèles.



# C'est quoi le Machine Learning ?

**Machine Learning:**  
"The field of study that gives computers the ability to learn without being explicitly programmed."

## ❑ DÉFINITION 1 :

Le **Machine Learning** consiste à laisser l'ordinateur **apprendre** quel calcul effectuer, plutôt que de lui donner ce calcul (c'est-à-dire le programmer de façon explicite). [**Arthur Samuel 1959**]





# C'est quoi le Machine Learning ?

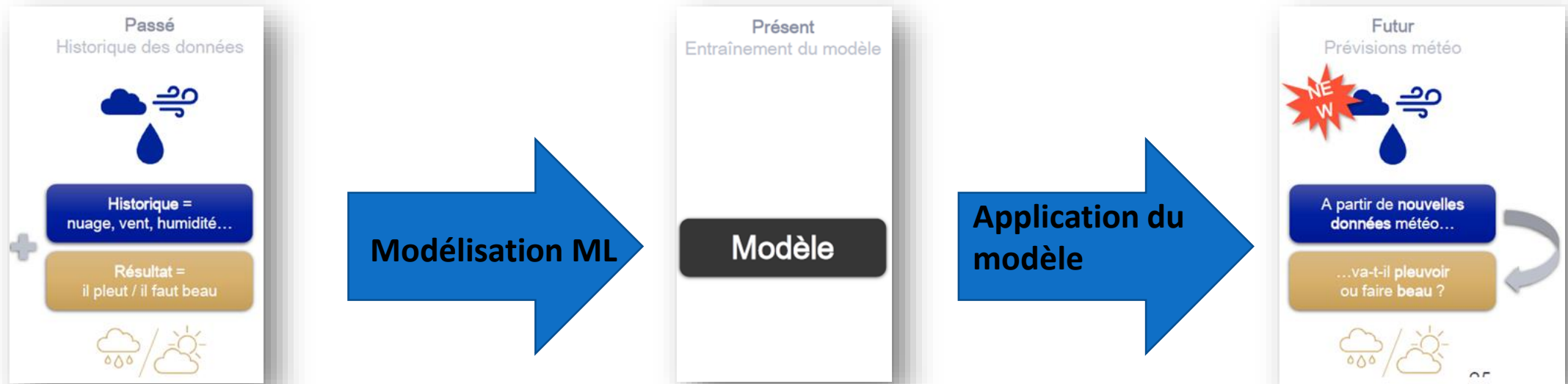
## ❑ Exemple : prédiction de la pluie

- ✓ On n'a pas de formule toute faite à donner à l'ordinateur pour qu'il nous donne une estimation si la pluie va tomber ou non.
- ✓ C'est dans ce cas précis qu'on peut faire appel au Machine Learning.
- ✓ On connaît les circonstances qui accompagnent généralement la pluie : la présence de nuages, la force du vent, le taux d'humidité...
- ✓ Le modèle s'entraîne sur les données. À la fin de son apprentissage, si on lui présente de nouvelles données météo, il sera capable de prédire s'il va pleuvoir ou pas.



# C'est quoi le Machine Learning ?

## ❑ Exemple : prédiction de la pluie

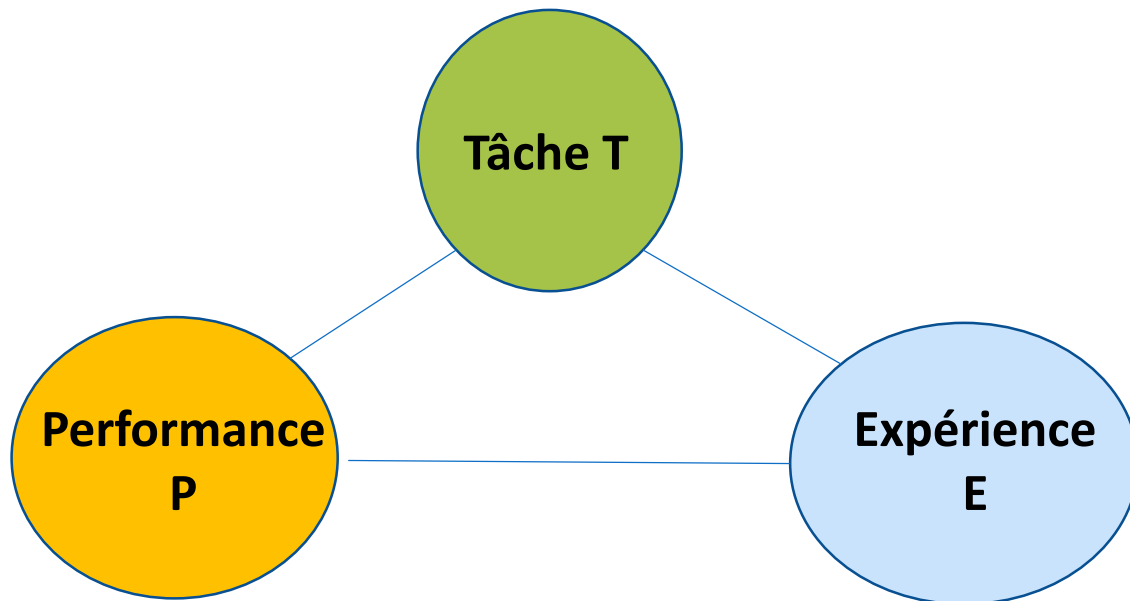


# C'est quoi le Machine Learning ?

❑ DÉFINITION 2 : Qu'est-ce qu'apprendre (learning) ?

“On dit qu'un programme informatique apprend de l'**expérience E** en référence à une certaine classe de **tâches T** et une mesure de **performance P**, si sa performance aux tâches de **T**, mesurée par **P**, s'améliore avec l'expérience **E**.”

---[Tom Mitchell, 1997]



- ✓ **Tâche T** : la tâche à traiter par le système
- ✓ **Performance P** : pour évaluer le système appris
- ✓ **Expérience E** : Training Data pour former le système d'apprentissage

# C'est quoi le Machine Learning ?

## ❑ Exemple 1 : Déterminer le prix de maison



- ✓ **Tâche T** : prédire le prix d'une maison en fonction des caractéristiques (ex: sa surface).
- ✓ **Performance P**: précision élevée des prix des maisons (accurate prices)
- ✓ **Expérience E**: un Data set de maisons avec leurs prix.

# C'est quoi le Machine Learning ?

## ❑ Exemple 2: Spam filtering



- ✓ **Tâche T**: identifier si un Email est un spam ou non.
- ✓ **Performance P**: % des email spam filtré et % des email non-spam qui ont été mal filtré.
- ✓ **Expérience E**: ensemble des emails "labelisées" par les utilisateurs.

# C'est quoi le Machine Learning ?

## ❑ **Exemple 3:** Reconnaissance d'objets dans une image



- ✓ **Tâche T:** dire si un objet est présent ou non dans l'image
- ✓ **Performance P:** nombre d'erreurs (ou pourcentage d'objets correctement identifiés)
- ✓ **Expérience E:** ensemble d'images "labelisées" avec informations sur les objets

# C'est quoi le Machine Learning ?

## ❑ Exemple 4: reconnaissance des caractères manuscrits



- ✓ **Tâche T:** la reconnaissance et la classification de caractères manuscrits présents dans une image.
- ✓ **Performance P:** % de caractères bien classés.
- ✓ **Expérience E:** Data set de caractères manuscrits étiquetés.



# Pourquoi / Quand utiliser du ML ?

❖ L'expertise humaine n'existe pas



❖ Les humains ne peuvent pas expliquer leur expertise



❖ Les solutions changent régulièrement



❖ Les solutions doivent s'adapter à des cas particuliers





# Pourquoi / Quand utiliser du ML ?

## ❑ Exemples :

- Parmi les problèmes suivants, lesquels se prêtent bien à être traités par le machine learning ?
  1. Déterminer l'horaire optimal pour poster un contenu sur une page web.
  2. Déterminer le chemin le plus court entre deux nœuds dans un graphe.
  3. Prédire le nombre de vélos à mettre en location à chaque station d'un système de location de vélos.
  4. Évaluer le prix qu'un tableau pourra atteindre lors d'une vente aux enchères.
  5. Débruiter un signal radio.

# Applications : Recommandation de produits

## Customers Who Bought This Item Also Bought

Page 1 of 15



**Data Science from Scratch:**  
First Principles with Python  
Joel Grus  
★★★★☆ 54  
#1 Best Seller in Data Mining  
Paperback  
\$33.99 ✓Prime



**Python for Data Analysis:**  
Data Wrangling with Pandas, NumPy, and...  
Wes McKinney  
★★★★☆ 118  
Paperback  
\$27.68 ✓Prime



**Data Science for Business:**  
What You Need to Know About Data Mining and Data-Analytic Thinking  
Foster Provost & Tom Fawcett  
★★★★☆ 135  
Paperback  
\$37.99 ✓Prime



**Reproducible Research with R and RStudio**  
Second Edition...  
Christopher Gandrud  
★★★★☆ 3  
Paperback  
\$51.97 ✓Prime



**An Introduction to Statistical Learning:**  
with Applications in R...  
Gareth James  
★★★★☆ 105  
Hardcover  
\$68.35 ✓Prime

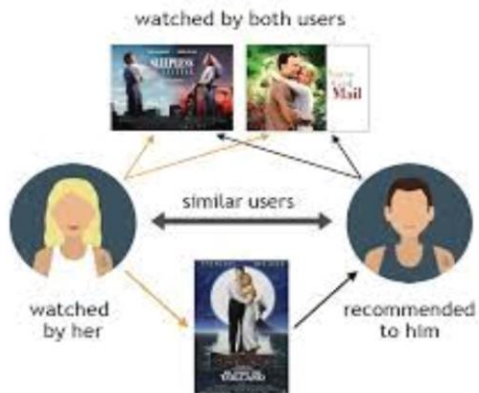


**Data Smart: Using Data Science to Transform Information into Insight**  
John W. Foreman  
★★★★☆ 99  
#1 Best Seller in Computer Simulation  
Paperback  
\$28.16 ✓Prime

















**The Statistical Sleuth: A Course in Methods of Data Analysis**  
Fred Ramsey  
★★★★☆ 6  
Hardcover  
\$284.42 ✓Prime

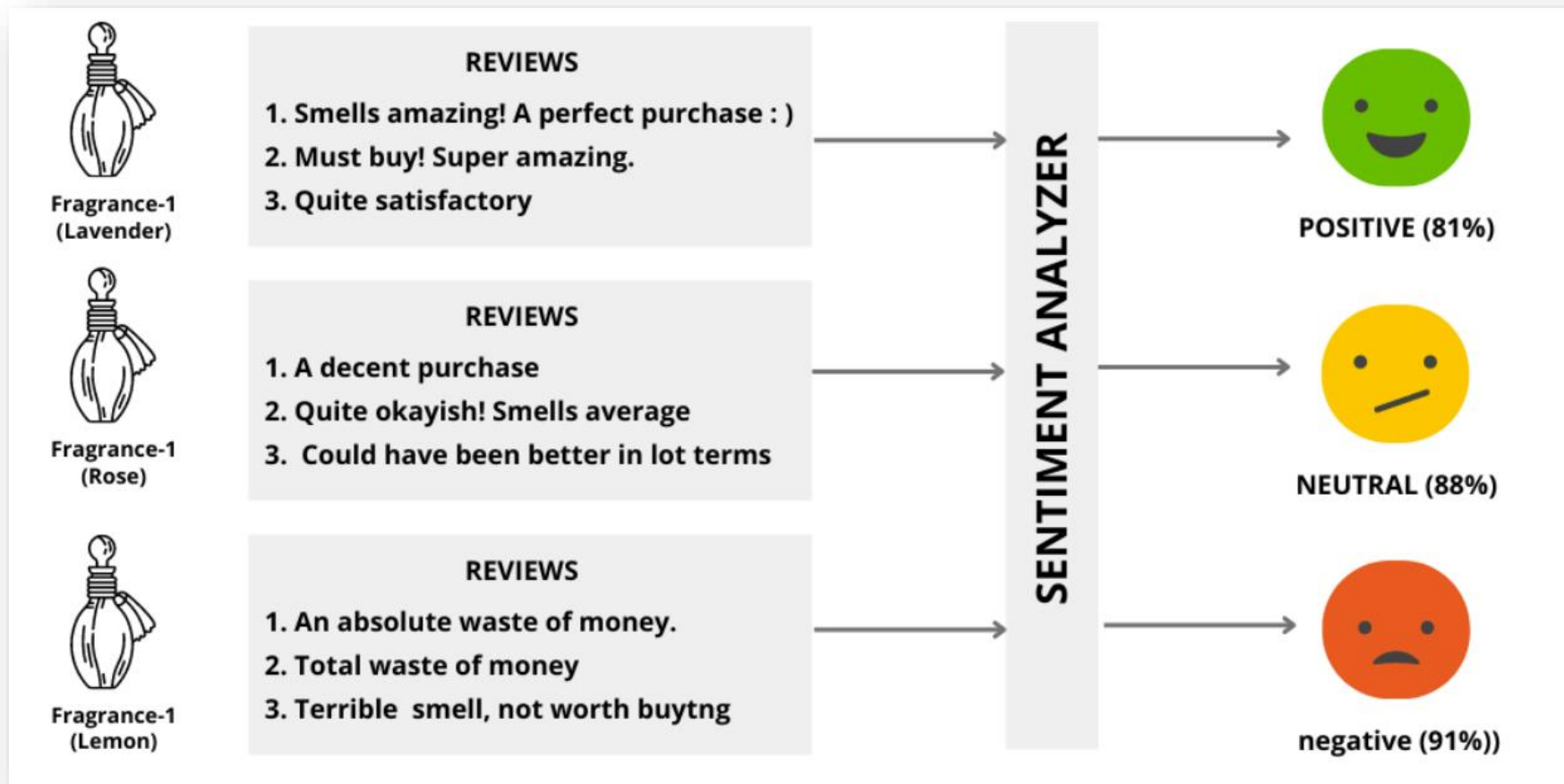
## Collaborative Filtering



Les utilisateurs similaires (en vert) n'ont pas aimé le produit que notre utilisateur n'a pas encore noté. L'algorithme aura donc tendance à prédire une mauvaise note et à ne pas recommander le produit ici.

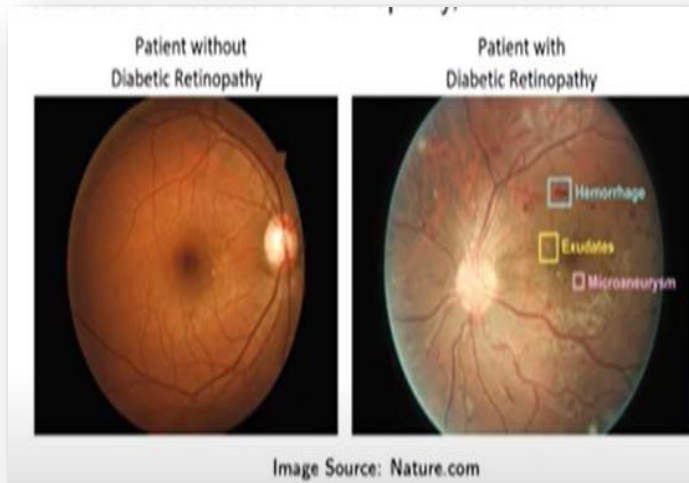
				
				
				
				
				
				

# Applications: Analyse de sentiments



Représentation schématique de l'analyse de sentiments sur les avis de trois parfums - Lavande, Rose et Citron.

# Applications : aide au diagnostic médical



**Détection de la rétinopathie  
diabétique**



**Détection de breast  
cancer**



**Détection de maladies  
cardiaques**

# Types de modèles ML

❑ 3 approches de machine learning :

**Apprentissage supervisé**



**Apprentissage non supervisé**



**Apprentissage par renforcement**



# Types de modèles ML

## ❖ Apprentissage supervisé

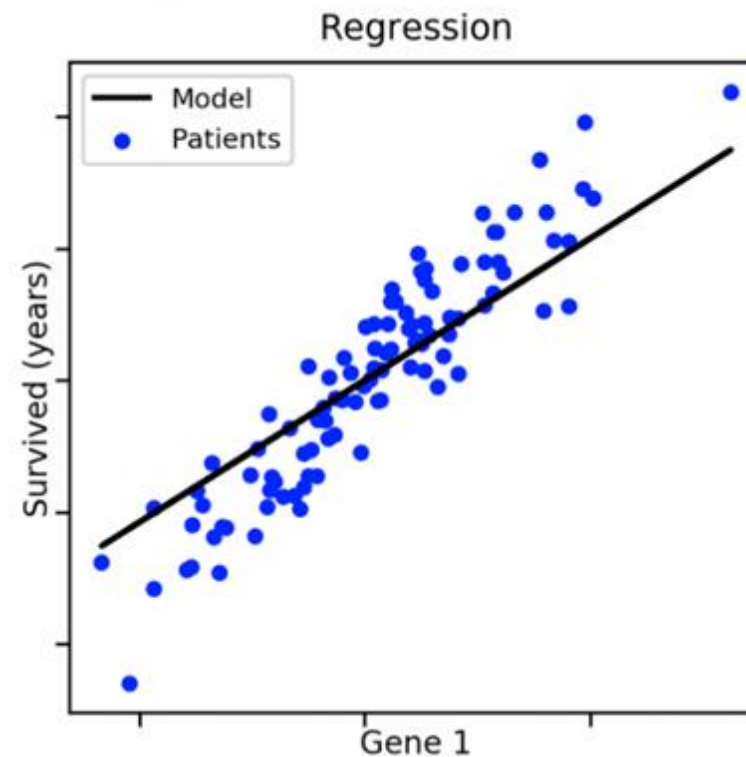
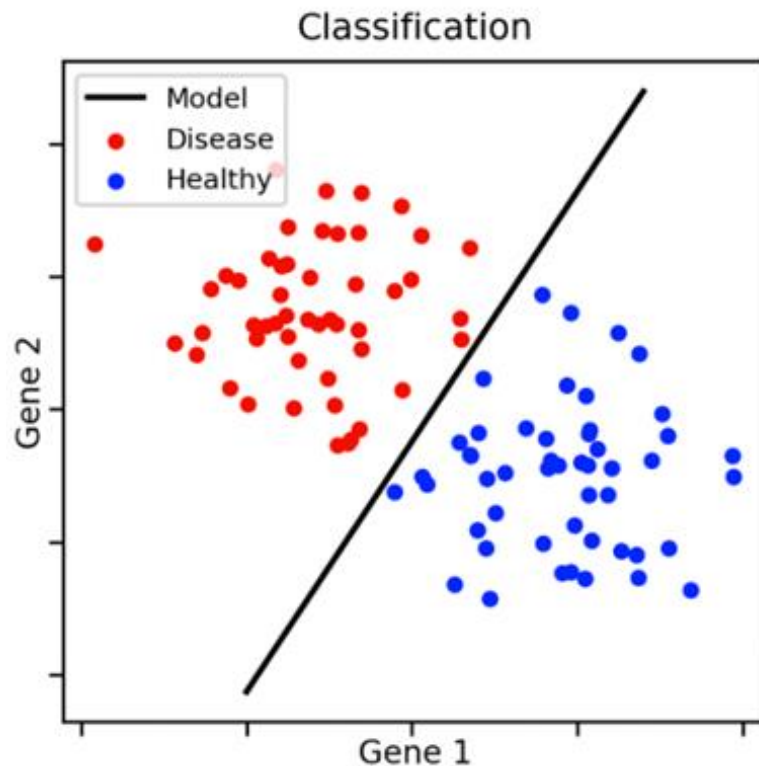
- L'approche de l'apprentissage supervisé (**supervised learning**) en ML consiste à utiliser des jeux de données étiquetés qui entraînent des algorithmes **pour classer** les données **ou prédire** des résultats avec précision.
- l'objectif de l'apprentissage supervisé est d'apprendre une fonction qui, à partir d'un échantillon de données et des résultats souhaités, se rapproche le mieux de la **relation entre entrée et sortie observable dans les données**.
- Par exemple, disons que nous formons un algorithme de reconnaissance d'images pour distinguer les chats des chiens.
- Chaque point de données d'apprentissage sera une paire constituée d'une image (données d'apprentissage) et de l'étiquette spécifiant s'il s'agit d'un chat ou d'un chien.



# Types de modèles ML

## ❖ Apprentissage supervisé

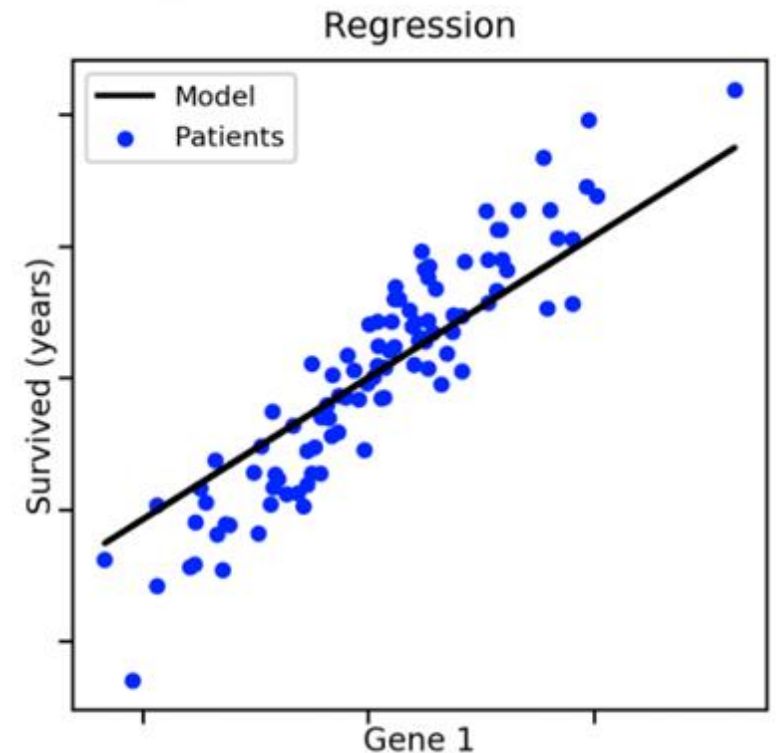
- Il existe deux grandes catégories d'apprentissage supervisé :



# Types de modèles ML

## ❖ Apprentissage supervisé

- **Les problèmes de régression** correspondent aux situations dans lesquelles la machine doit prédire la valeur d'une **variable quantitative** (variable continue).
- Un problème de régression avec plusieurs variables d'entrée est appelé un problème de **régression multiple**
- **Exemples :**
  - ✓ Prédire le prix de l'immobilier
  - ✓ Prédire la consommation électrique
  - ✓ Prédire la durée de vie d'un patient

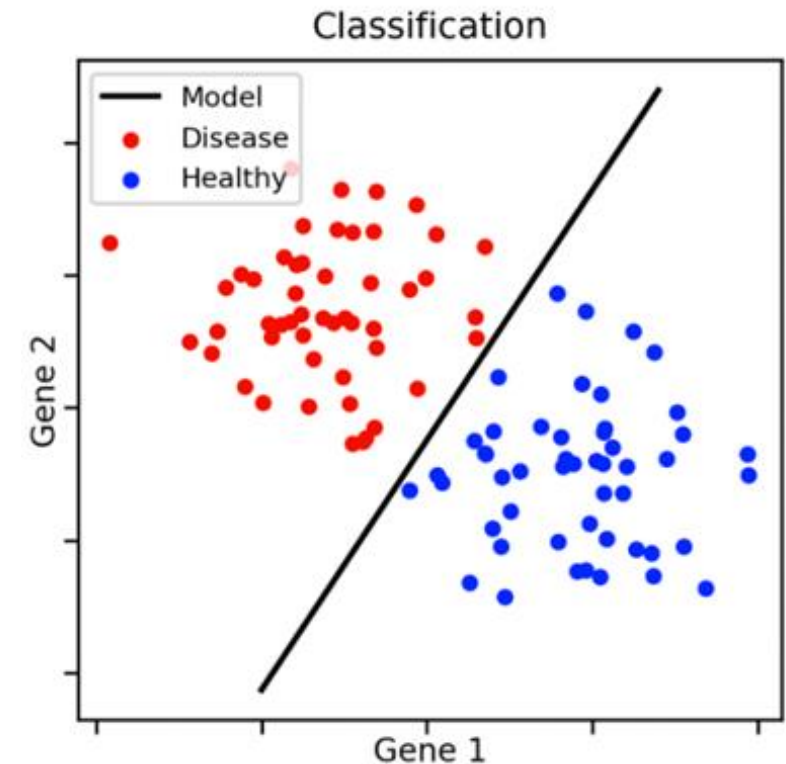




# Types de modèles ML

## ❖ Apprentissage supervisé

- **La classification** correspondent aux situations dans lesquelles la machine doit prédire la valeur d'une **variable qualitative** (variable discrète). Autrement dit, la machine doit classer ce qu'on lui donne dans des classes.
- Un problème de classification peut être **binaire** (deux classes) ou **multi-class** (plus de deux classes).
- **Exemples :**
  - ✓ Email Spam / non Spam
  - ✓ Cancer / non Cancer
  - ✓ Photo de Chat / Chien



# Types de modèles ML

## ❖ Apprentissage supervisé

### Regression



What will be the temperature tomorrow?

84°



Fahrenheit

### Classification



Will it be hot or cold tomorrow?

COLD

HOT



Fahrenheit

# Types de modèles ML

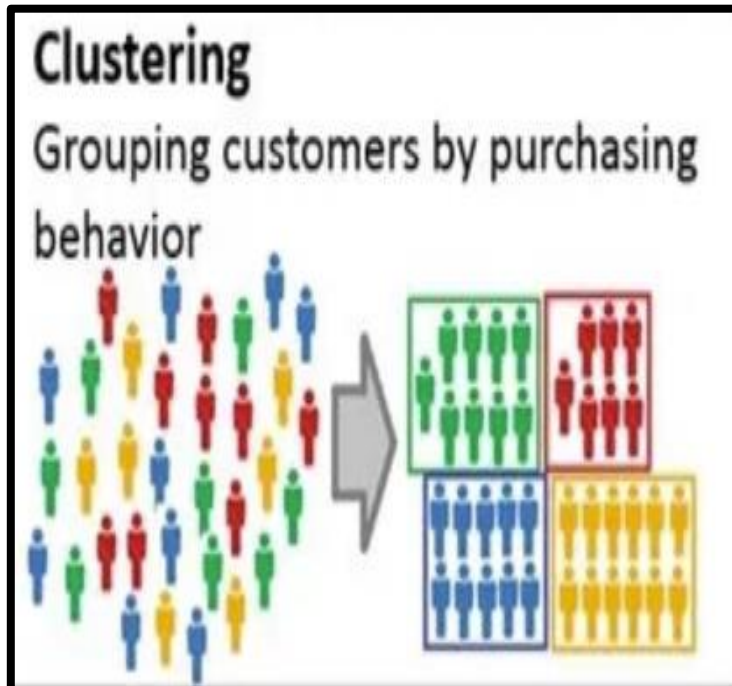
## ❖ Apprentissage non supervisé

- L'approche de l'apprentissage non supervisé (**Unsupervised Learning**) consiste à ne disposer que de données d'entrée (X) et pas de variables de sortie correspondantes.
- On l'appelle apprentissage non supervisé car, contrairement à l'apprentissage supervisé, il n'y a pas de réponse correcte ni d'enseignant. Les algorithmes sont laissés à leurs propres mécanismes pour découvrir et présenter la structure intéressante des données.
- L'objectif de l'apprentissage non supervisé est de modéliser la structure ou la distribution sous-jacente dans les données afin d'en apprendre davantage sur les données.
- Par exemple, un ensemble de données qui contient les préférences d'achat des utilisateurs du e-commerce. Les utilisateurs formeront probablement des groupes qui se comportent de manière similaire, et ils peuvent avoir un comportement d'achat similaire.

# Types de modèles ML

## ❖ Apprentissage non supervisé

- Il existe deux grandes catégories d'apprentissage non supervisé :



# Types de modèles ML

## ❖ Apprentissage non supervisé

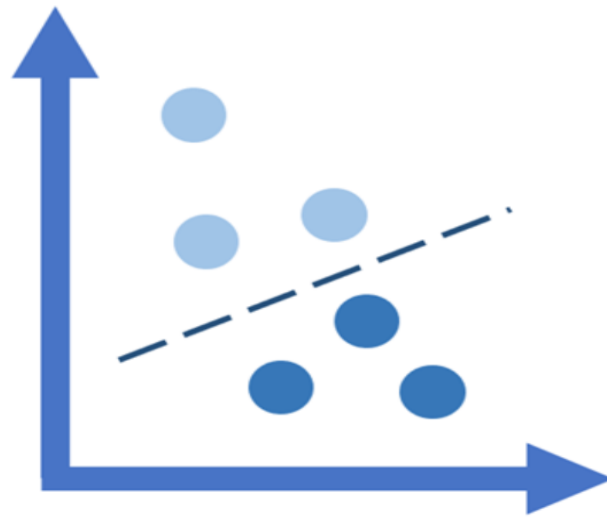
- **Le Clustering ou Regroupement** consiste à séparer ou à diviser un ensemble de données en un certain nombre de groupes, de sorte que les ensembles de données appartenant aux mêmes groupes se ressemblent davantage que ceux d'autres groupes.
- Cette technique permet aux spécialistes du marketing digital d'adapter leurs contenus en fonction des audiences identifiées.
- Le Clustering est **de la classification non supervisée**: pas de classes prédéfinies
- **Exemples :**
  - ✓ Segmentation du marché en fonction de l'emplacement, des niveaux de revenus, de l'âge des acheteurs ou de n'importe quelle autre variable.



# Types de modèles ML

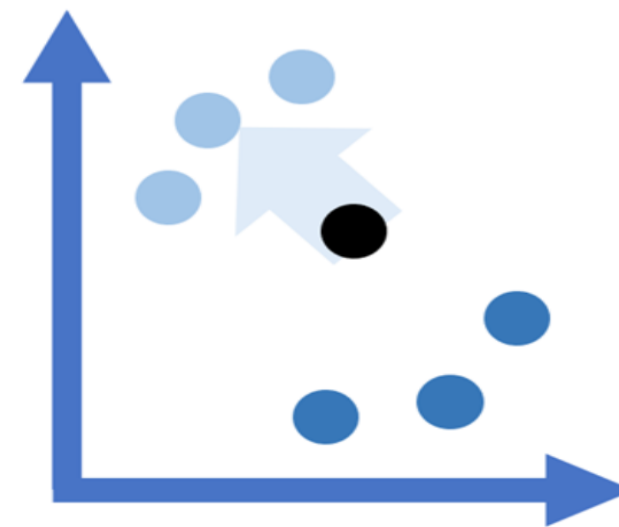
## ❖ Apprentissage non supervisé

### ■ Clustering Vs Classification



Regroupements  
*clustering*

Le regroupement (ou *clustering*) permet d'identifier des groupes d'individus au sein d'un jeu de données.



Classifications  
*catégorisations*

La classification ou catégorisation permet d'attribuer des individus à des groupes déjà définis.

# Types de modèles ML

## ❖ Apprentissage non supervisé

- **L'association** consiste à découvrir des relations intéressantes entre des variables dans de grandes bases de données.
- Ces méthodes sont souvent utilisées pour l'analyse du panier de consommation, qui permet aux compagnies de comprendre la relation entre l'achat de différents produits. Cela permet d'établir une relation de la forme : « **Les individus qui achètent le produit X ont également tendance à acheter le produit Y** ».
- **Exemples :**
  - ✓ les personnes qui achètent une nouvelle maison ont aussi tendance à acheter de nouveaux meubles
  - ✓ Si 15 clients ayant acheté un nouveau téléphone ont également commandé un casque, les algorithmes recommandent un casque à tous les clients qui mettent un téléphone dans leur panier





# Types de modèles ML

## ❖ Apprentissage par renforcement

- L'apprentissage par renforcement (**Reinforcement learning**) consiste à laisser les ordinateurs apprendre de **leurs expériences** grâce à un système de récompense ou de pénalité.
- Son objectif est de créer des **agents** capables de prendre les meilleures décisions possibles.
- Le RL se base donc sur un **système de récompenses** : à chaque **action** effectuée, **l'agent** reçoit une **récompense** (ou une pénalité dans le cas où l'action est mal réalisée) de la part de **l'environnement**, qui va lui permettre **d'ajuster sa stratégie**.
- L'agent apprend donc à agir de façon à maximiser les récompenses espérées sur le long terme.
- **Exemples :**
  - ✓ apprendre à un robot à marcher en terrain difficile
  - ✓ La voiture autonome



# Types de modèles ML

## ❖ Apprentissage par renforcement

### ▪ Exemple : Concevoir une voiture autonome

- ✓ l'**agent** est le logiciel qui conduit la voiture,
- ✓ l'**environnement** est l'espace virtuel dans lequel le véhicule évolue,
- ✓ l'**action** à un instant  $t$ , c'est la possibilité de changer de direction, de ralentir, d'accélérer etc...,
- ✓ la **récompense** est associée dans le cas où le véhicule circule correctement dans son environnement (à défaut, une **pénalité** est associée si le véhicule se heurte à un obstacle, par exemple),
- ✓ On parle de **politique** pour désigner le mécanisme qui lie les récompenses observées dans le passé et l'état du système à l'instant  $t$

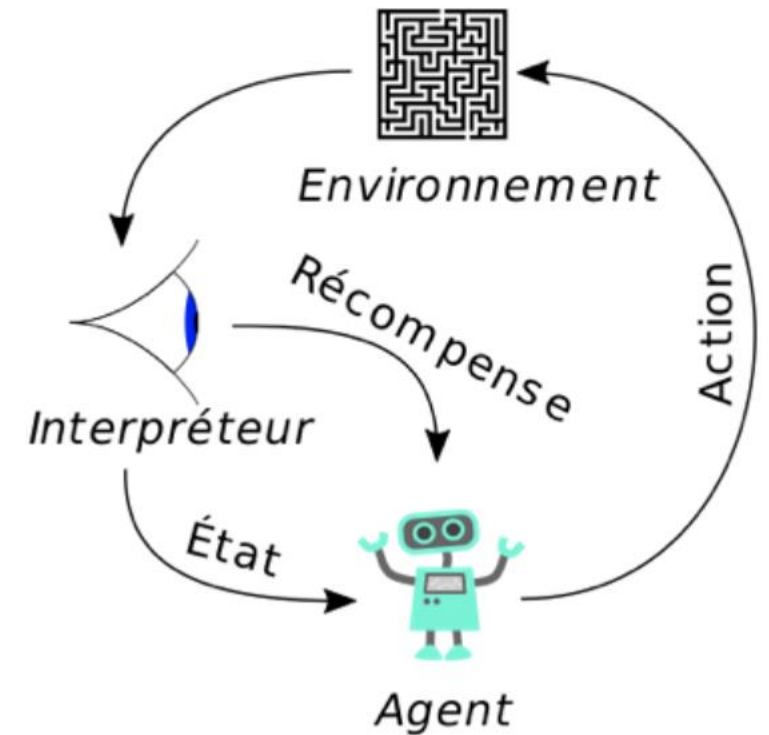


Fig. 1: Scénario typique de RL (Source : Wikipédia)

# Types de modèles ML : Résumé

## Définition

*Supervised learning is a method in which we teach the machine using labelled data*



*In unsupervised learning the machine is trained on unlabelled data without any guidance*

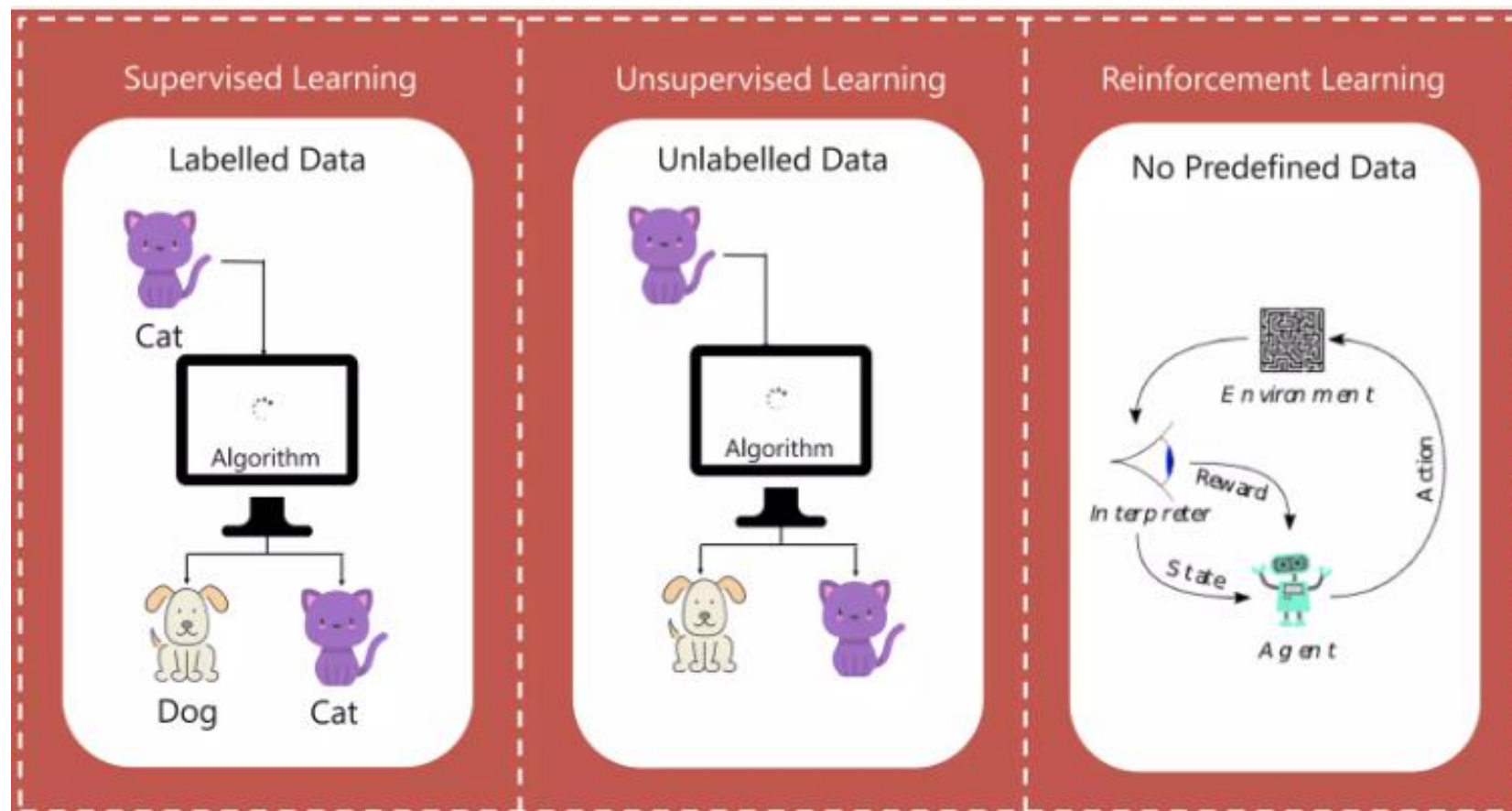


*In Reinforcement learning an agent interacts with its environment by producing actions & discovers errors or rewards*



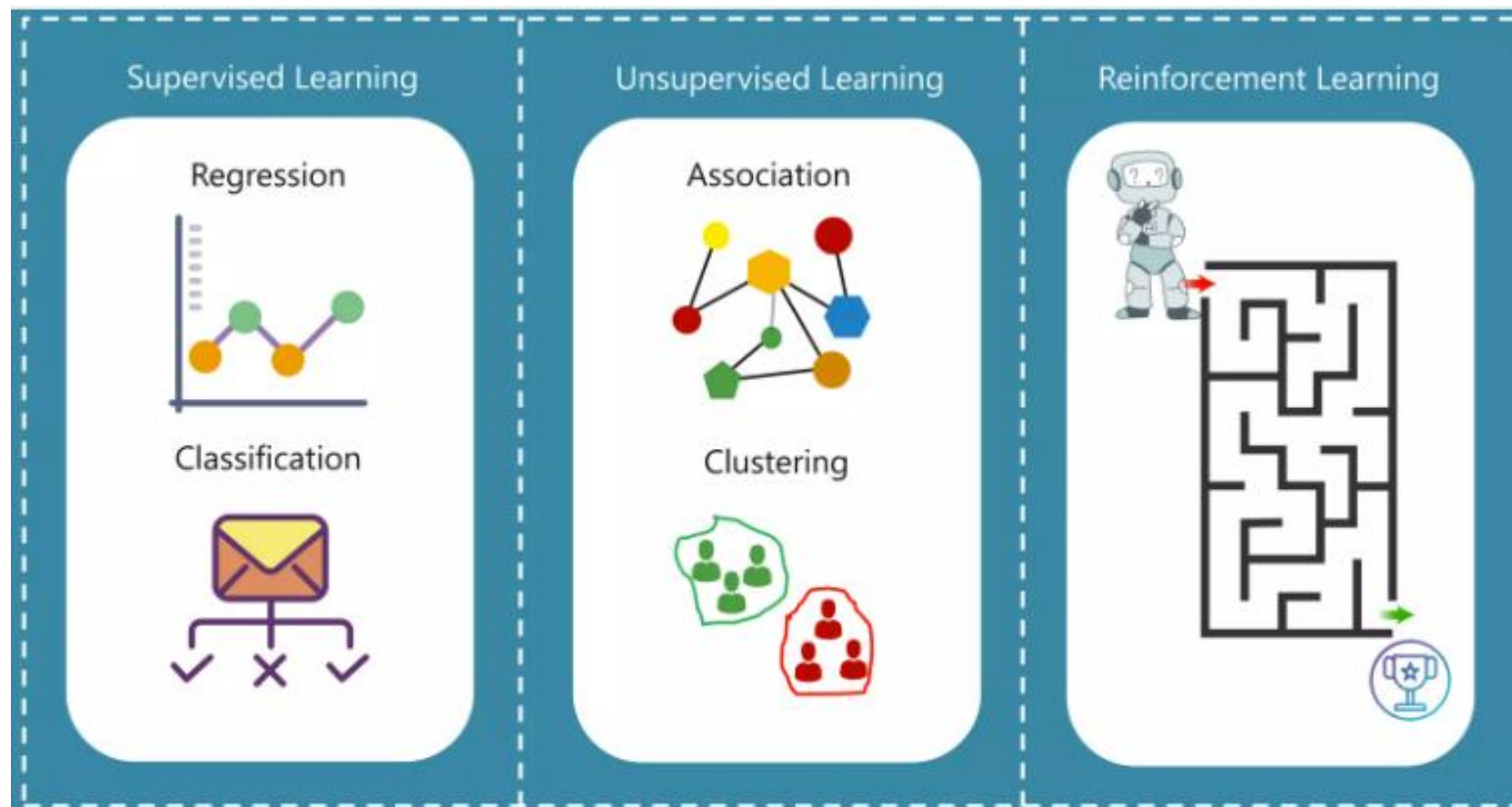
# Types de modèles ML : Résumé

Type de données



# Types de modèles ML : Résumé

## Type de problèmes



# Types de modèles ML : Résumé

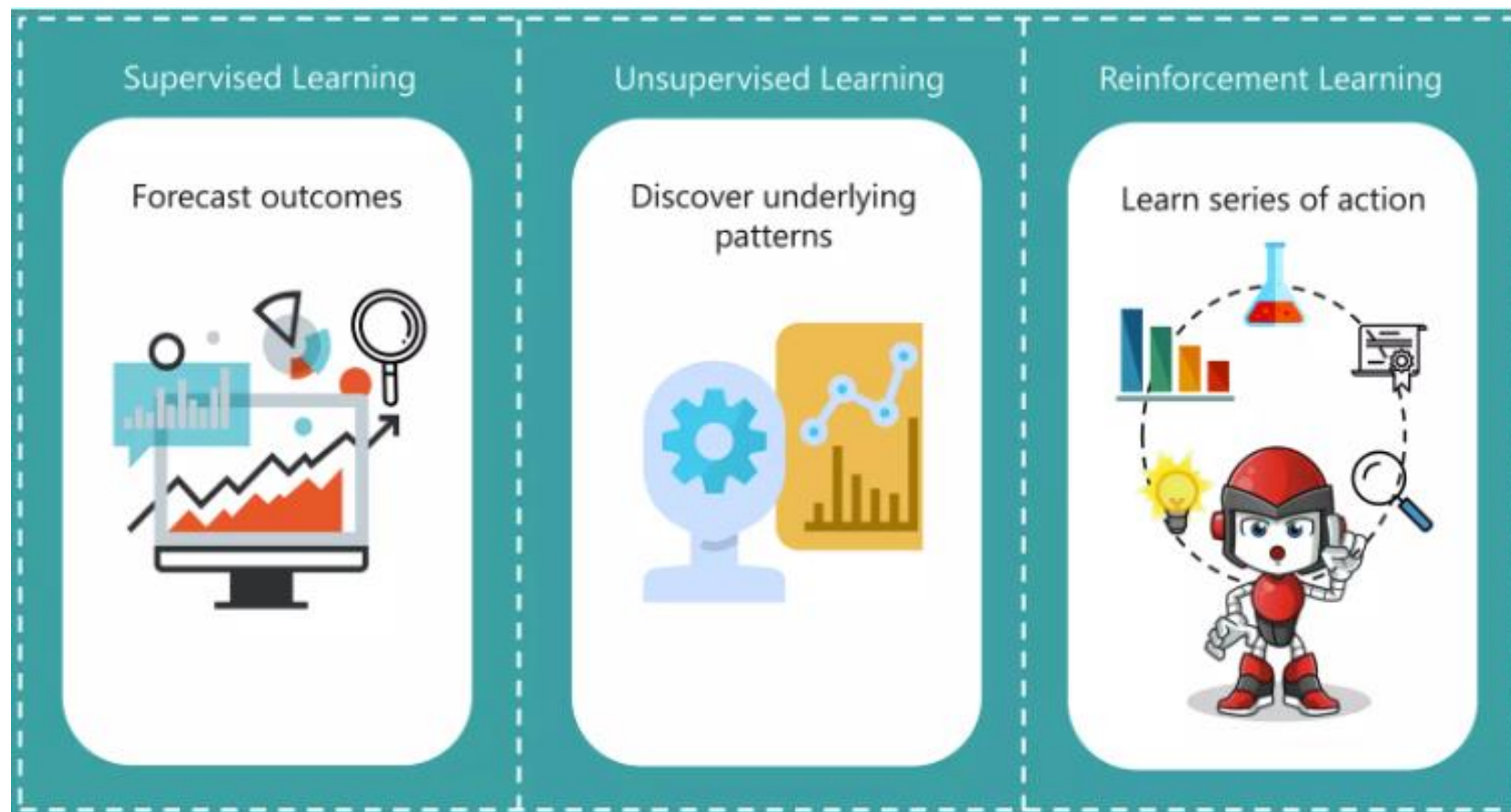
**Supervision**





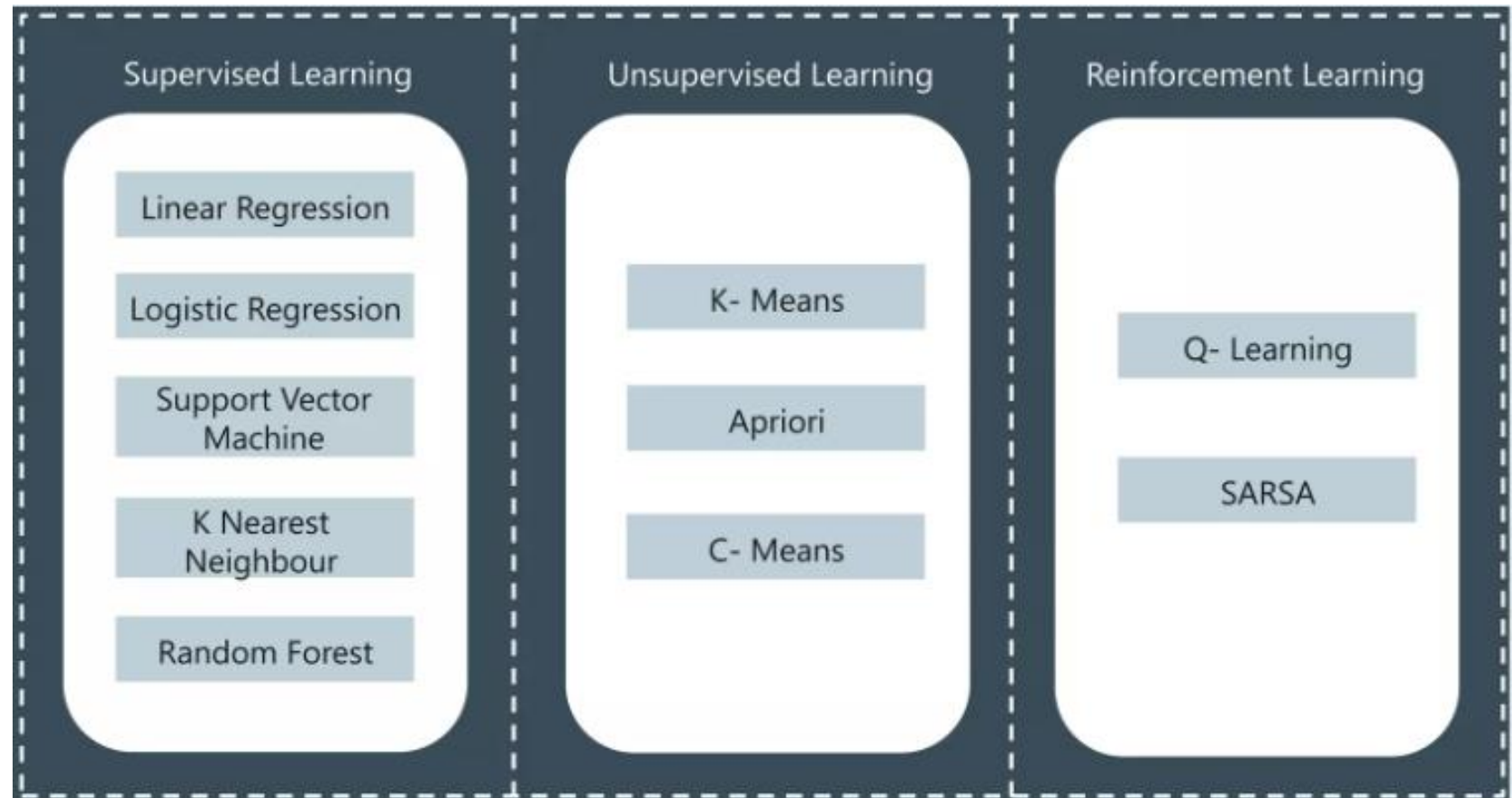
# Types de modèles ML : Résumé

## Objectif



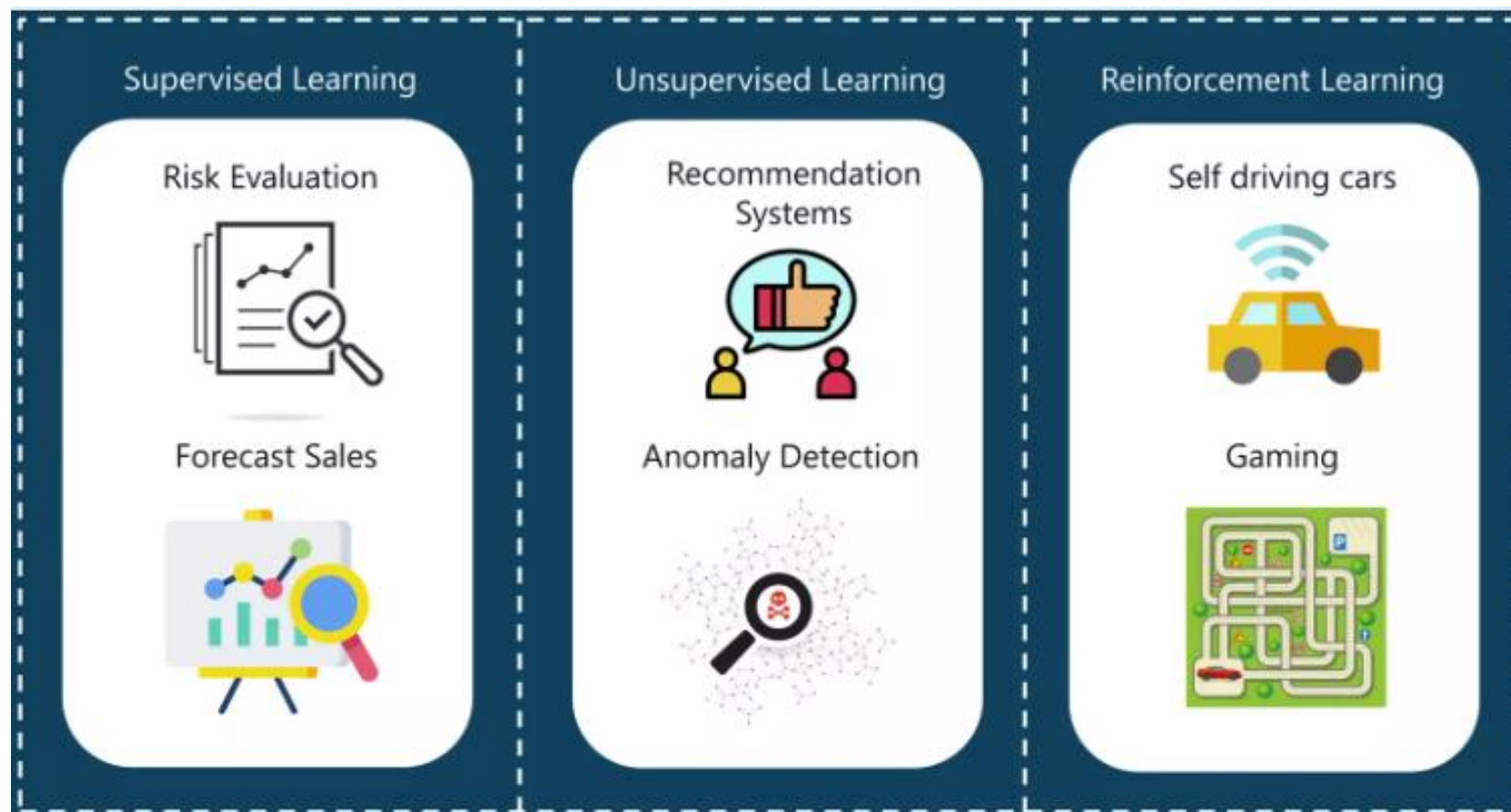
# Types de modèles ML : Résumé

## Algorithmes



# Types de modèles ML : Résumé

## Applications





# Types de modèles ML : Etude de cas

- **Énoncé du problème :** Étudier un ensemble de données sur le crédit bancaire et prendre la décision d'approuver ou non le prêt d'un demandeur en fonction de son profil.

\$ Account.Balance	: int	1 1 2 1 1 1 1 1 4 2 ...
\$ Duration.of.Credit..month.	: int	18 9 12 12 12 10 8 6 18 24 ..
\$ Payment.Status.of.Previous.Credit	: int	4 4 2 4 4 4 4 4 4 2 ...
\$ Purpose	: int	2 0 9 0 0 0 0 0 3 3 ...
\$ Credit.Amount	: int	1049 2799 841 2122 2171 2241
\$ Value.Savings.Stocks	: int	1 1 2 1 1 1 1 1 1 3 ...
\$ Length.of.current.employment	: int	2 3 4 3 3 2 4 2 1 1 ...
\$ Instalment.per.cent	: int	4 2 2 3 4 1 1 2 4 1 ...
\$ Sex...Marital.Status	: int	2 3 2 3 3 3 3 3 2 2 ...
\$ Guarantors	: int	1 1 1 1 1 1 1 1 1 1 ...
\$ Duration.in.Current.address	: int	4 2 4 2 4 3 4 4 4 4 ...
\$ Most.valuable.available.asset	: int	2 1 1 1 2 1 1 1 3 4 ...
\$ Age..years.	: int	21 36 23 39 38 48 39 40 65 23
\$ Concurrent.Credits	: int	3 3 3 3 1 3 3 3 3 3 ...
\$ Type.of.apartment	: int	1 1 1 1 2 1 2 2 2 1 ...
\$ No.of.Credits.at.this.Bank	: int	1 2 1 2 2 2 2 1 2 1 ...
\$ Occupation	: int	3 3 2 2 2 2 2 2 1 1 ...
\$ No.of.dependents	: int	1 2 1 2 1 2 1 2 1 1 ...
\$ Telephone	: int	1 1 1 1 1 1 1 1 1 1 ...
\$ Foreign.Worker	: int	1 1 1 2 2 2 2 2 1 1 ...

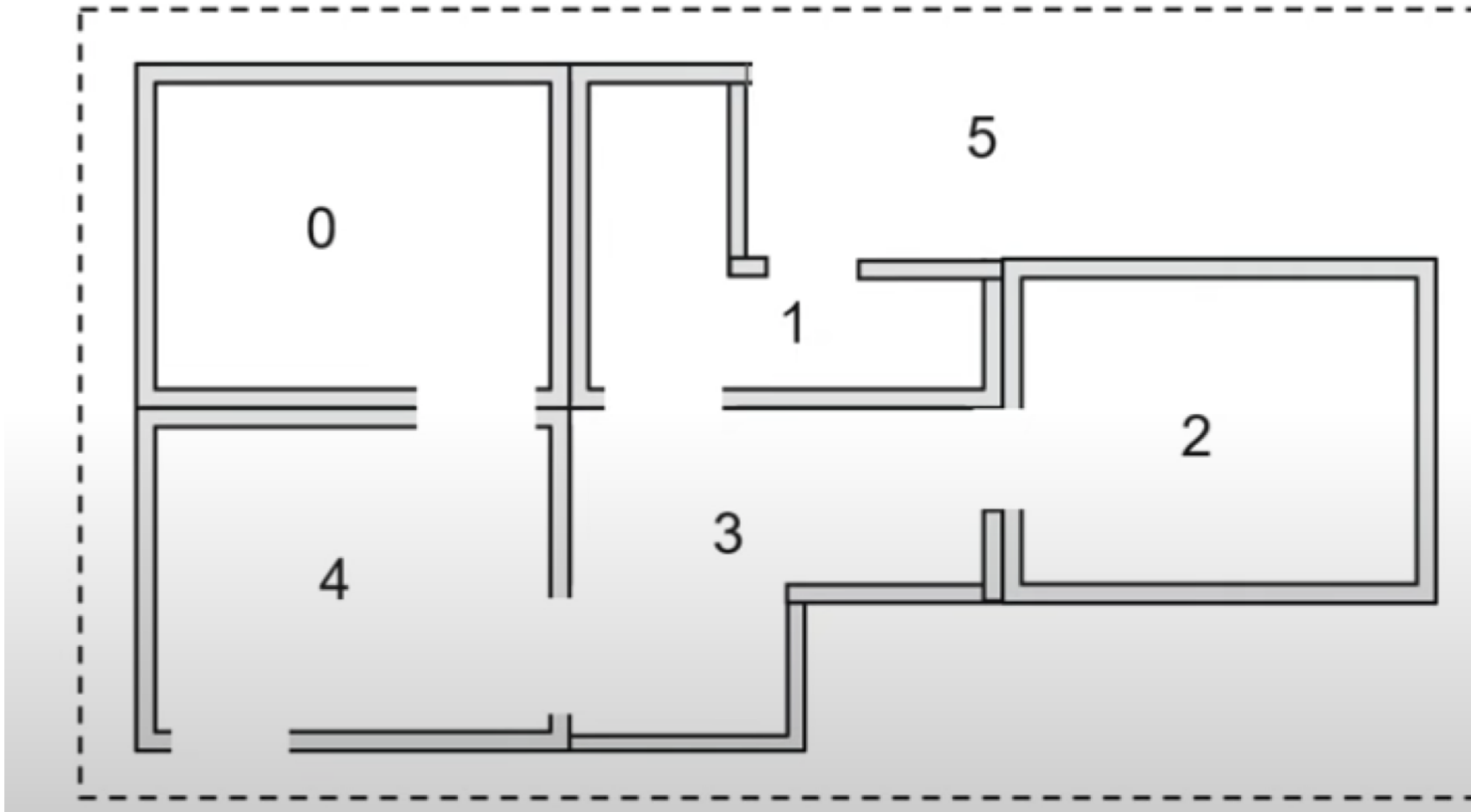
# Types de modèles ML : Etude de cas

- **Énoncé du problème :** Effectuer une analyse du panier en trouvant la relation entre les articles achetés à l'épicerie.

	A	B	C	D	E	F	G	H
1	citrus fruit	semi-finish	margarine	ready soups				
2	tropical fr	yogurt	coffee					
3	whole milk							
4	pip fruit	yogurt	cream che	meat spreads				
5	other vege	whole mill	condense	long life bakery product				
6	whole mill	butter	yogurt	rice	abrasive cleaner			
7	rolls/buns							
8	other vege	UHT-milk	rolls/buns	bottled be	liquor (appetizer)			
9	pot plants							
10	whole mill	cereals						
11	tropical fr	other vege	white bre	bottled w	chocolate			
12	citrus fruit	tropical fr	whole mill	butter	curd	yogurt	flour	bottled w
13	beef							
14	frankfurte	rolls/buns	soda					
15	chicken	tropical fruit						

# Types de modèles ML : Etude de cas

- **Énoncé du problème :** Placez un agent dans l'une des pièces (0,1,2,3,4) et le but est d'atteindre l'extérieur du bâtiment (pièce 5).

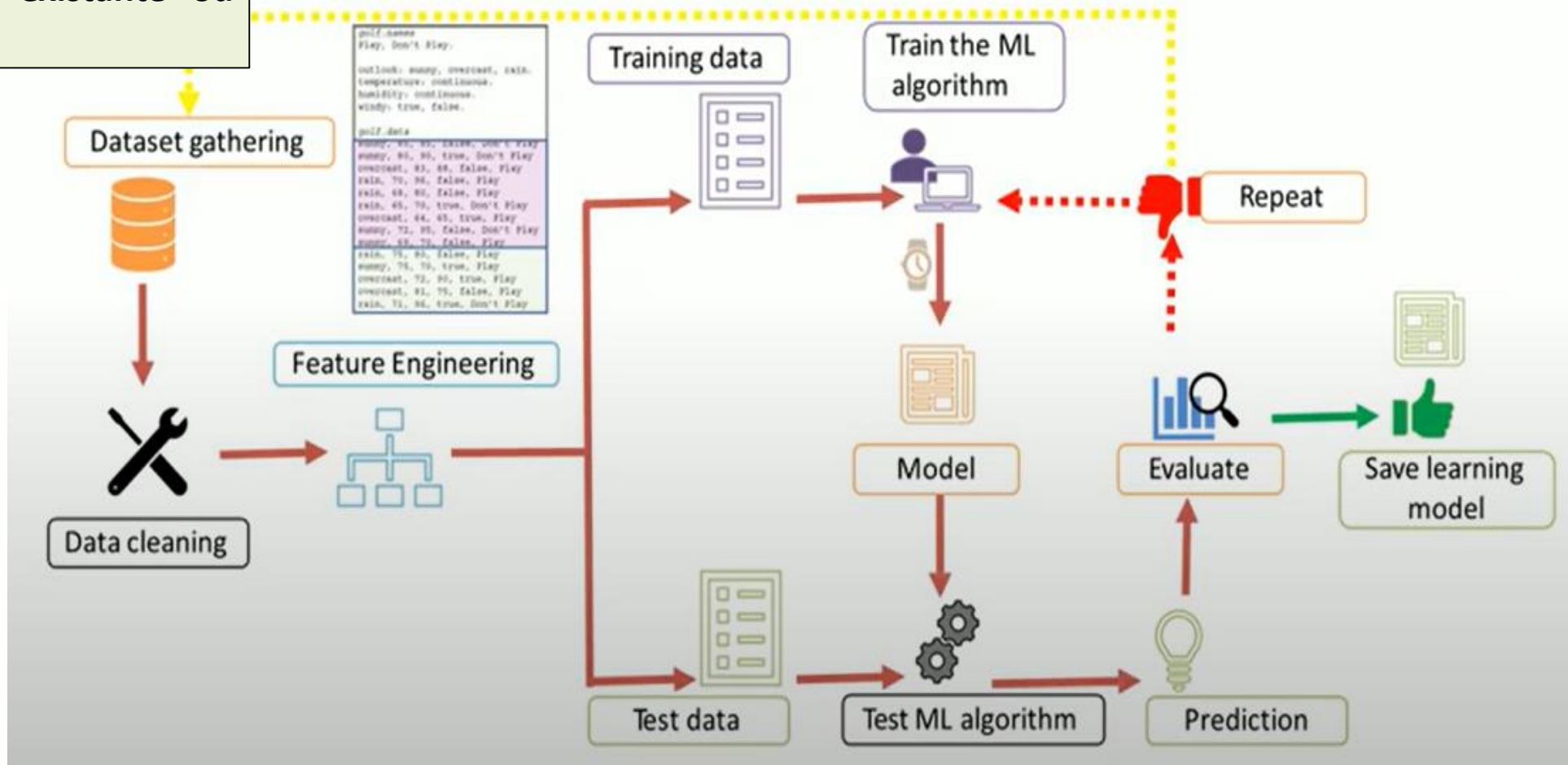


# Types de modèles ML : Exercice

- 1) Je veux écrire un programme qui utilise la fréquence des mots « science », « public », « accès », « université », « gouvernement », « financer », « éducation », « budget », « justice » et « loi » pour déterminer si un article traite ou non de politique scientifique. J'ai commencé par annoter un millier d'articles selon leur sujet. Quel genre de problème d'apprentissage automatique dois-je résoudre ?
- 2) Je dispose de 10 000 articles de journaux que je souhaite classer par leur thématique. Dois-je utiliser un algorithme supervisé ou non supervisé ?
- 3) Je gère un outil qui permet d'organiser les liens HTML qui ont été sauvegardés. Je souhaite suggérer des catégories auxquelles affecter un nouveau lien, en fonction des catégories déjà définies par l'ensemble des utilisateurs du service. Quel type d'algorithme d'apprentissage dois-je utiliser ?
- 4) Je veux examiner mes spams pour déterminer s'il existe des sous-types de spams. Quel type d'algorithme d'apprentissage dois-je utiliser ?

# Machine learning process

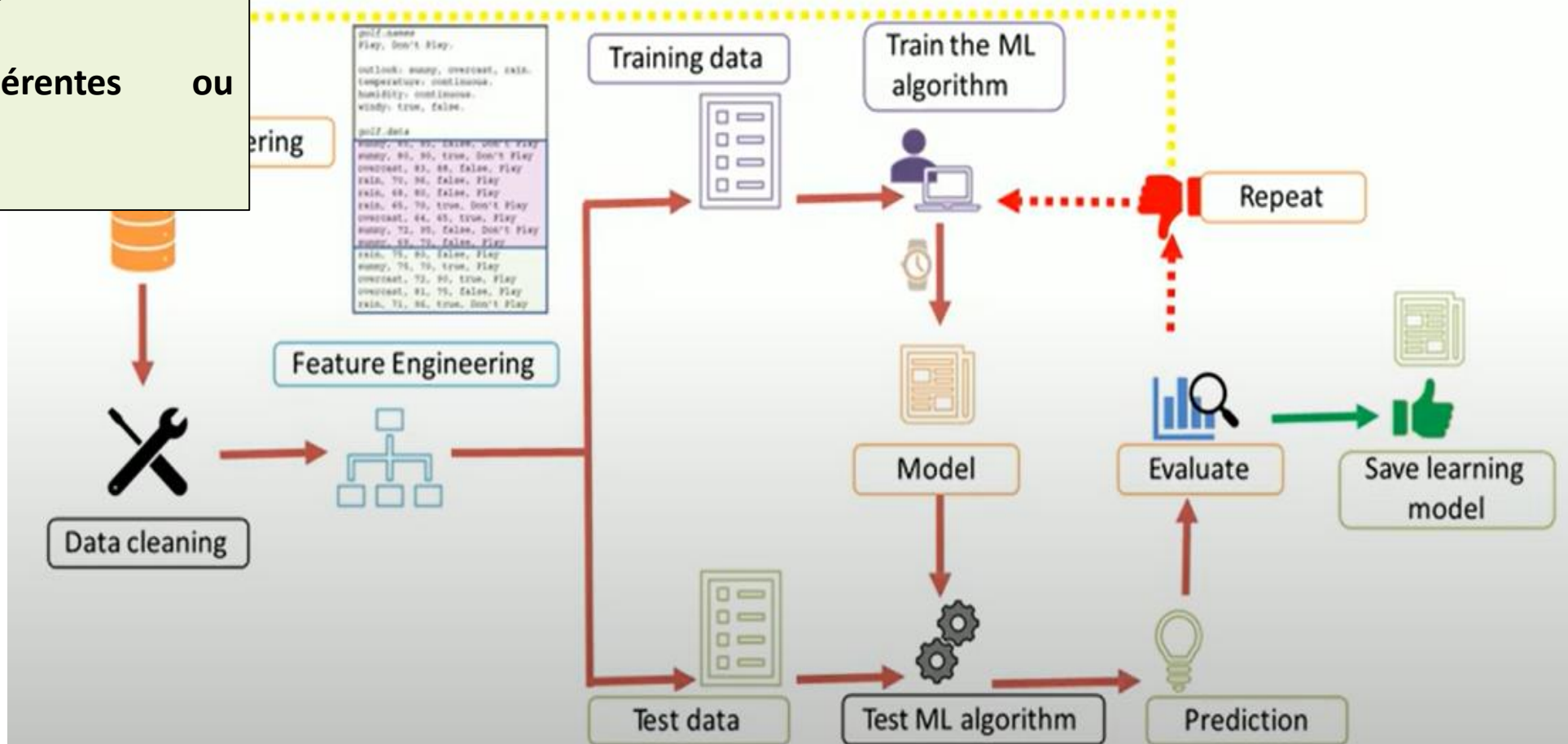
1- Obtenir une BD sur laquelle on va créer notre algorithme. Soit on va prendre une base déjà existante ou créer votre propre BD.



# Machine learning process

2- Nettoyer les données brutes afin d'extraire tout le potentiel :

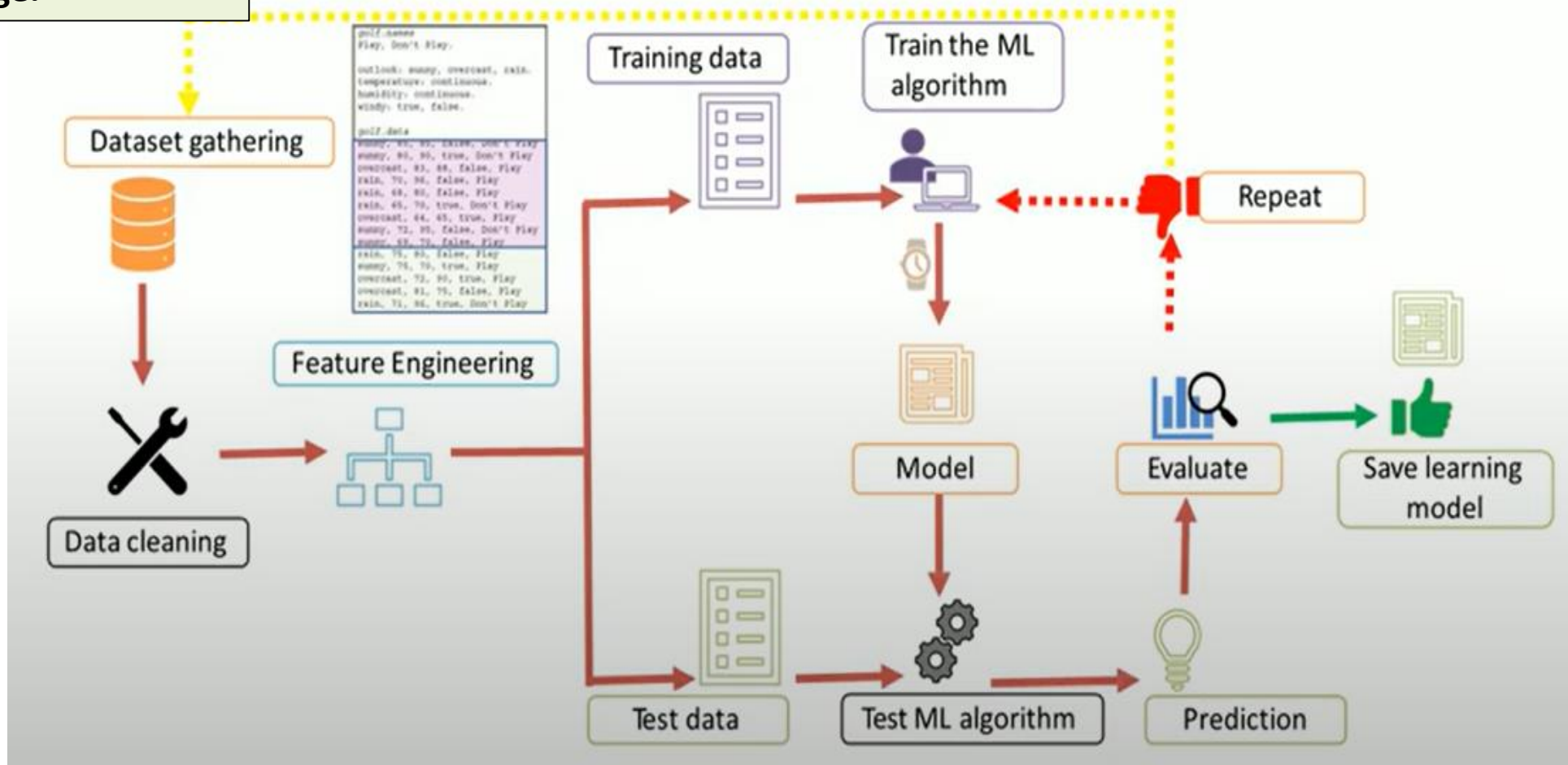
- Données mal annotées
- Doublons
- Informations incohérentes ou superflues
- ...





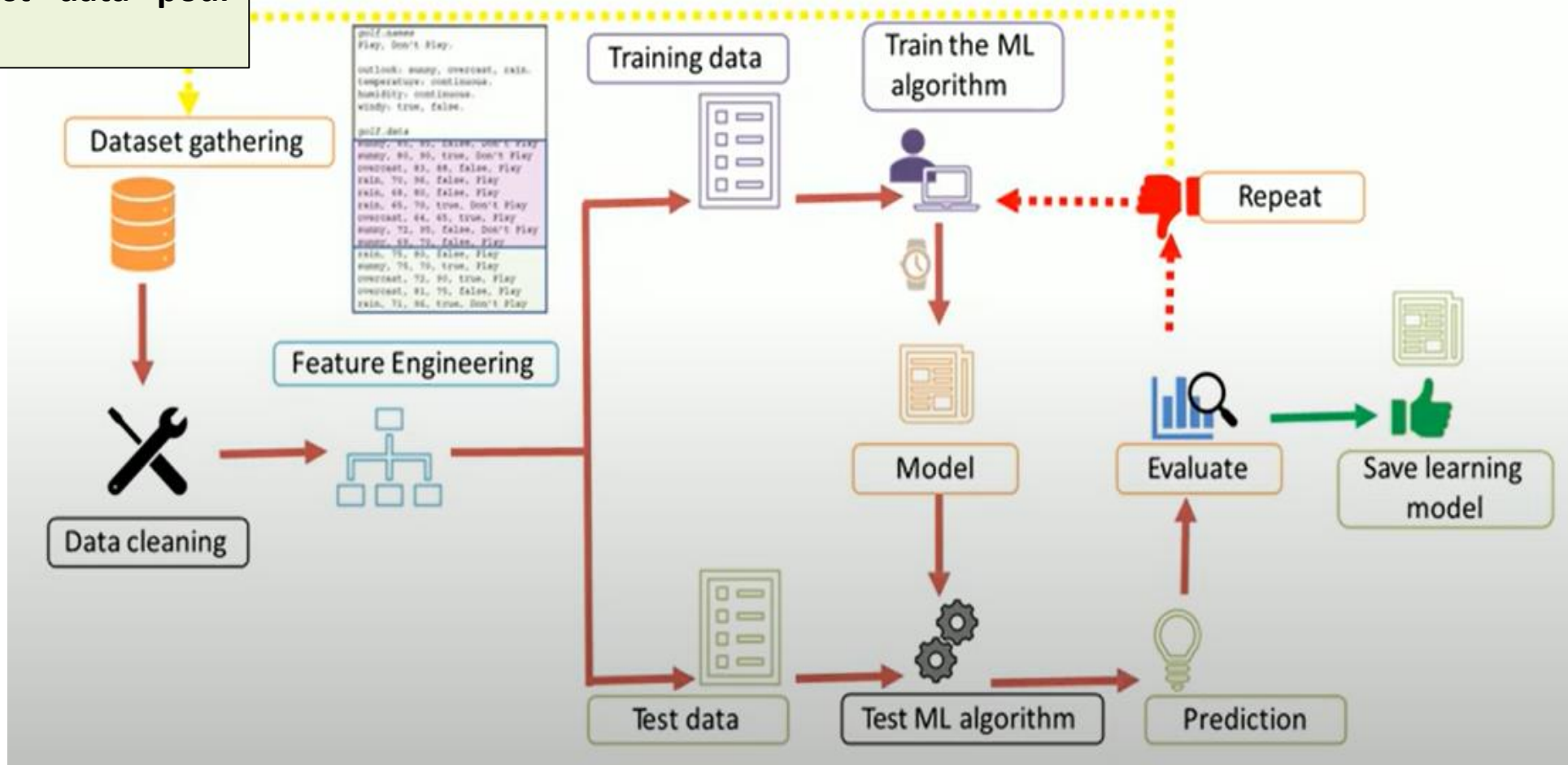
# Machine learning process

3- Extraire un ensemble de caractéristiques qui seront les entrées au système d'apprentissage.



# Machine learning process

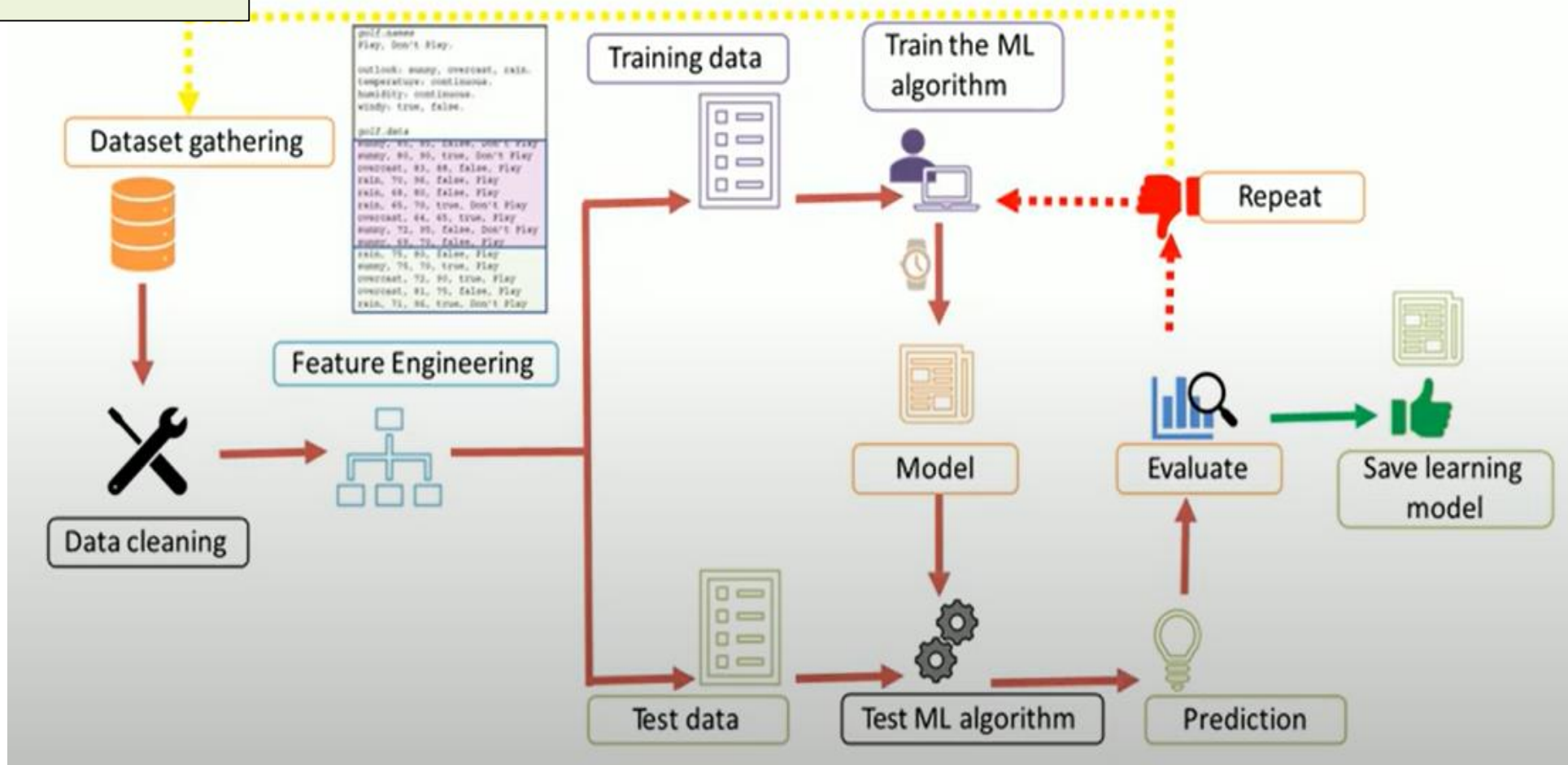
4- Diviser le data set en 2 sous-ensembles, le training data pour l'apprentissage et le test data pour l'évaluation.





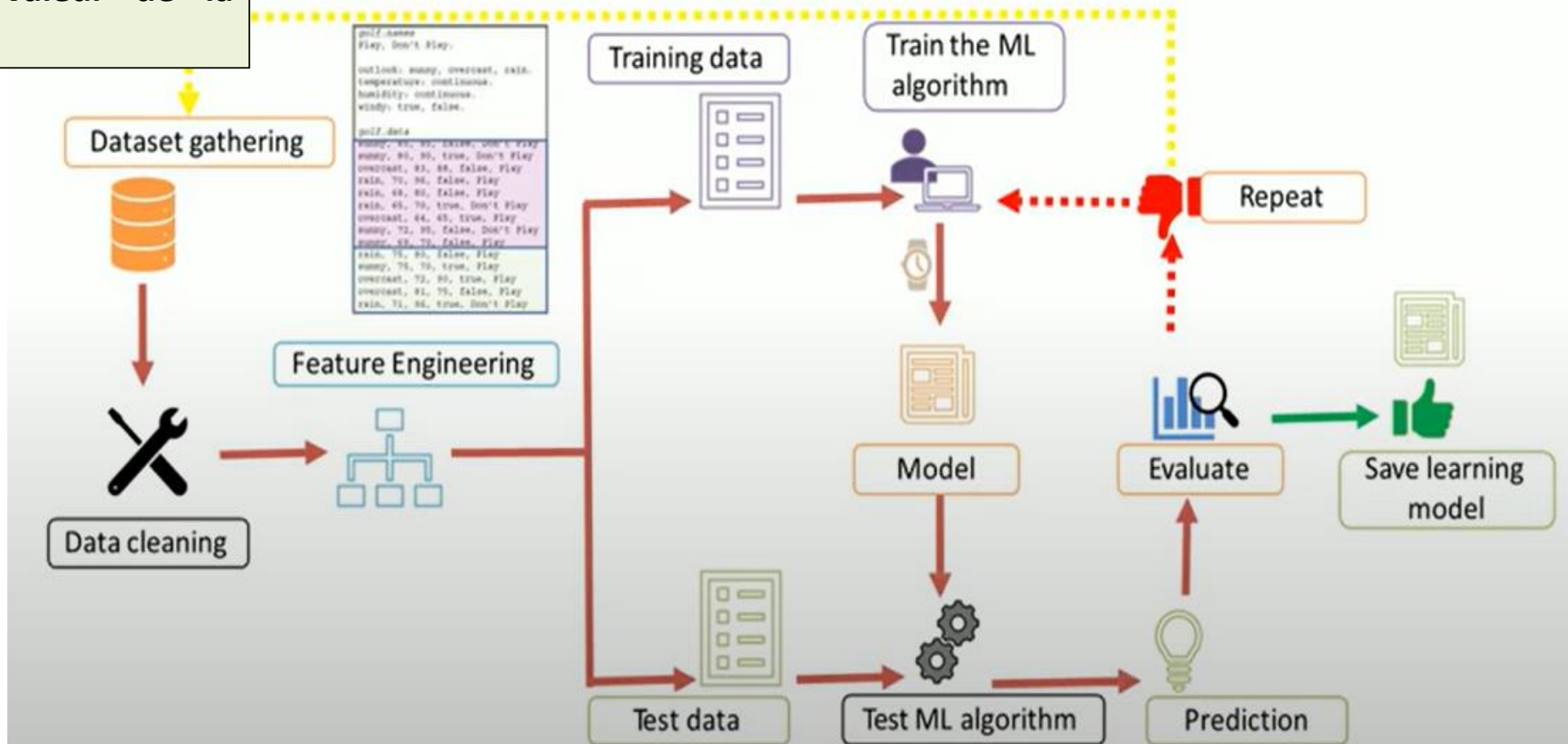
# Machine learning process

5- choisir le type de modèle d'apprentissage et puis on fait l'entraînement.



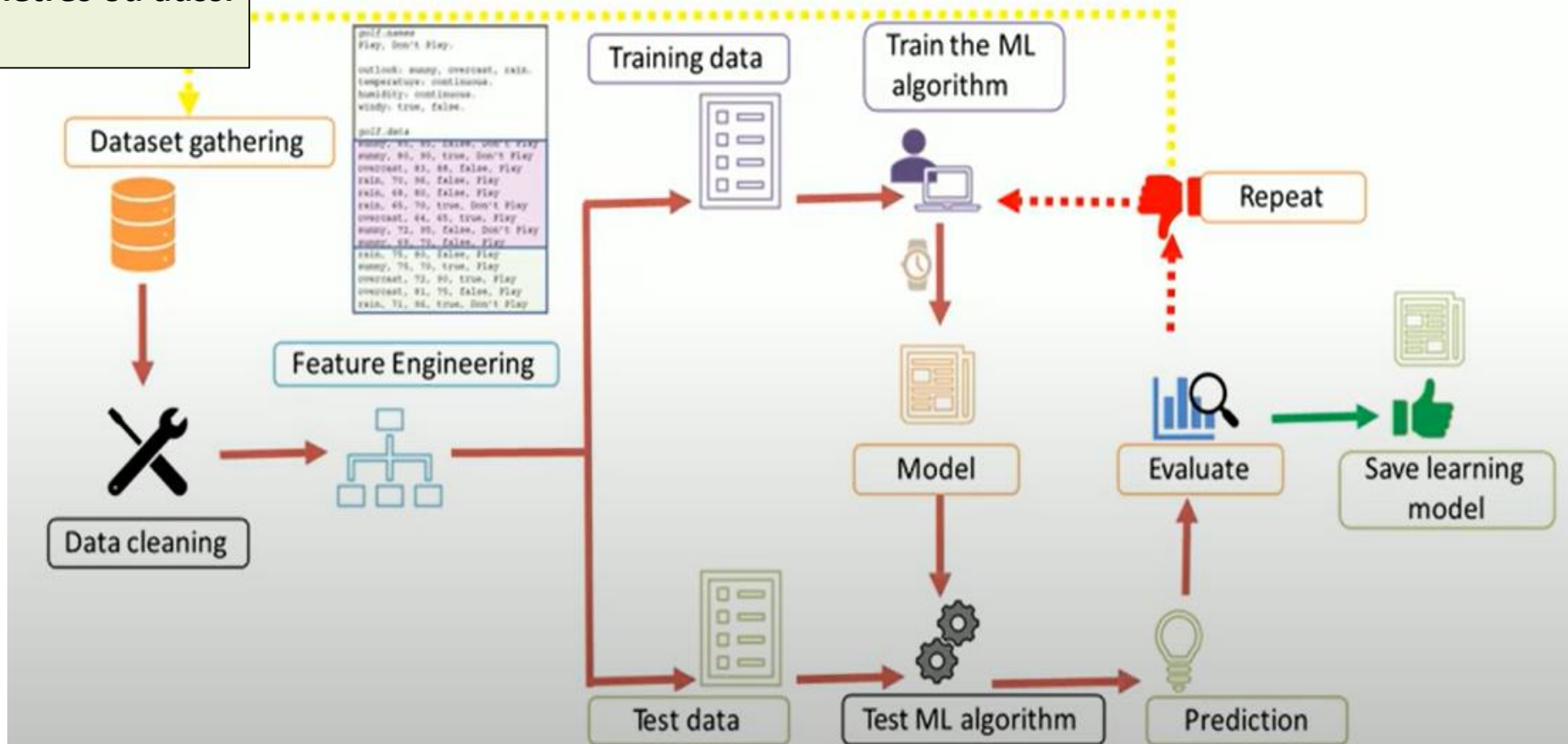
# Machine learning process

6- Evaluer l'efficacité du modèle on faisons une comparaison entre les prédictions et la vrai valeur de la variable cible.



# Machine learning process

7- Si le modèle n'est pas efficace, vous devez revenir vers l'étape de choix d'algorithmes et de paramètres ou aussi vers la BD.

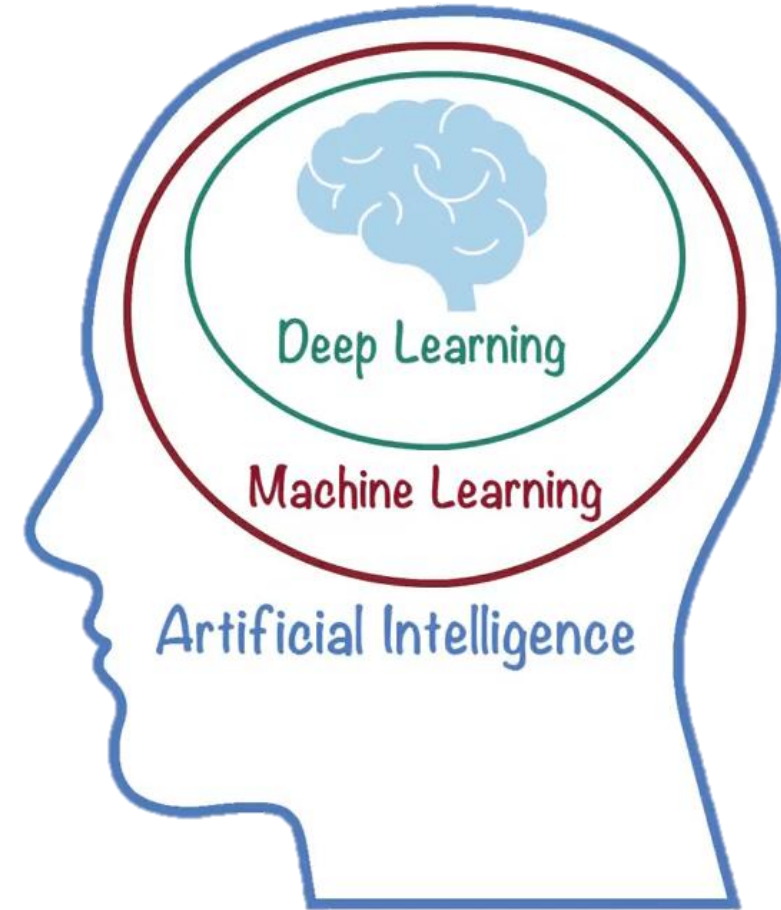


# Machine learning process

❖ **Exemple**

# CHAPITRE 2 :

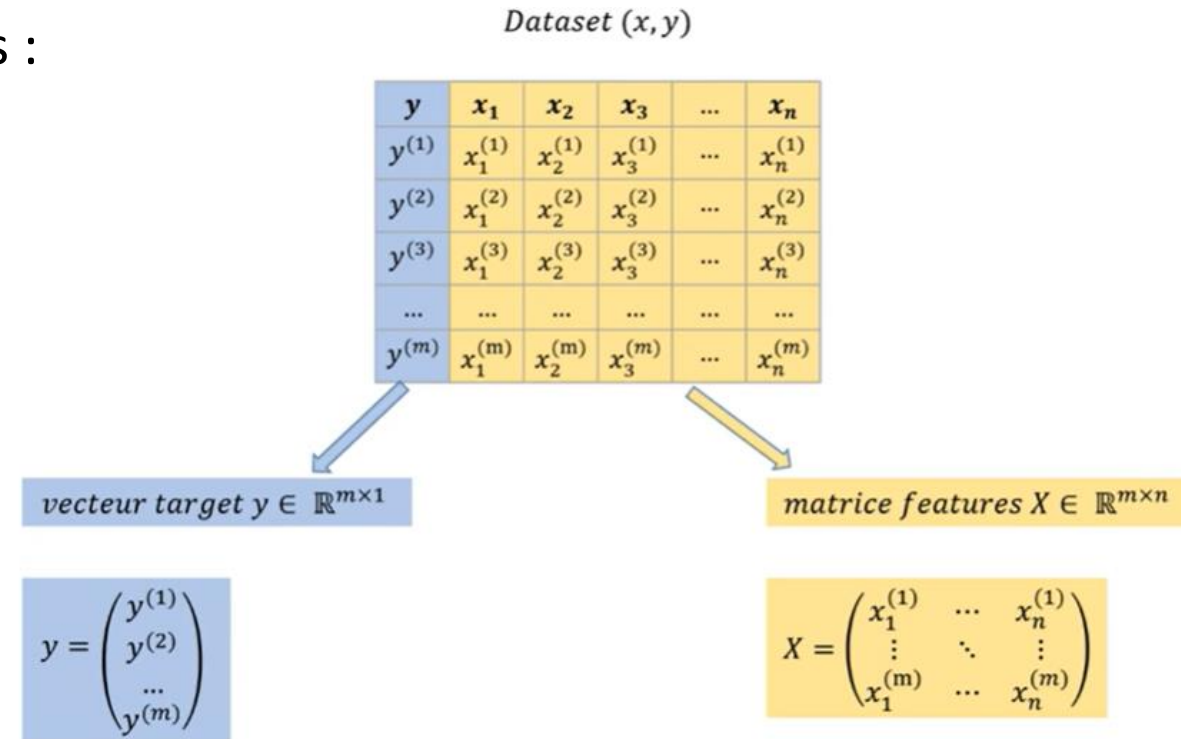
Régression linéaire simple et multiple



# Notions fondamentales de l'apprentissage supervisé

L'apprentissage supervisé fonctionne en 4 étapes :

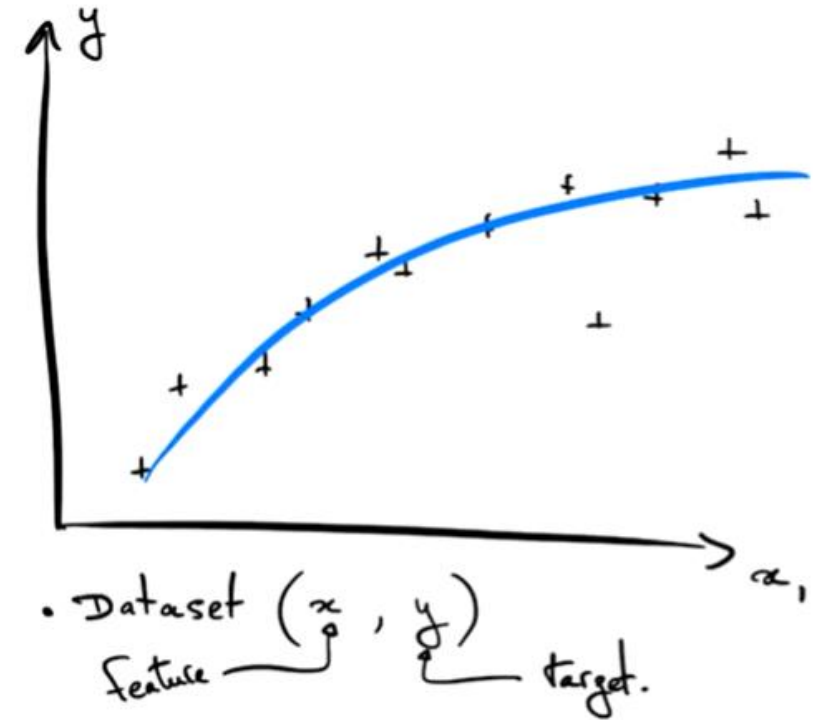
1. Importer les données (Dataset)



# Notions fondamentales de l'apprentissage supervisé

L'apprentissage supervisé fonctionne en 4 étapes :

1. Importer les données (Dataset)
2. Créer un modèle

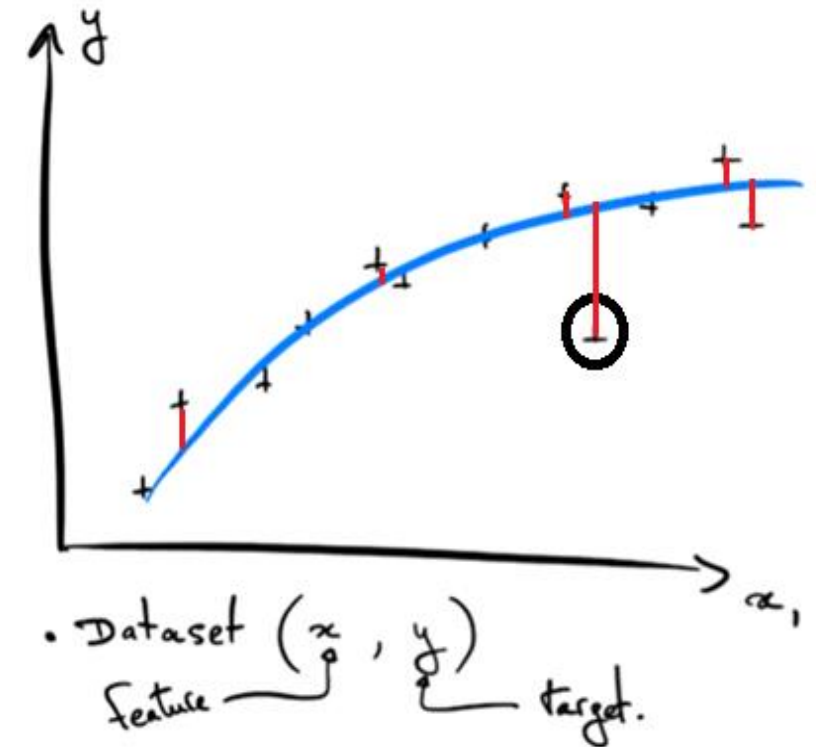


$$• f(x) = ax^2 + bx + c$$

# Notions fondamentales de l'apprentissage supervisé

L'apprentissage supervisé fonctionne en 4 étapes :

1. Importer les données (Dataset)
2. Créer un modèle
3. Calculer la fonction coût



$$• f(x) = ax^2 + bx + c$$



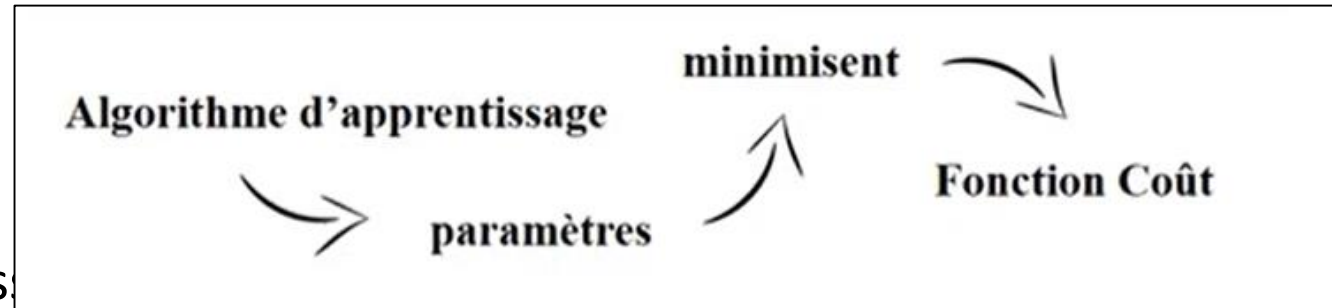
# Notions fondamentales de l'apprentissage supervisé

L'apprentissage supervisé fonctionne en 4 étapes :

1. Importer les données (Dataset)
2. Créer un modèle
3. Calculer la fonction coût
4. Développer un algorithme d'apprentissage

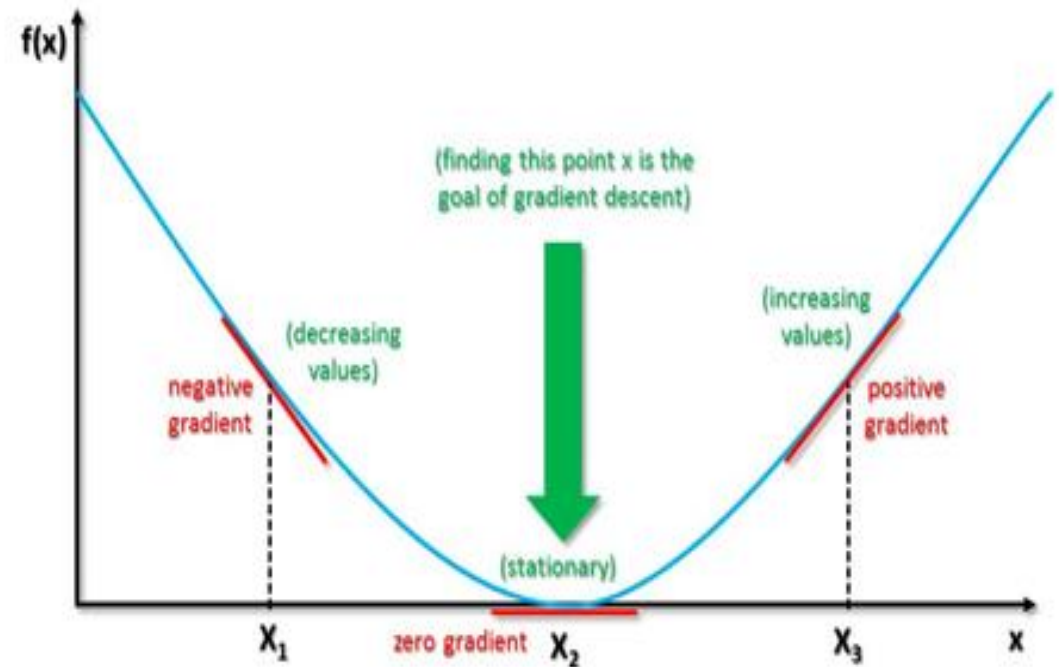
(algorithme de minimisation, par exemple :

**méthode des moindres carrés, descente de gradient)**



# Descente de gradient

- L'apprentissage par descente de gradient permet de modifier les paramètres de la fonction de la régression linéaire afin de minimiser le coût d'erreur.
- Pour minimiser l'erreur, elle a besoin de deux points de données : une direction et un **taux d'apprentissage** (learning rate).
- Ces facteurs déterminent les calculs de dérivée partielle des itérations futures, ce qui lui permet d'arriver progressivement au minimum local ou global (c'est-à-dire au point de convergence).



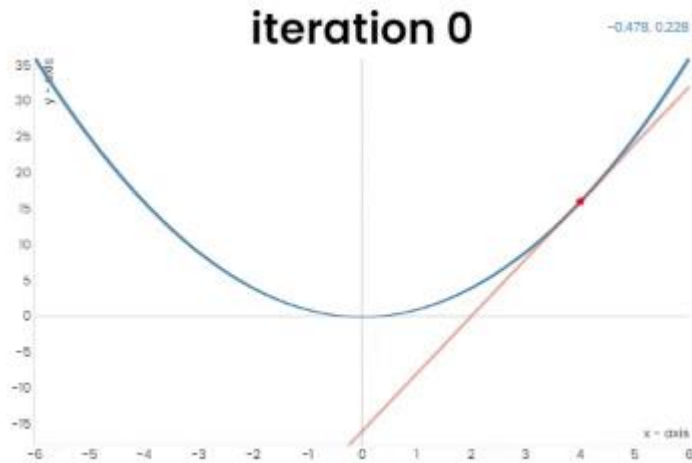
# Descente de gradient

## Comment décider du pas avec lequel je vais avancer?

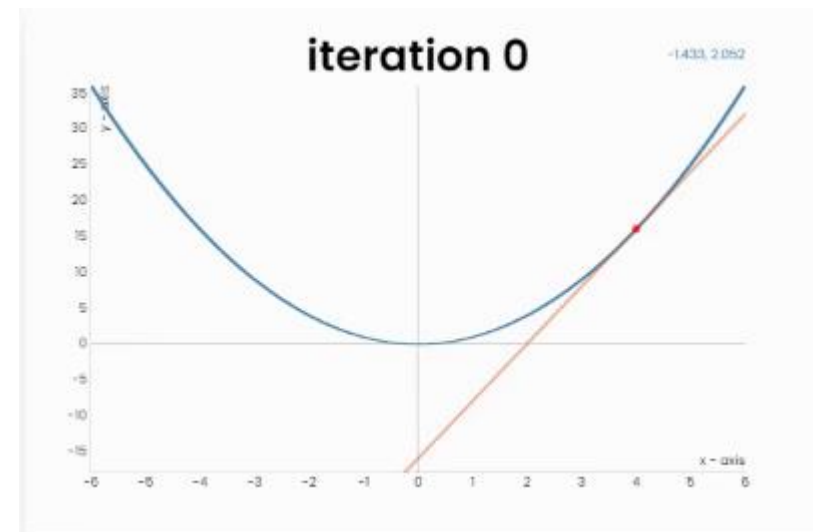
- **Le taux d'apprentissage  $\alpha$  (learning rate)** est un hyper paramètre à régler. Il représente la taille des pas entre deux itérations.
- Le dilemme caché derrière ce paramètre est qu'il permet une convergence rapide pour un pas élevé, mais de petites valeurs assurent plus de stabilité à l'algorithme.
- Un **taux d'apprentissage élevé** se traduit par des pas plus grands, mais risque de dépasser le minimum.
- Un **taux d'apprentissage faible** entraîne des pas de petite taille. S'il présente l'avantage d'une plus grande précision, le nombre d'itérations compromet l'efficacité globale car il faut plus de temps et de calculs pour atteindre le minimum.

# Descente de gradient

Un taux d'apprentissage faible



Un taux d'apprentissage élevé



# Descente de gradient pour la régression linéaire simple

1. **Le dataset** :  $(x, y)$  avec  $m$  exemples

2. **Le modèle** :  $f(x) = ax + b$

3. **La fonction coût** :  $J(a, b) = \frac{1}{2m} \sum_{i=1}^m (ax^{(i)} + b - y^{(i)})^2$

4. **Gradients** :

$$\frac{\partial J(a, b)}{\partial a} = \frac{1}{m} \sum_{i=1}^m x^{(i)} (ax^{(i)} + b - y^{(i)})$$
$$\frac{\partial J(a, b)}{\partial b} = \frac{1}{m} \sum_{i=1}^m (ax^{(i)} + b - y^{(i)})$$

5. **Algorithme de descente de gradient**

$$a = a - \alpha \frac{\partial J(a, b)}{\partial a}$$

$$b = b - \alpha \frac{\partial J(a, b)}{\partial b}$$

