



INTRODUCTION AU MACHINE LEARNING ET À L'APPRENTISSAGE PAR RENFORCEMENT



Formateur



Maxime Carrere
Data Scientist
Scalian
Maxime.carrere@scalian.com

Formateur



Andrea Cortesi
Data Scientist
Scalian
Andrea.cortesi@scalian.com

Formateur



Pierre Leroy
Data Scientist
Scalian
Pierre.leroy@scalian.com

LES OBJECTIFS DE CETTE INTRODUCTION

- Démystifier l'IA et ses nombreux buzzwords
- Découvrir ce qu'est la Datascience, et la positionner par rapport aux nombreuses disciplines du domaine:
 - DEEP LEARNING, MACHINE LEARNING, BIG DATA
- Acquérir une culture générale du machine learning, nécessaire pour comprendre l'intérêt des traitement « mathématiques » de la datascience.



QU'EST CE QUE L'IA ?

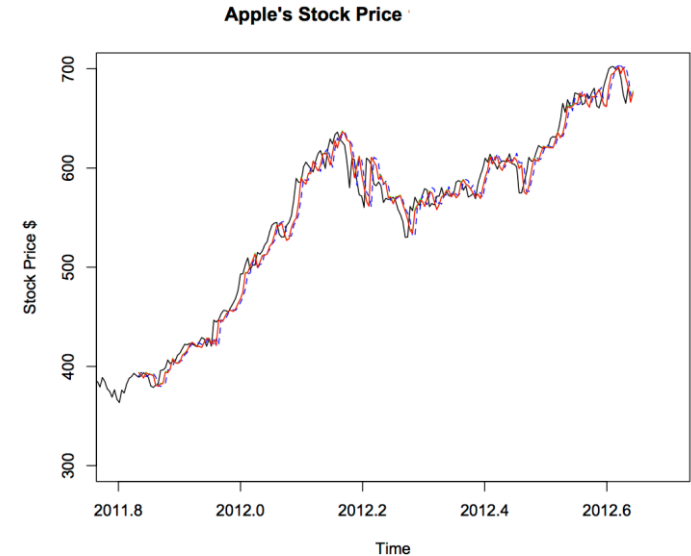
QU'EST CE QUE L'IA ?

INTELLIGENCE ARTIFICIELLE

Statistiques classiques

- Visualisation des données
- Préparation des données

- Utilisation de statistiques pour mieux décrire, « faire parler » et comprendre les données.
- Souvent des méthodes simples comme la moyenne, l'écart-type, la variance, ...
- Permet des tâches simples, comme certaines détections d'anomalies
- Utilisation pour préparer les données avant des traitements plus complexes
- Ex : supprimer des corrélations, normaliser les données, ...



Statistiques classiques

- Visualisation des données
- Préparation des données

INTELLIGENCE ARTIFICIELLE

Systèmes de règles de décision

- Règles pour résoudre la tâche écrite par un expert

Statistiques classiques

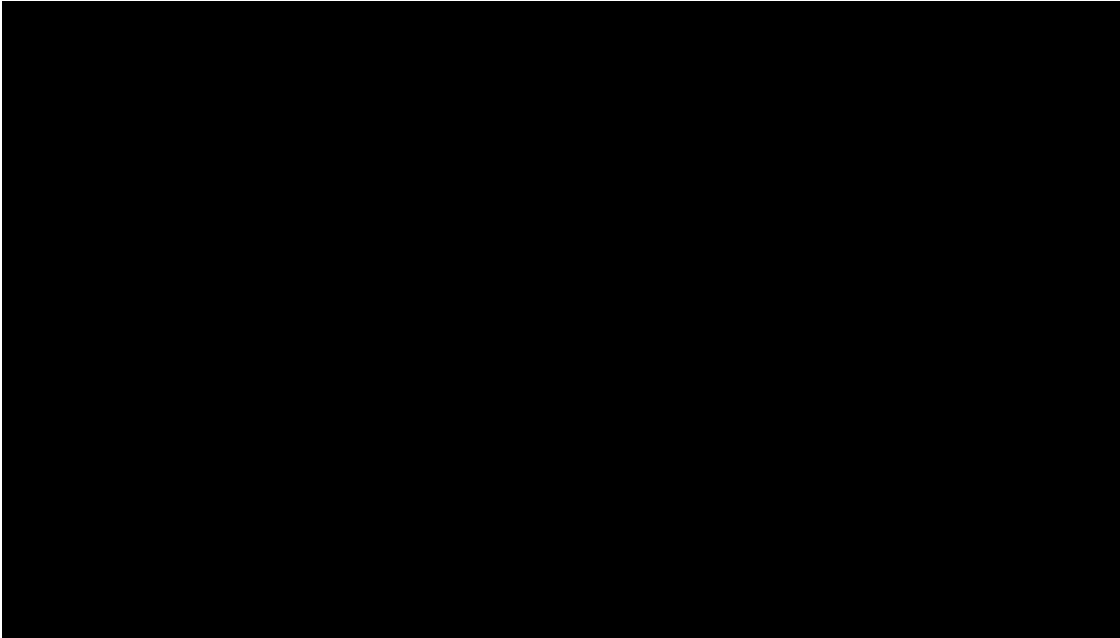
- Visualisation des données
- Préparation des données

INTELLIGENCE ARTIFICIELLE

Systèmes de règles de décision

- Règles pour résoudre la tâche écrite par un expert
- Souvent une simple succession de conditions « if », « then », « else »

EXEMPLE : BOIDS



Fonctionnent uniquement sur 3 règles simples :

- **Séparation** : s'éloigne dans la direction opposée si trop près d'un voisin ou obstacle.
- **Alignement** : prend la direction moyenne de ses voisins.
- **Cohésion** : se positionne pour être au barycentre de ses voisins.

<https://www.youtube.com/watch?v=rw7072URg1A>

Statistiques classiques

- Visualisation des données
- Préparation des données

INTELLIGENCE ARTIFICIELLE

Systèmes de règles de décision

- Règles pour résoudre la tâche écrites par un expert
- Souvent une simple succession de conditions « if », « then », « else »
- Pénible à mettre en place et maintenir
- Pas adapté à des tâches complexes

MOTIVATION OF MACHINE LEARNING

*EXPERT
KNOWLEDGE*



STATISTICS



*IF-ELSE
RULES*

1000 situations = 1000 règles

Difficile à créer, maintenir, faire évoluer

Pourrait-on faire apprendre automatiquement ces règles par la machine ?

«L'apprentissage automatique est la science permettant aux ordinateurs d'accomplir des tâches sans avoir été explicitement programmé dans ce sens.»



Andrew Ng

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .

Données d'entraînement :

9 6 6 5 4 0

Généralisation

6

Statistiques classiques

- Visualisation des données
- Préparation des données

MACHINE LEARNING

- Apprentissage par la machine des règles pour résoudre la tâche

INTELLIGENCE ARTIFICIELLE

Systèmes de règles de décision

- Règles pour résoudre la tâche écrites par un expert
- Souvent une simple succession de conditions « if », « then », « else »
- Pénible à mettre en place et maintenir
- Pas adapté à des tâches complexes

Statistiques classiques

- Visualisation des données
- Préparation des données

MACHINE LEARNING

- Apprentissage par la machine des règles pour résoudre la tâche

DEEP LEARNING

- Par des réseaux de neurones

INTELLIGENCE ARTIFICIELLE

Systèmes de règles de décision

- Règles pour résoudre la tâche écrites par un expert
- Souvent une simple succession de conditions « if », « then », « else »
- Pénible à mettre en place et maintenir
- Pas adapté à des tâches complexes

DATA SCIENCE

Statistiques classiques

- Visualisation des données
- Préparation des données

MACHINE LEARNING

- Apprentissage par la machine des règles pour résoudre la tâche

DEEP LEARNING

- Par des réseaux de neurones

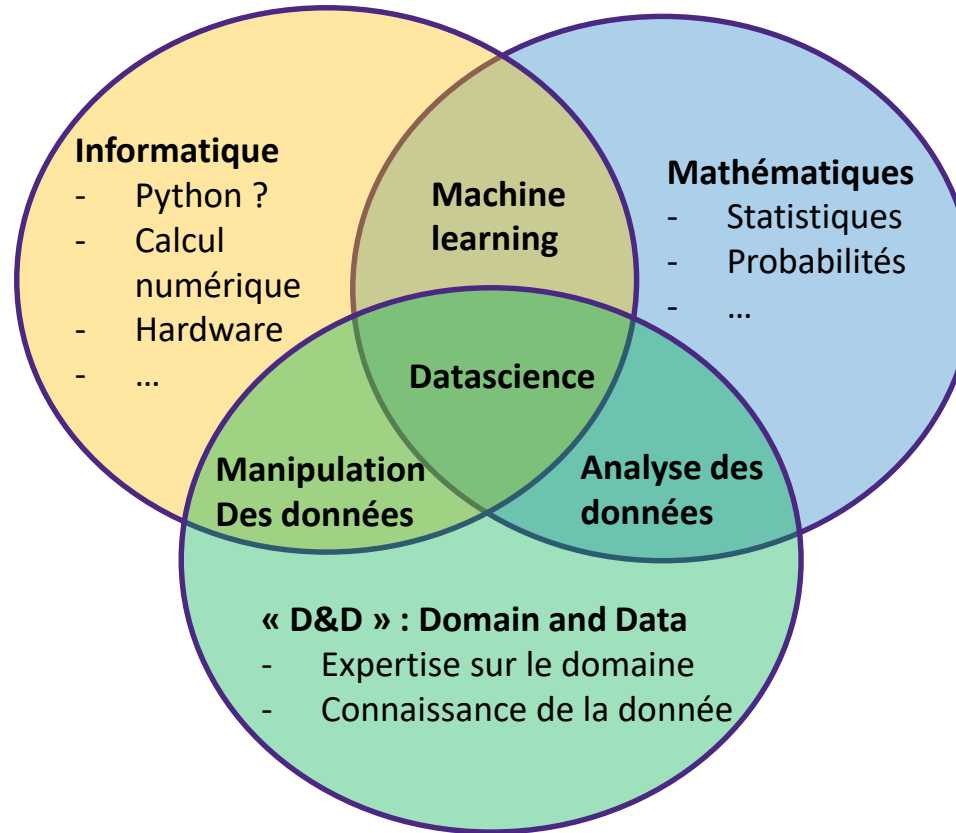
INTELLIGENCE ARTIFICIELLE

Systèmes de règles de décision

- Règles pour résoudre la tâche écrites par un expert
- Souvent une simple succession de conditions « if », « then », « else »
- Pénible à mettre en place et maintenir
- Pas adapté à des tâches complexes

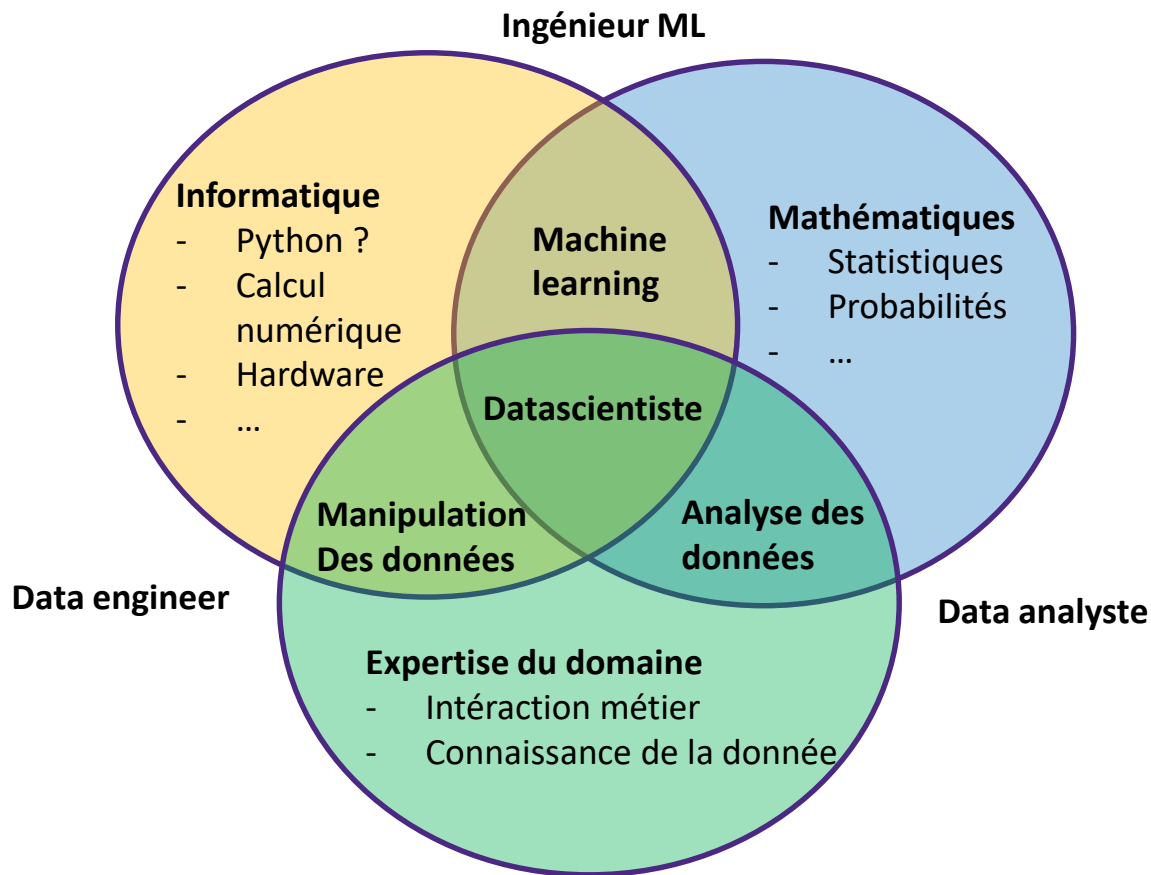
LA DATASCIENCE

UN MÉLANGE DE PLUSIEURS DISCIPLINES



LA DATASCIENCE

UN DOMAINE EN SPÉCIALISATION



QU'EST CE QUE L'IA ?

DATA SCIENCE

BIG DATA

Statistiques classiques

- Visualisation des données
- Préparation des données

MACHINE LEARNING

- Apprentissage par la machine des règles pour résoudre la tâche

DEEP LEARNING

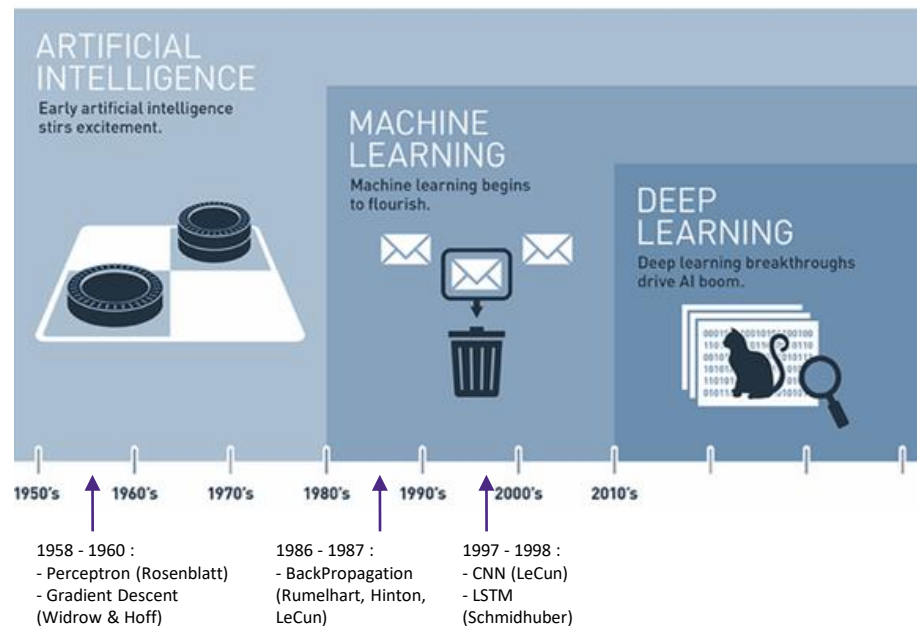
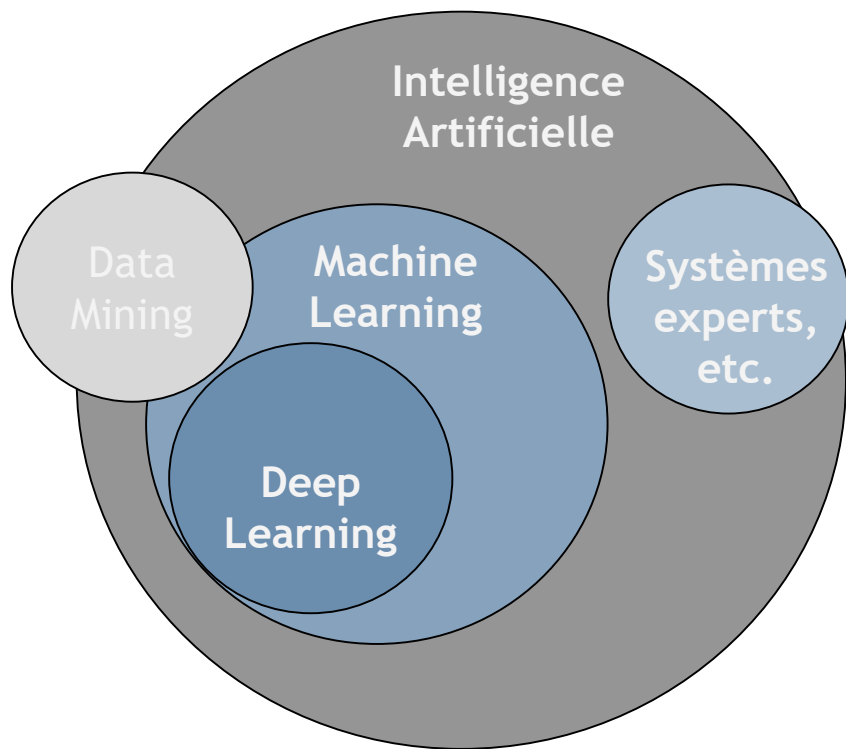
- Par des réseaux de neurones

INTELLIGENCE ARTIFICIELLE

Systèmes de règles de décision

- Règles pour résoudre la tâche écrites par un expert
- Souvent une simple succession de conditions « if », « then », « else »
- Pénible à mettre en place et maintenir
- Pas adapté à des tâches complexes

HISTORIQUE





01

**INTRODUCTION AU MACHINE
LEARNING**

QU'EST CE QUE LE MACHINE LEARNING ?

Définition « large »: Permettre à des programmes d'apprendre des éléments pour lesquels ils n'ont pas été spécifiquement programmés.

Définition « pratique » : des programmes que l'on entraîne avant de faire fonctionner.

Par rapport au méthodes précédentes :

- De meilleurs performances et capable de traiter des problèmes très complexes
- Plus compliqués à mettre en œuvre

Plus de contraintes sur les données et leur volume

Couteux en temps d'entraînement/puissance de calculs

- Plus compliqués à interpréter (mais il existe des solutions)
- Complémentaires

1957 : premier modèle de machine learning apprenant, le perceptron.

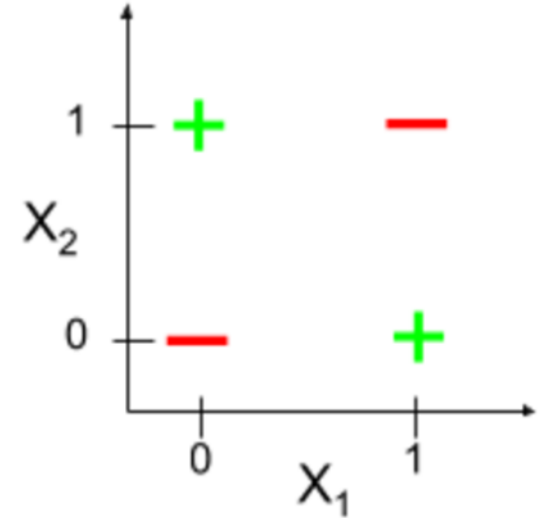
Limites : le XOR

Années 70 : hivers de l'AI : pas de ML !

Années 80: Arbre de décision et ANN.

1990-2010 : Variété de modèles (SVM, Random Forest, Adaboost).

2006 : Deep learning. Atteint des résultats à l'état de l'art dans des domaines très variés(images, NLP....)



1957 : premier modèle de machine learning apprenant, le perceptron.

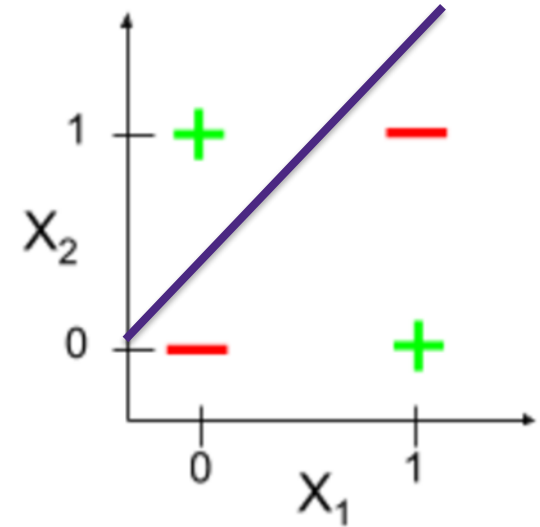
Limites : le XOR

Années 70 : hivers de l'AI : pas de ML !

Années 80: Arbre de décision et ANN.

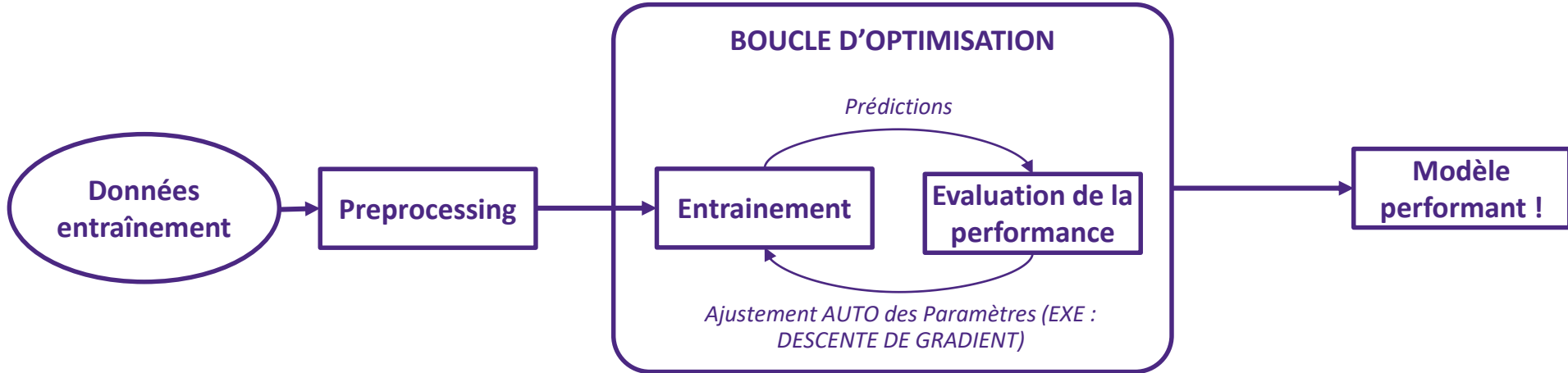
1990-2010 : Variété de modèles (SVM, Random Forest, Adaboost).

2006 : Deep learning. Atteint des résultats à l'état de l'art dans des domaines très variés (images, NLP....)



- **Modèle** : notre algorithme apprenant
- **Poids** : Variables stockant ce qu'apprend le modèle lors de l'entraînement.
- **Input/entrées** : données telles que passées en entrée du modèle.
- **Features** (« caractéristiques ») : les propriétés encodant chaque entrée.
- **Output/prediction** : Sortie du modèle sur une entrée.
- **Label** (« étiquette ») : Valeur à prédire associée à chaque entrée
- **Fonction de coût/de perte/ objective** : fonction à optimiser pendant l'entraînement
- **Outliers** (anomalies) : valeurs « différentes », qui contrastent avec l'ensemble des données.

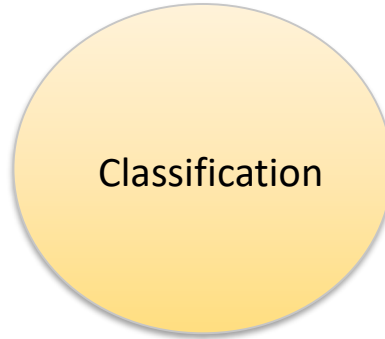
LE PIPELINE CLASSIQUE



- **Dépend des données et de l'algorithme choisi.**
- **Suivant le projet, le pré-processing peut être la tâche la plus importante en terme de temps consacré mais aussi d'impact sur les performance.**
- **Peut contenir :**
 - **Nettoyage des données**
 - **Equilibrage du jeu de donnée**
 - **Génération de donnée**
 - **Modification des données**
 - **Extraction de features**
- **Bénéficie souvent des connaissances et intuitions d'un expert métier**

Il faut faire la différence entre

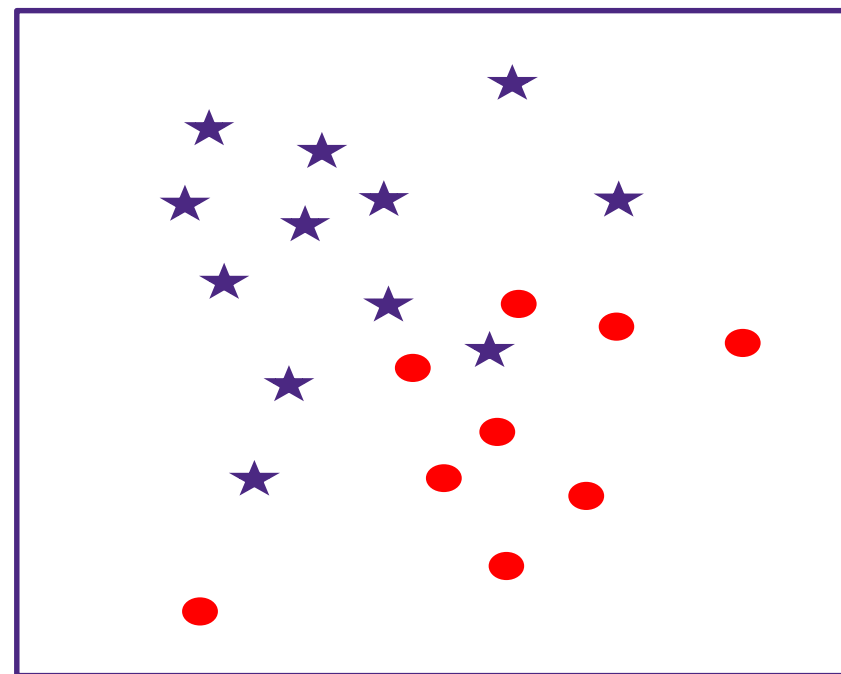
- les différentes **catégories de problèmes/tâches** que peut résoudre le ML
 - Trier les patates
 - Faire la différence entre chien et chat
 - Prédire la météo
 - La conduite d'une voiture autonome
 - ...
- les **méthodes d'apprentissage** pour apprendre résoudre ces problèmes
 - Cf slides suivantes
 - L'apprentissage par renforcement !



LA CLASSIFICATION

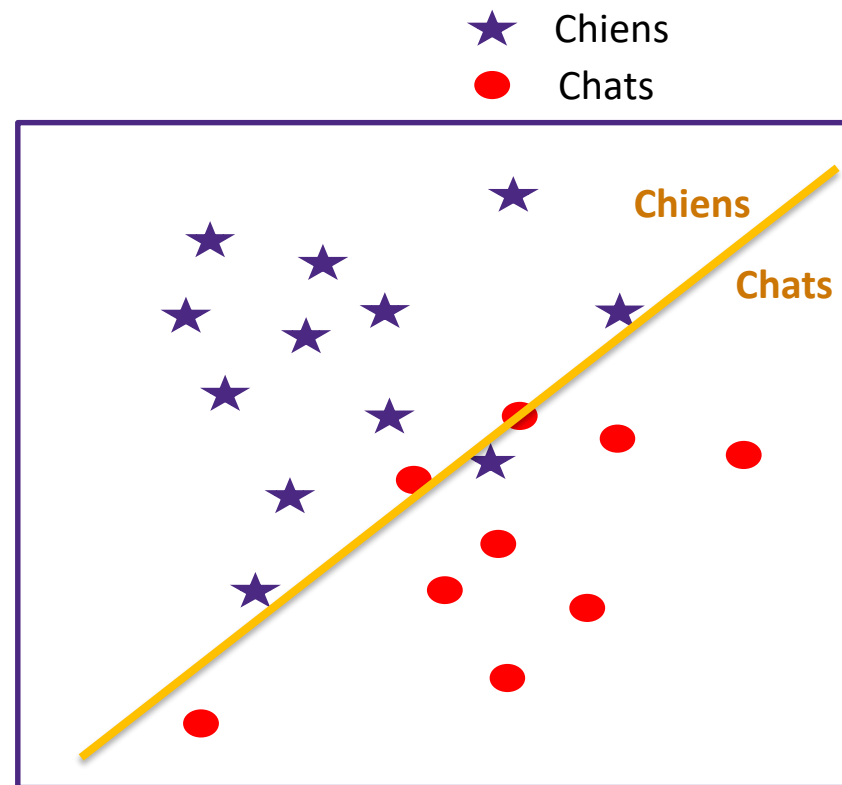
Classer les données en catégories, ou classes, connues au préalable.

★ Chiens
● Chats



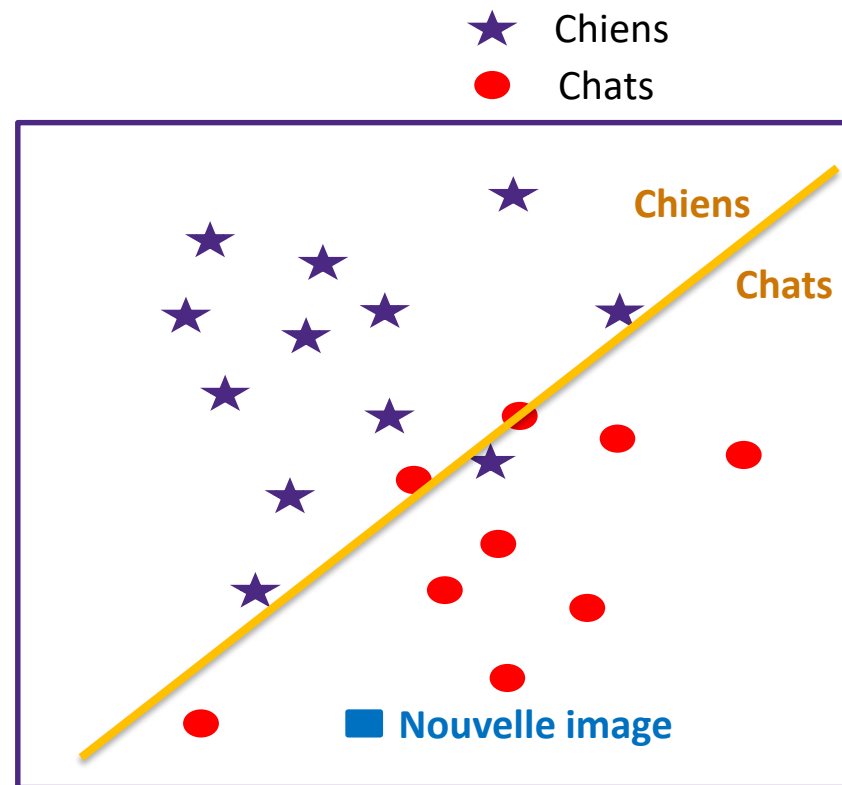
LA CLASSIFICATION

Classer les données en catégories, ou classes, connues au préalable.



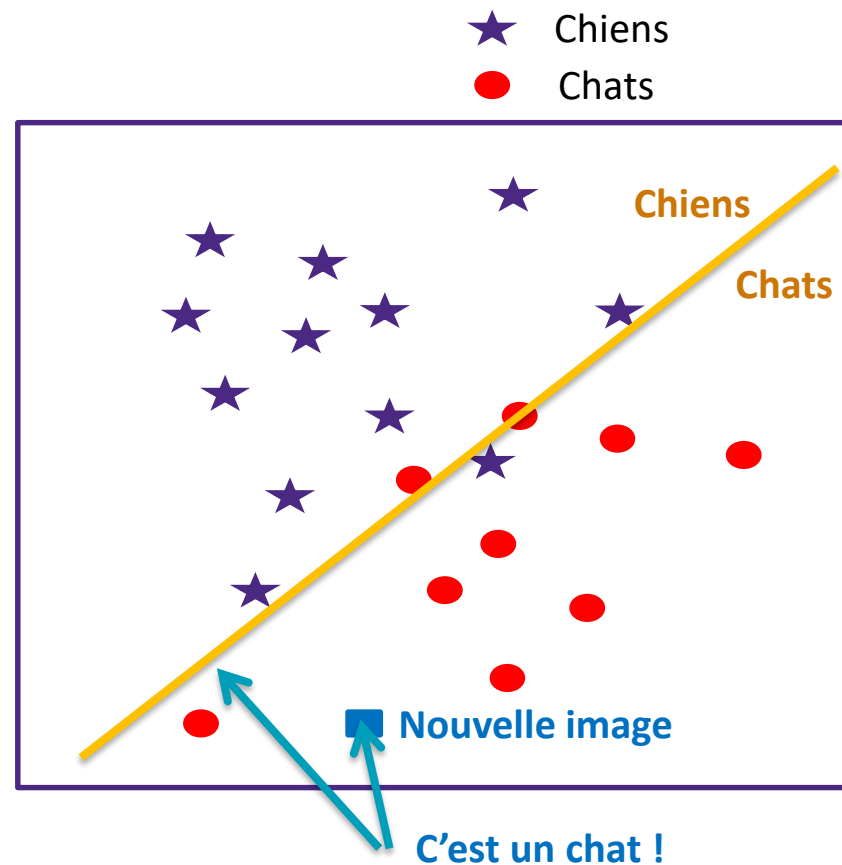
LA CLASSIFICATION

Classer les données en catégories, ou classes, connues au préalable.



LA CLASSIFICATION

Classer les données en catégories, ou classes, connues au préalable.

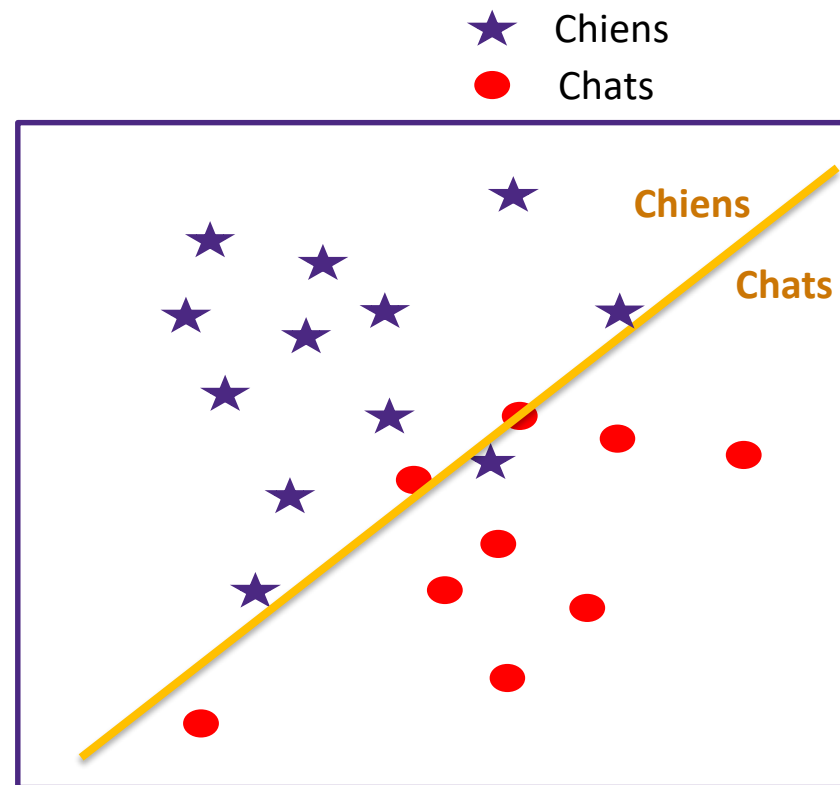


LA CLASSIFICATION

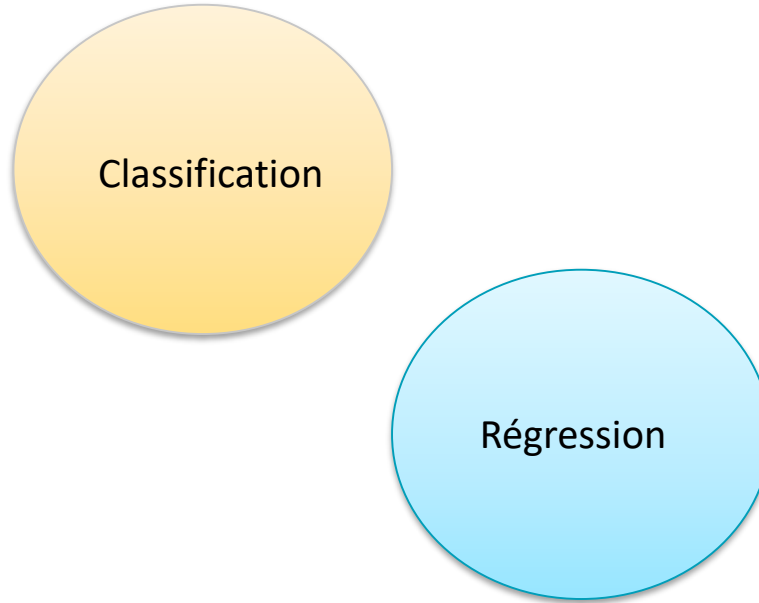
Classer les données en catégories, ou classes, connues au préalable.

Quelques exemples d'applications :

- Identifier l'animal présent dans une image.
- Séparer les spams des mails normaux
- Classer les utilisateurs des sites webs par profil d'intérêts.
- ...



LES DIFFÉRENTES TÂCHES DU ML

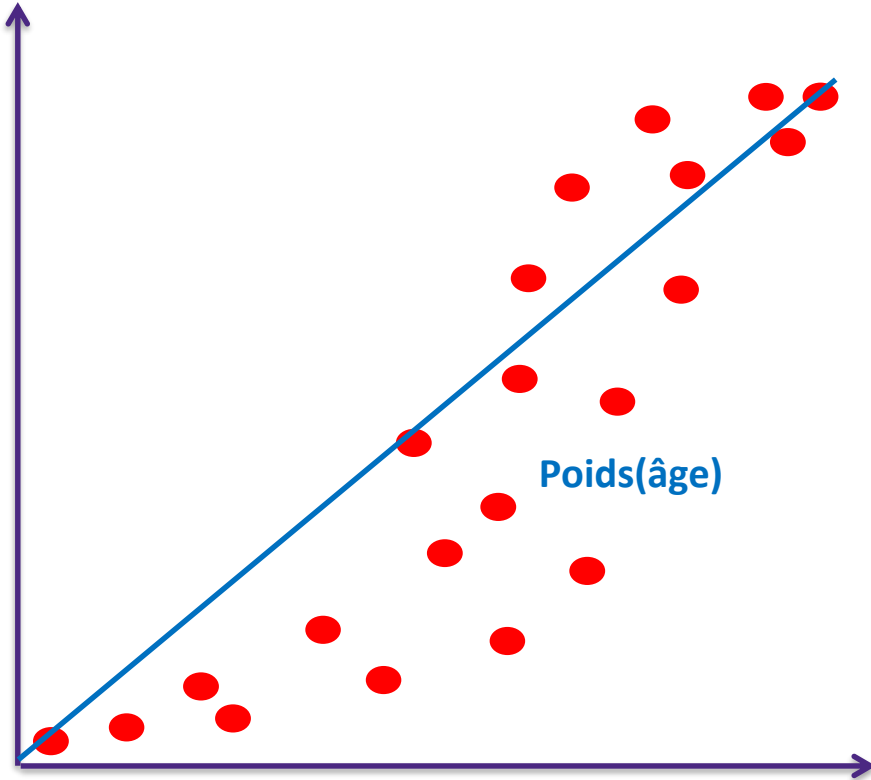


LA REGRESSION LINÉAIRE

● Chats

Estimer une valeur à partir des données

Poids du
chat



Poids(âge)

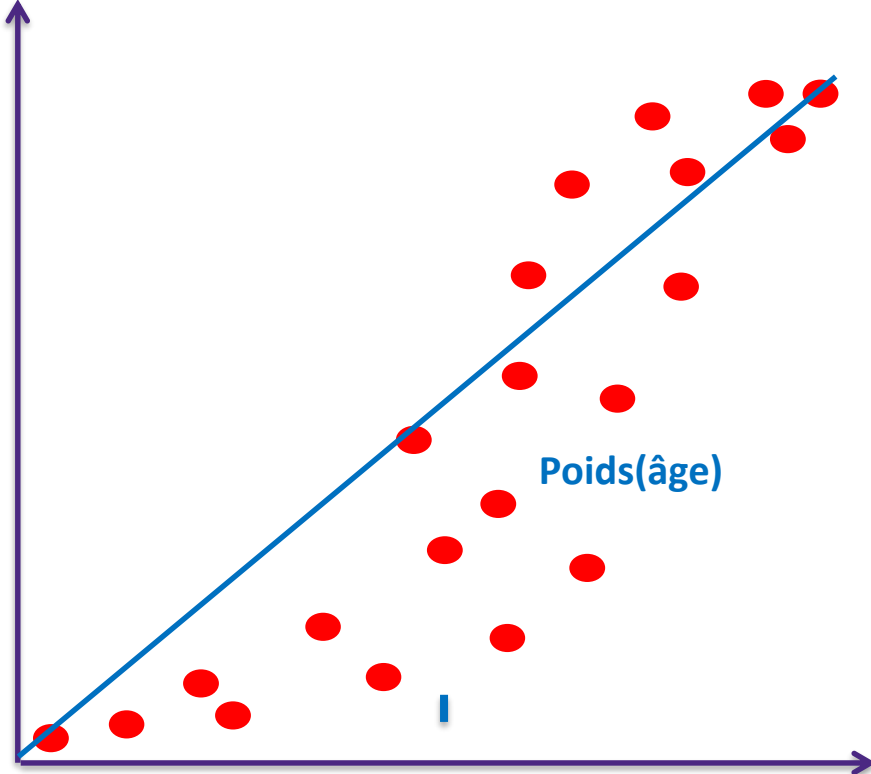
Âge du chat

LA REGRESSION LINÉAIRE

● Chats

Estimer une valeur à partir des données

Poids du
chat



Nouveau chat

Âge du chat

LA REGRESSION LINÉAIRE

● Chats

Estimer une valeur à partir des données

Voici son poids estimé !

Poids du
chat

Poids(âge)

Nouveau chat

Âge du chat

Régression

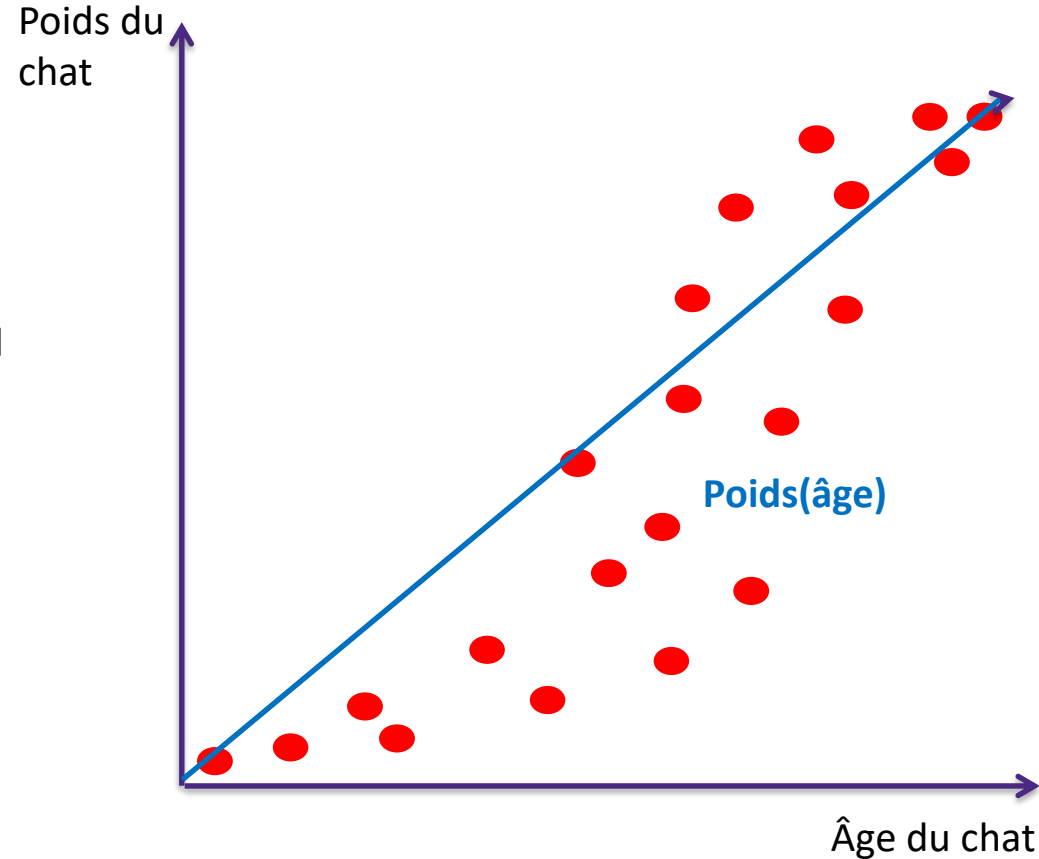
LA REGRESSION LINÉAIRE

● Chats

Estimer une valeur continue à partir des données

Quelques exemples d'applications :

- Prédire la météo et la température qu'il fera demain
- Prédire l'effet d'une campagne de publicité sur les achats effectués
- Prédire les performances de différents modèles pour utiliser le meilleur sur une entrée donnée.

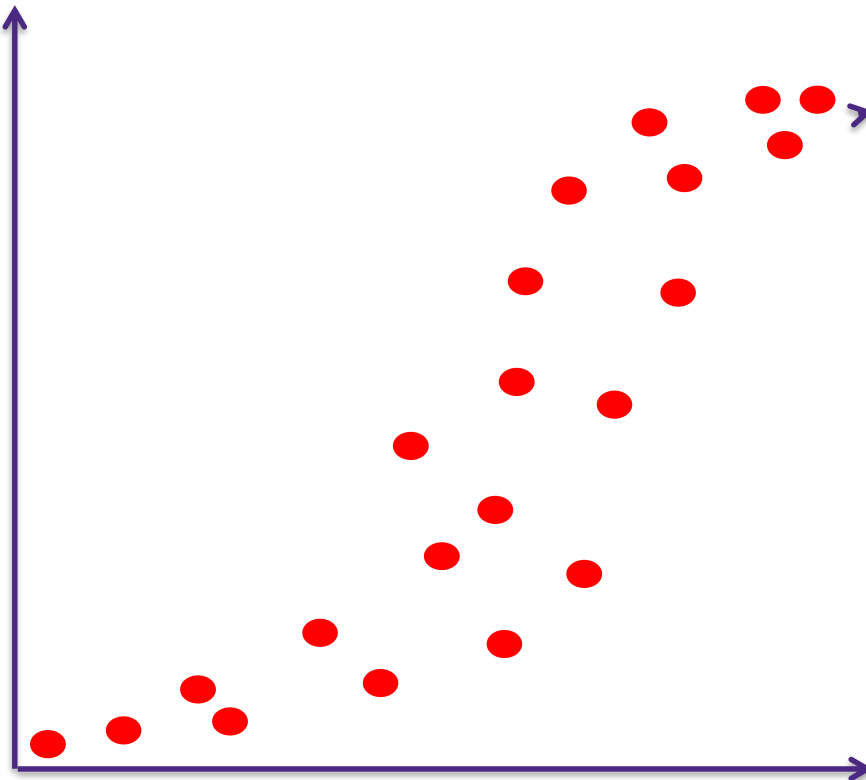


LA REGRESSION

● Chats

Estimer une valeur à partir des données

Poids du
chat



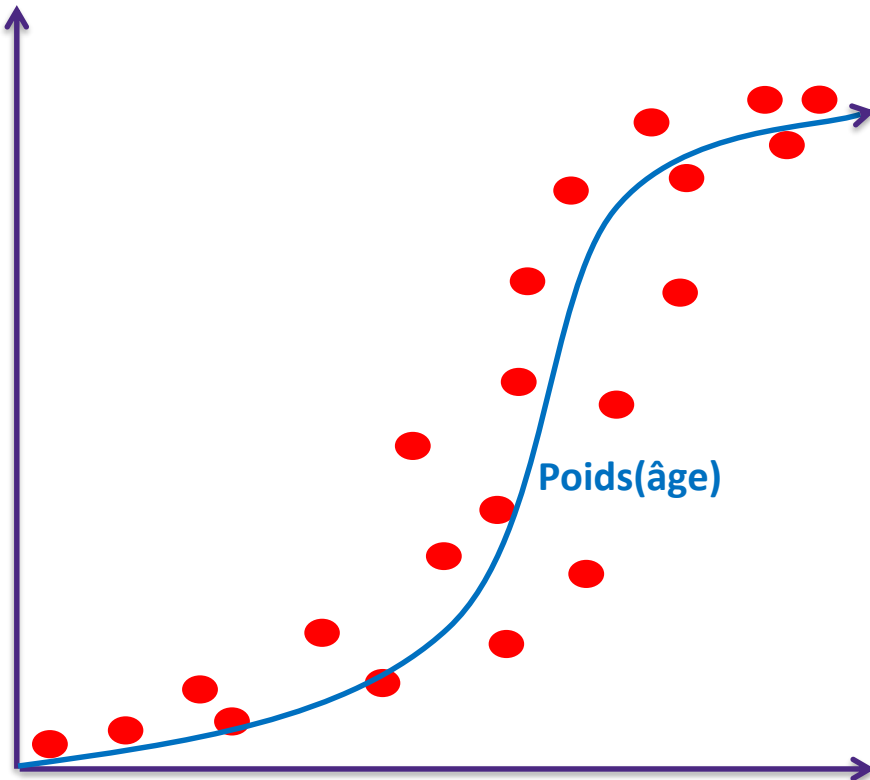
Âge du chat

LA REGRESSION

● Chats

Estimer une valeur à partir des données

Poids du
chat



Poids(âge)

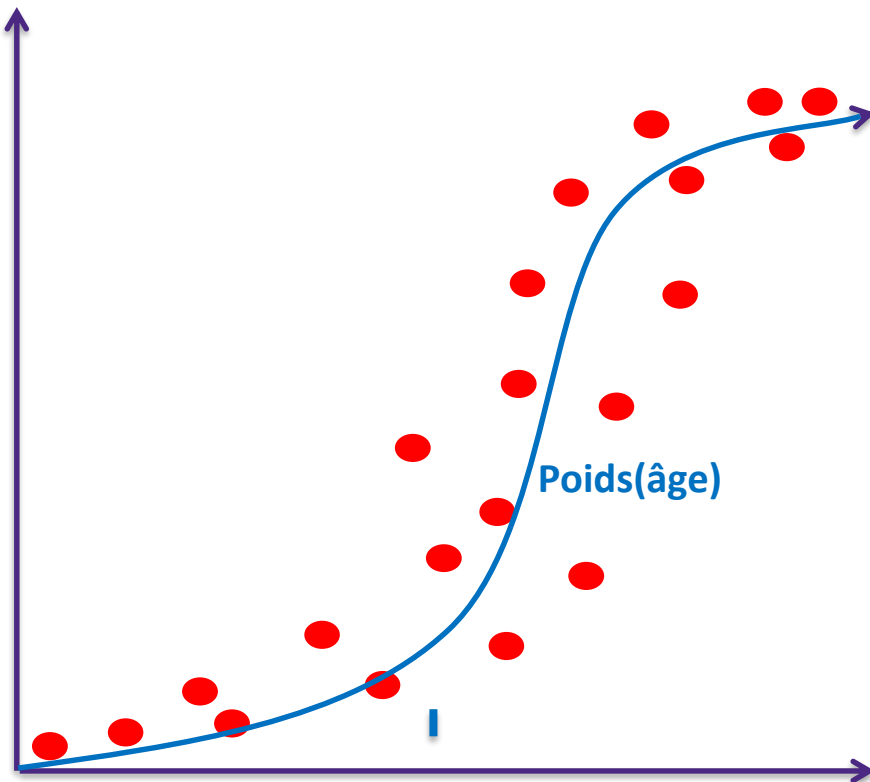
Âge du chat

LA REGRESSION

Estimer une valeur à partir des données

● Chats

Poids du
chat



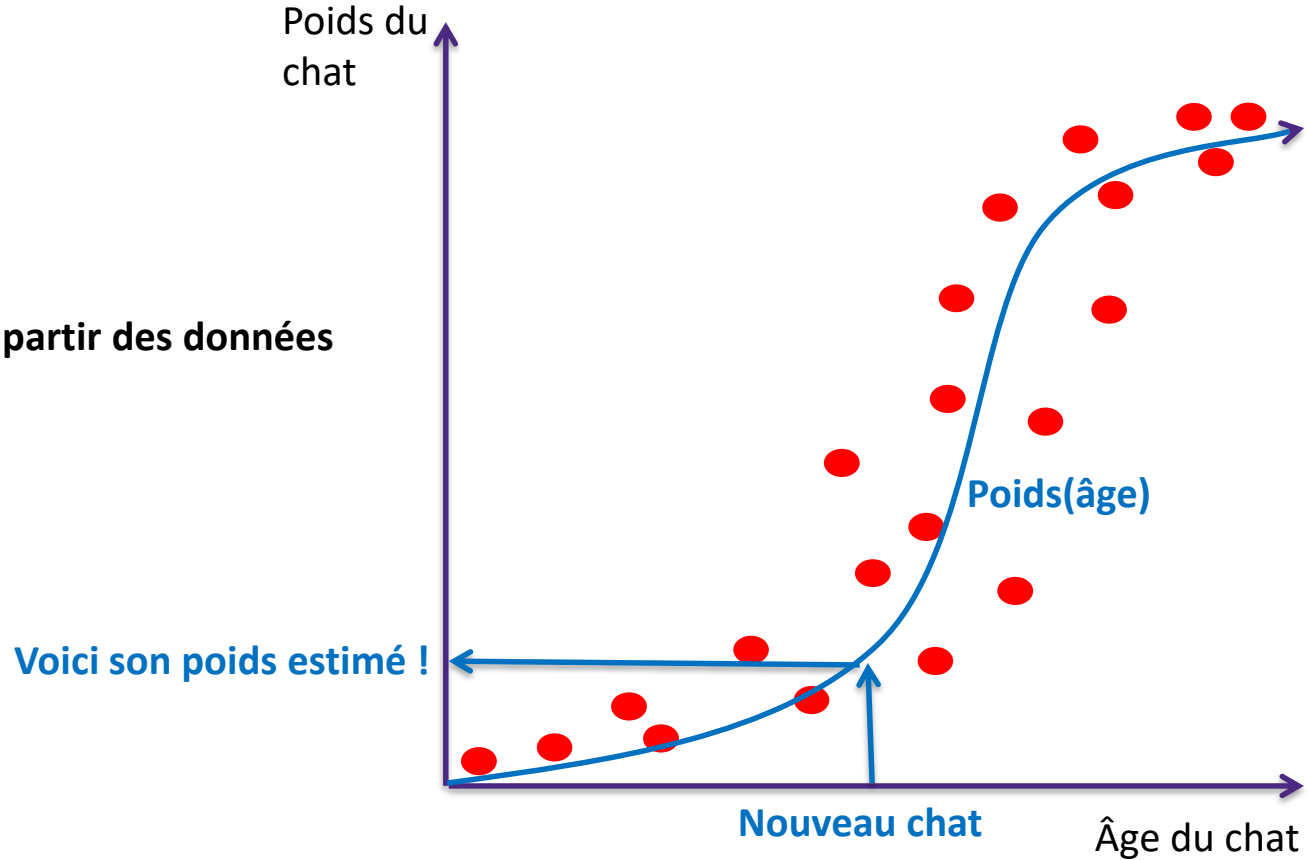
Âge du chat

Nouveau chat

LA REGRESSION

● Chats

Estimer une valeur à partir des données



Régression

LA REGRESSION

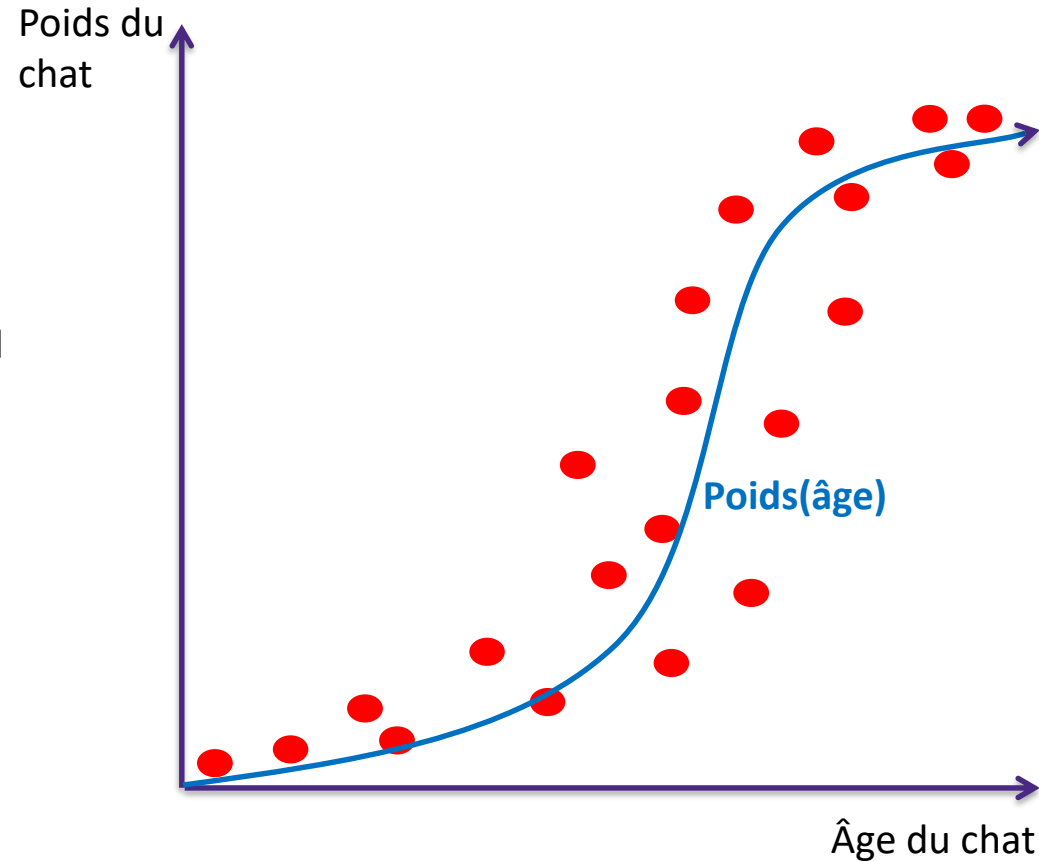
Estimer une valeur continue à partir des données

Quelques exemples d'applications :

- Prédire la météo et la température qu'il fera demain
- Prédire l'effet d'une campagne de publicité sur les achats effectués
- Prédire les performances de différents modèles pour utiliser le meilleur sur une entrée donnée.

• ...

● Chats



CLASSIFICATION / REGRESSION

QUIZZ

Classification

Regression

Connaître l'âge de mon prochain acheteur

Prédire le temps de rentabilité d'une action

Connaître le genre (masculin féminin) de mon prochain acheteur

Estimer le prix d'une voiture d'occasion

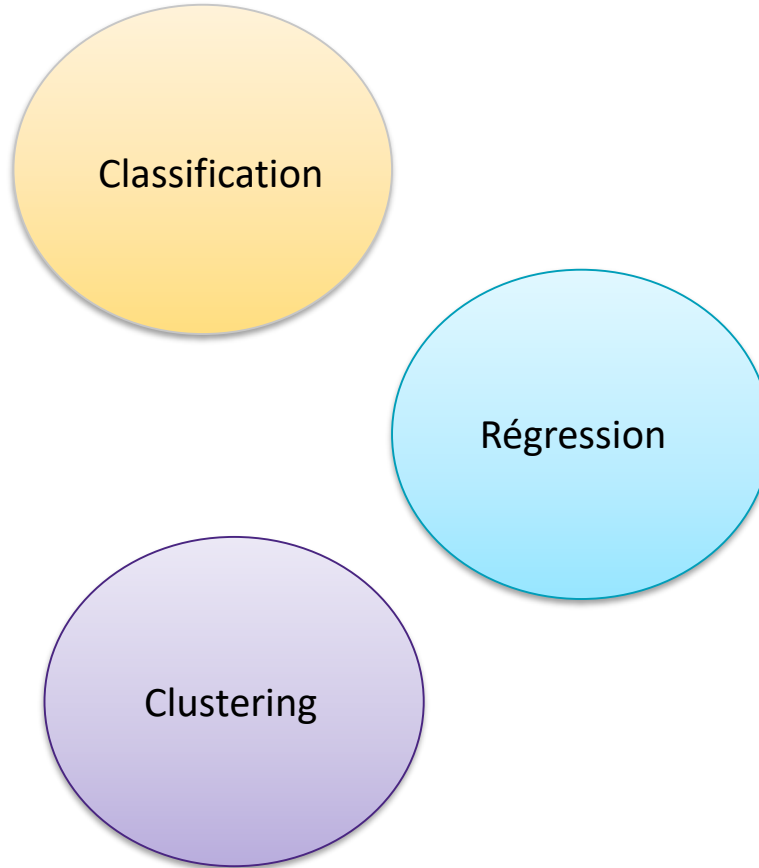
Détecter les objets dans une image

CLASSIFICATION / RÉGRESSION

QUIZZ

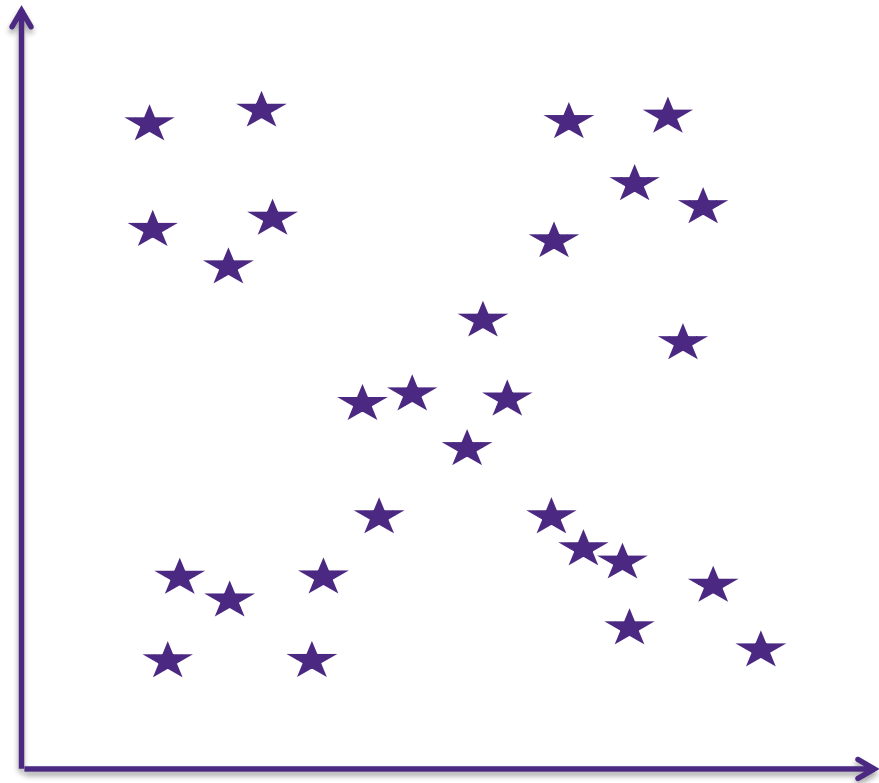
	Classification	Régression
Connaître l'âge de mon prochain acheteur		✓
Prédire le temps de rentabilité d'une action		✓
Connaître le genre (masculin féminin) de mon prochain acheteur	✓	
Estimer le prix d'une voiture d'occasion		✓
Détecter les chats et les chiens dans une image	✓	✓

LES DIFFÉRENTES TÂCHES DU ML



Partitionner les données en différents groupes ou « cluster »

Satisfaction



Montant dépensé

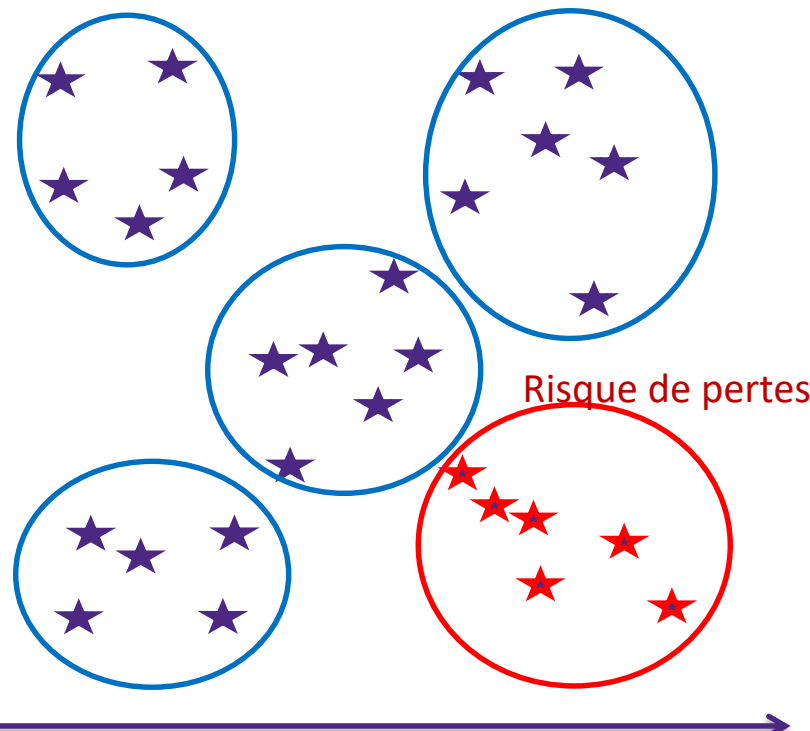
Clustering

LE CLUSTERING

★ Utilisateurs

Partitionner les données en différents groupes ou « cluster »

Satisfaction
sur le dernier
achat



Montant mensuel dépensé

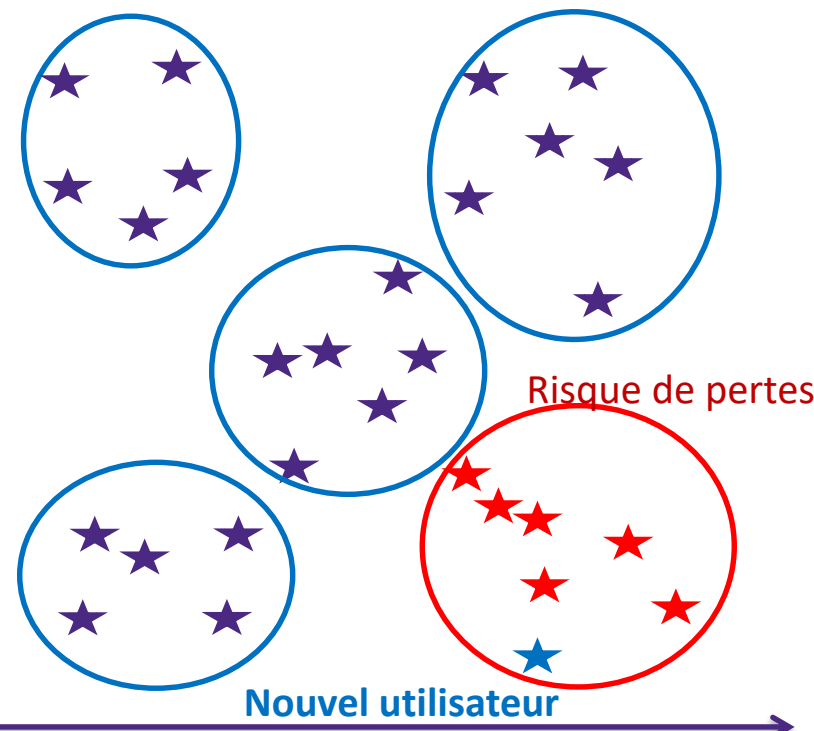
Clustering

LE CLUSTERING

Partitionner les données en différents groupes ou « cluster »

★ Utilisateurs

Satisfaction
sur le dernier
achat



Nouvel utilisateur

Montant mensuel dépensé

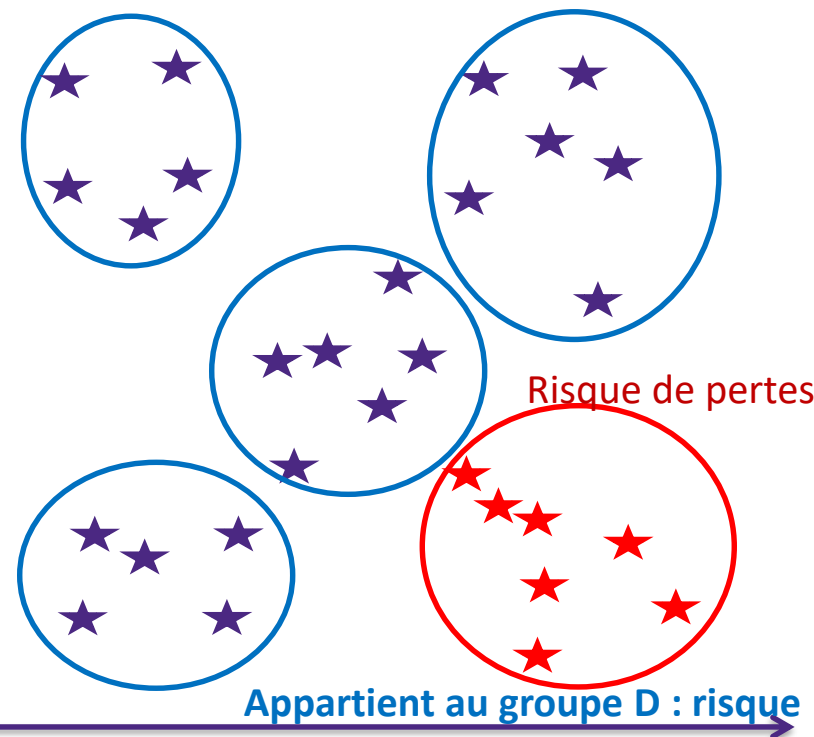
Clustering

LE CLUSTERING

Partitionner les données en différents groupes ou « cluster »

★ Utilisateurs

Satisfaction
sur le dernier
achat



Clustering

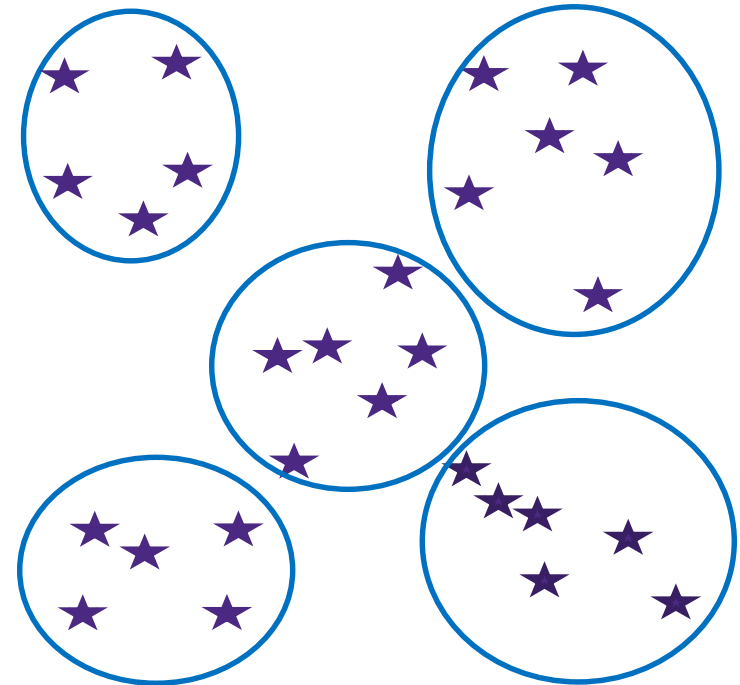
LE CLUSTERING

Partitionner les données en différents groupes ou « cluster »

Exemples d'applications:

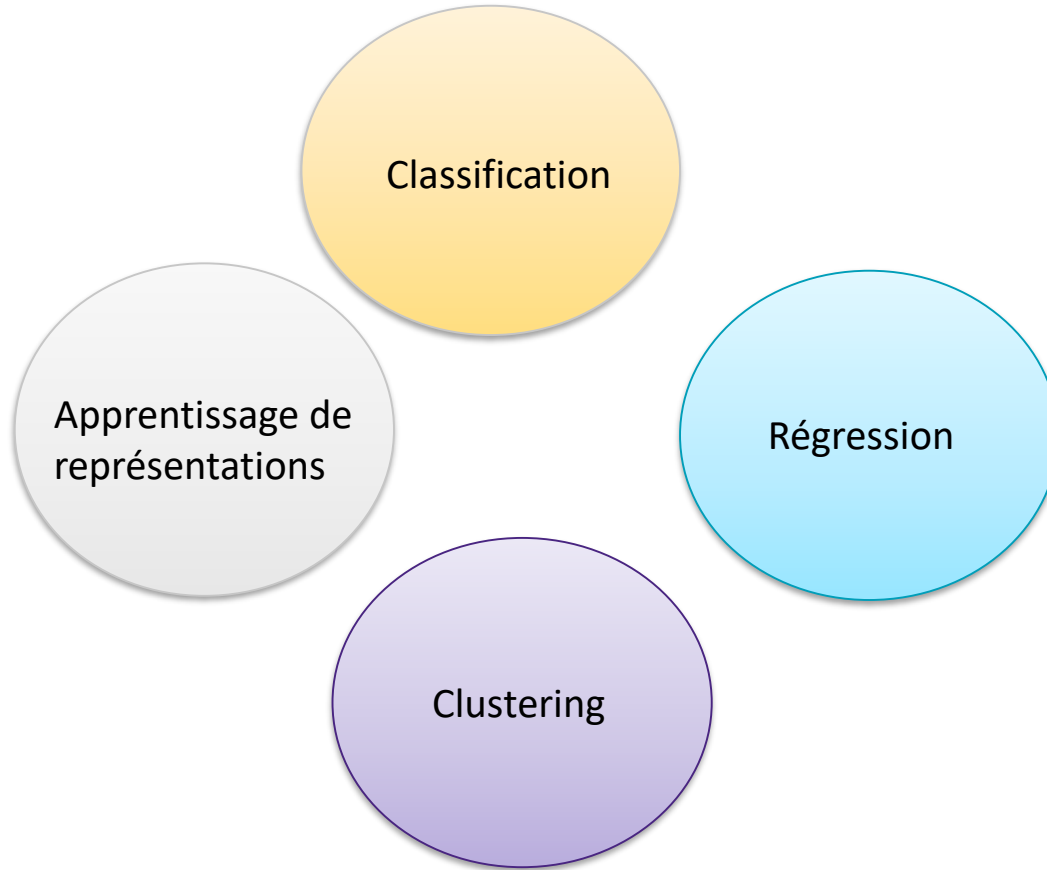
- **Profilage d'utilisateurs**
- **Détection de communautés dans les réseaux sociaux**
- **Détection d'anomalies !**
- ...

Satisfaction
sur le dernier
achat



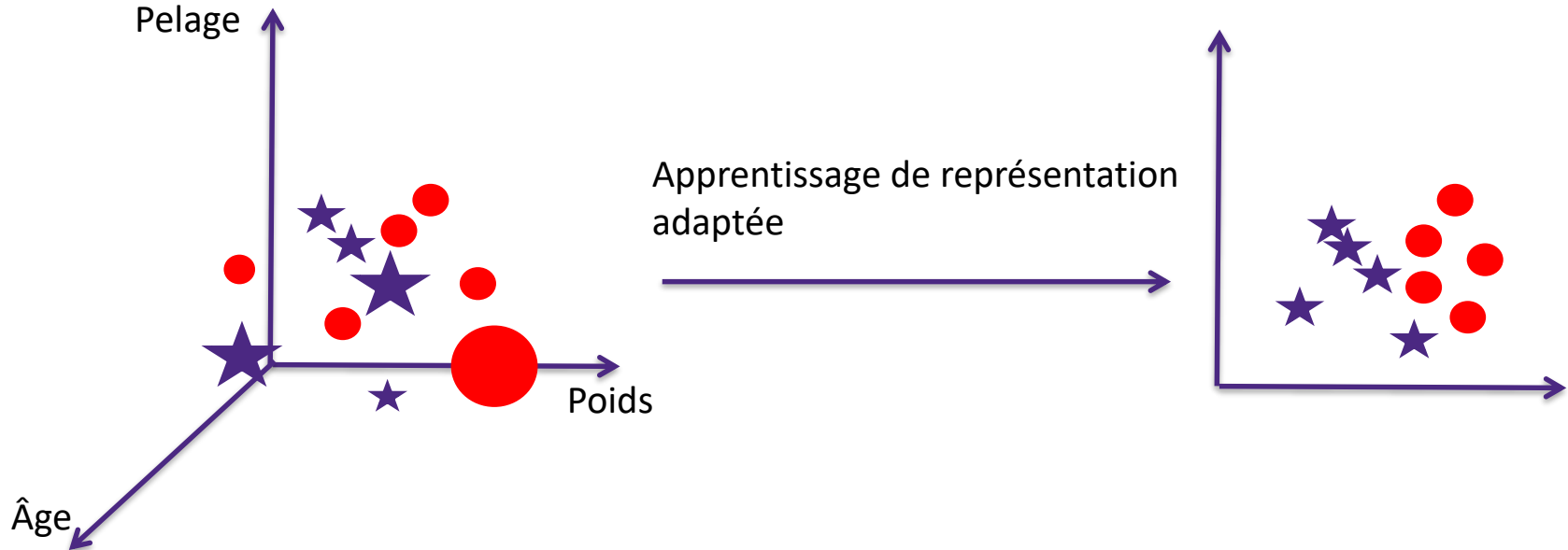
★ Utilisateurs

Montant mensuel dépensé



APPRENTISSAGE DE REPRÉSENTATIONS

★ Chien
● Chat



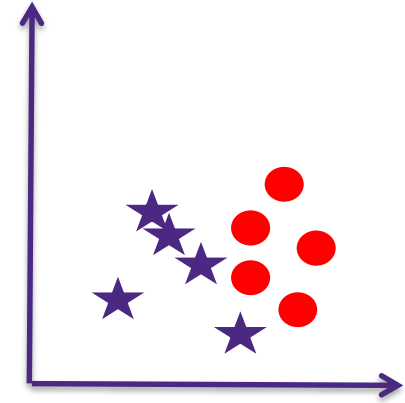
- **Beaucoup de dimensions**
 - Difficile de visualiser les données
- **Séparation difficile**

- Moins de dimensions
- Représentation facilitant la séparation

Apprentissage de représentation des données pour faciliter la représentation, le stockage ou le calcul sur les données.

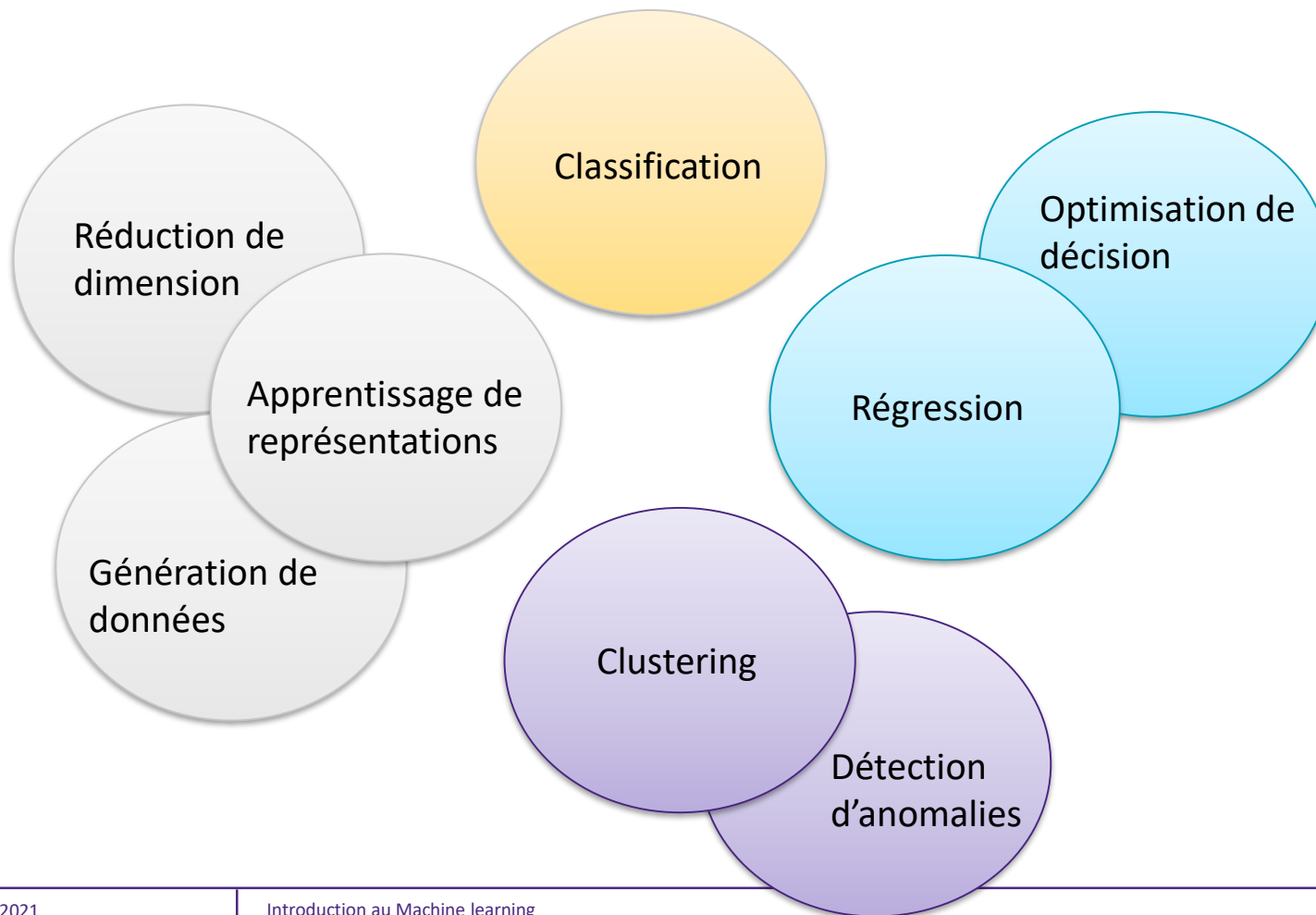
Exemples d'applications:

- **Encodage des mots pour le traitement de texte**
- **Réduction du nombre de dimension pour visualisation**
- **Réduction du nombre de dimension pour faciliter les calculs**
- **Générer de nouvelles données**



LES DIFFÉRENTES TÂCHES DU ML

Les principales tâches et leurs équivalents



LES DIFFÉRENTS TYPES D'APPRENTISSAGE

Il existe différents types d'apprentissage pour les algos de ML.

Le choix de celui-ci dépend du problème et des données.

Types
d'apprentissages les
plus utilisés

Apprentissage
supervisé

Apprentissage
non supervisé

LES DIFFÉRENTS TYPES D'APPRENTISSAGE

Il existe **différents types d'apprentissage** pour les algos de ML.

Le choix de celui-ci dépend **du problème et des données**.

Types
d'apprentissages les
plus utilisés

Apprentissage
supervisé

Apprentissage
non supervisé

Un peu moins
utilisé

Apprentissage par
renforcement

Apprentissage
semi supervisé

Apprentissage
actif

Beaucoup
moins utilisés

L'APPRENTISSAGE SUPERVISÉ

- Le modèle est entraîné en lui indiquant quels sont les **bonnes réponses sur les exemples utilisés pour son entraînement**:

Entraînement

Labels = bonnes réponses

Label Donnée 1	Verbe
Label Donnée 2	Nom
Label Donnée 3	Verbe

Exemples d'entrainement

Donnée 1	Manger
Donnée 2	Cuisine
Donnée 3	Boire
	⋮

Modèle

Prédictions

Input 1: Nom
Input 2: Verbe
Input 3: Verbe

Production

L'APPRENTISSAGE SUPERVISÉ

- Le modèle est entraîné en lui indiquant quels sont les **bonnes réponses sur les exemples utilisés pour son entraînement**:

Entraînement

Labels = bonnes réponses

Label Donnée 1	Verbe
Label Donnée 2	Nom
Label Donnée 3	Verbe

Exemples d'entrainement

Donnée 1	Manger
Donnée 2	Cuisine
Donnée 3	Boire
	⋮

Modèle

Prédictions

Input 1: Nom
Input 2: Verbe
Input 3: Verbe

Production

Données réelles

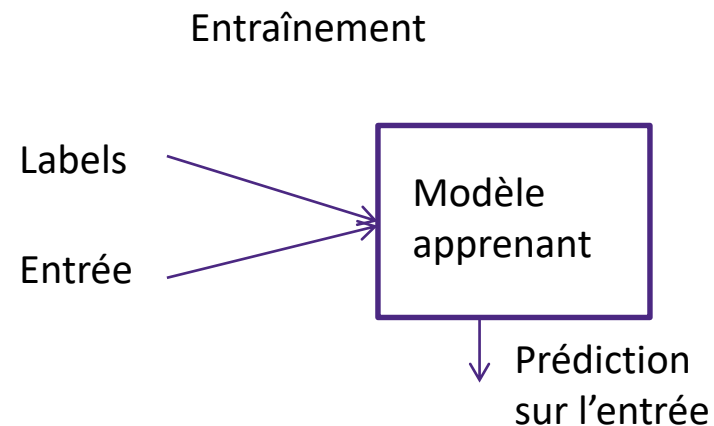
Donnée 1	Dancer
Donnée 2	Salon
Donnée 3	Marché
	⋮

Modèle
entraîné

Prédictions

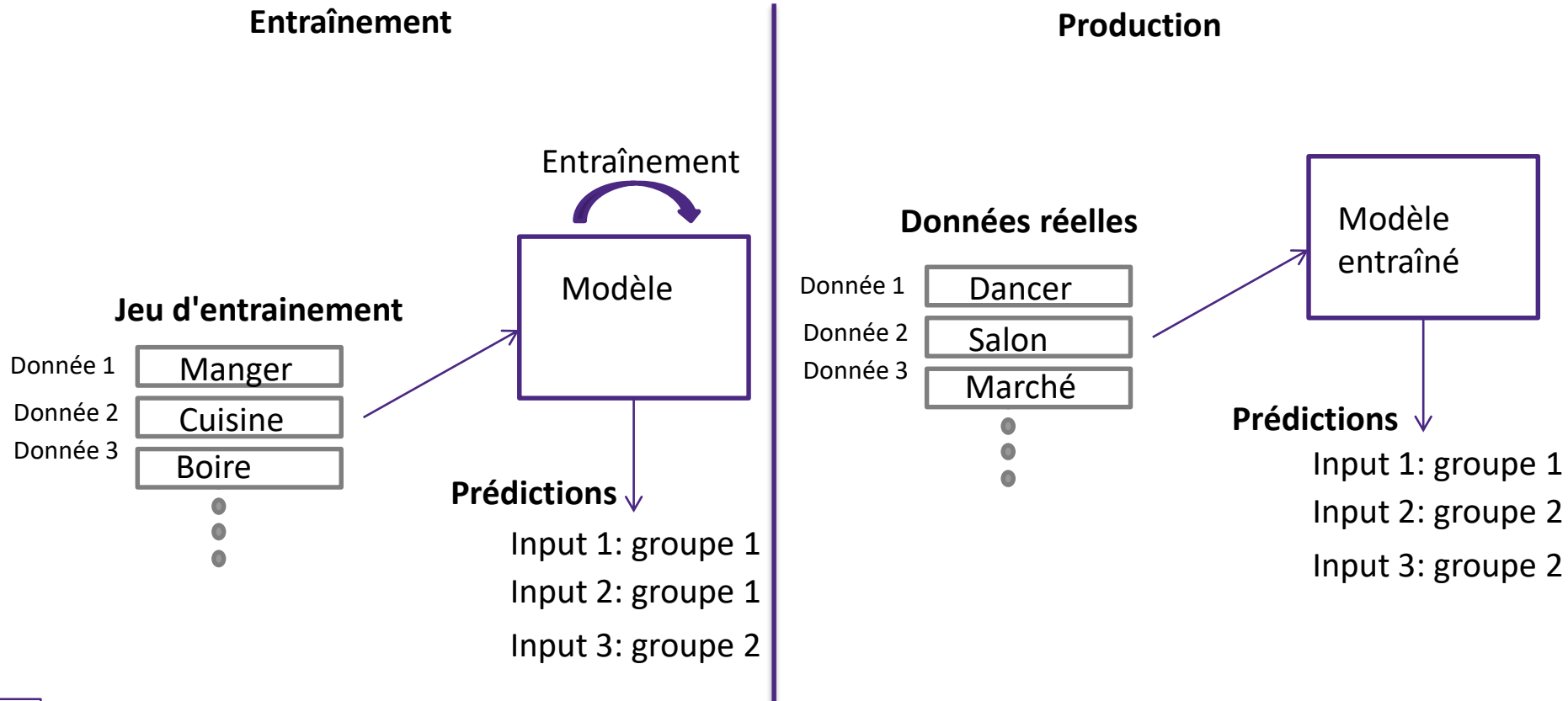
Input 1: Nom
Input 2: Verbe
Input 3: Nom

- **Utilisé principalement pour la classification et la régression car on connaît généralement le label à associer à chaque entrée**
- **Peu utilisé pour l'apprentissage de représentations et le clustering car on ne connaît en général pas la sortie optimale pour chaque entrée.**
- ✓ **Permet généralement un meilleur modèle entraîné**
- **Plus de contraintes** car nécessite de fournir la réponse désirée pour chaque exemple d'entraînement.
 - **Pas toujours possible**
 - **Possiblement coûteux** si cela doit être fait à la main par un spécialiste sur de gros volumes de données

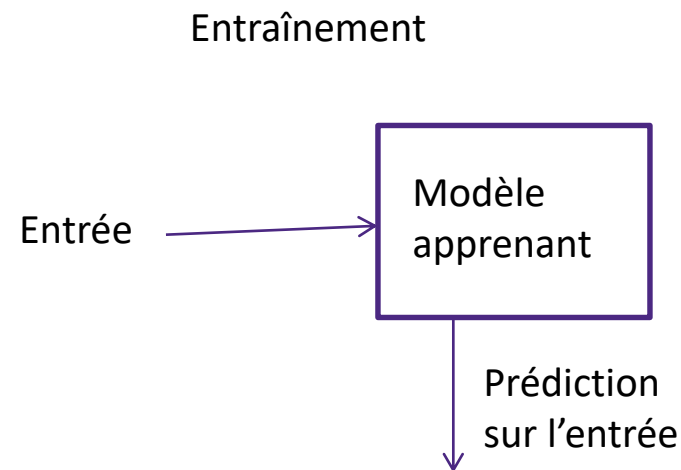


L'APPRENTISSAGE NON SUPERVISÉ

- Le modèle est entraîné sans recevoir de feedback sur ses prédictions :

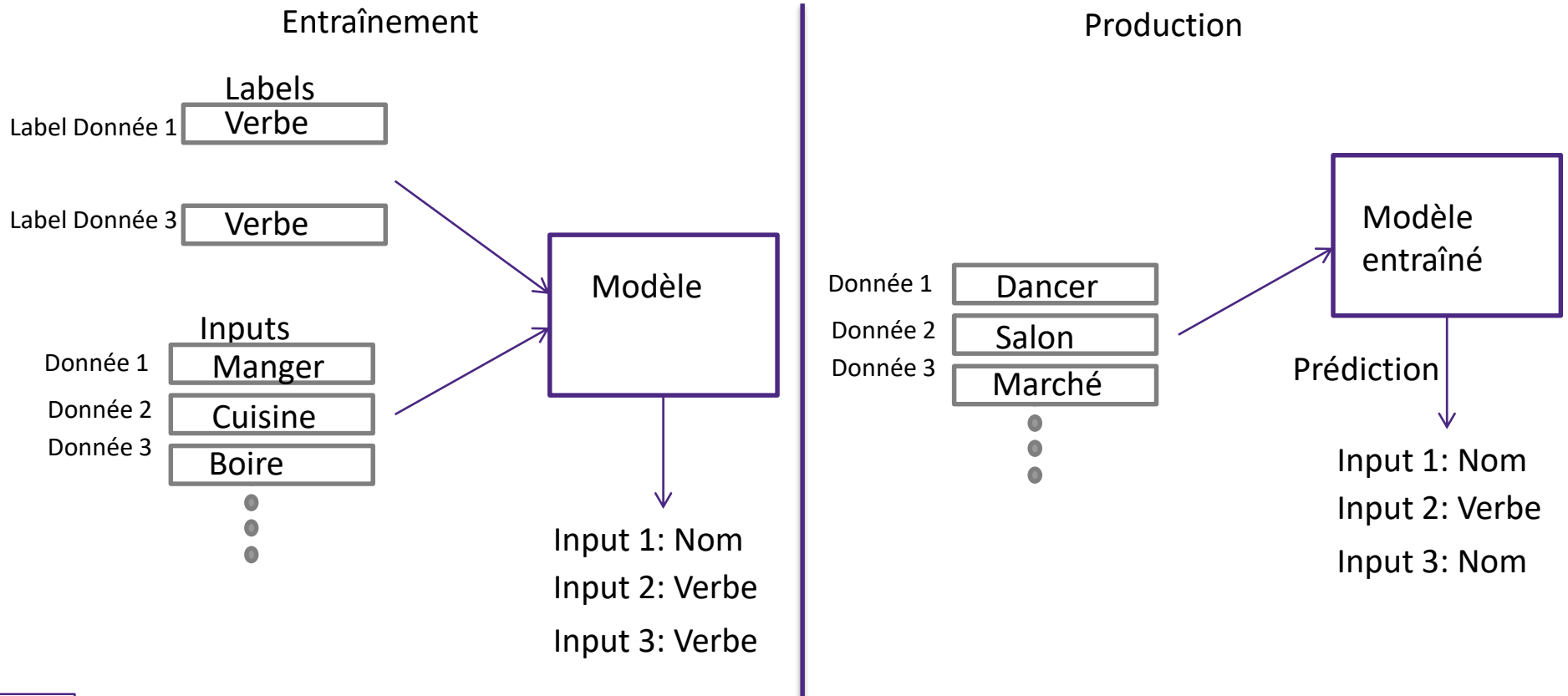


- Apprentissage adapté **quand on ne connaît pas la "bonne réponse"** pour aucun des exemples du jeu d'entraînement.
- **Très utilisé pour le clustering**
- **Une option pour la détection d'anomalie**
- **Très utilisé dans l'apprentissage de représentation**
- **Exemples :**
 - Trouver des clusters d'utilisateurs ou de communautés dans des réseaux sociaux
 - Représentation des mots telle que des synonymes aient une représentation similaire
 - Trouver des anomalies dans les relevés de capteurs embarqués

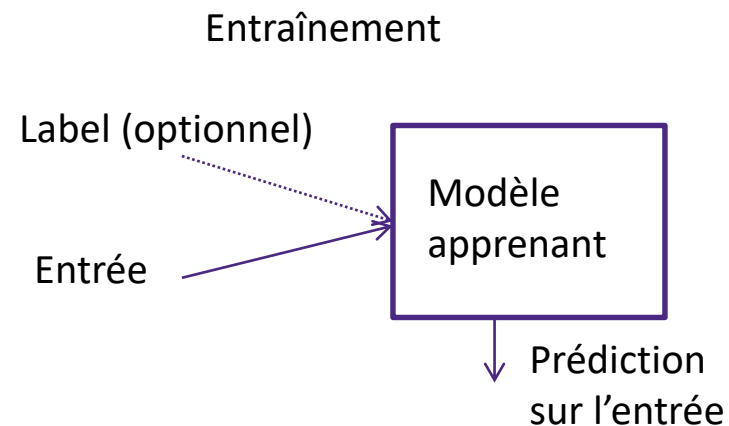


L'APPRENTISSAGE SEMI-SUPERVISÉ

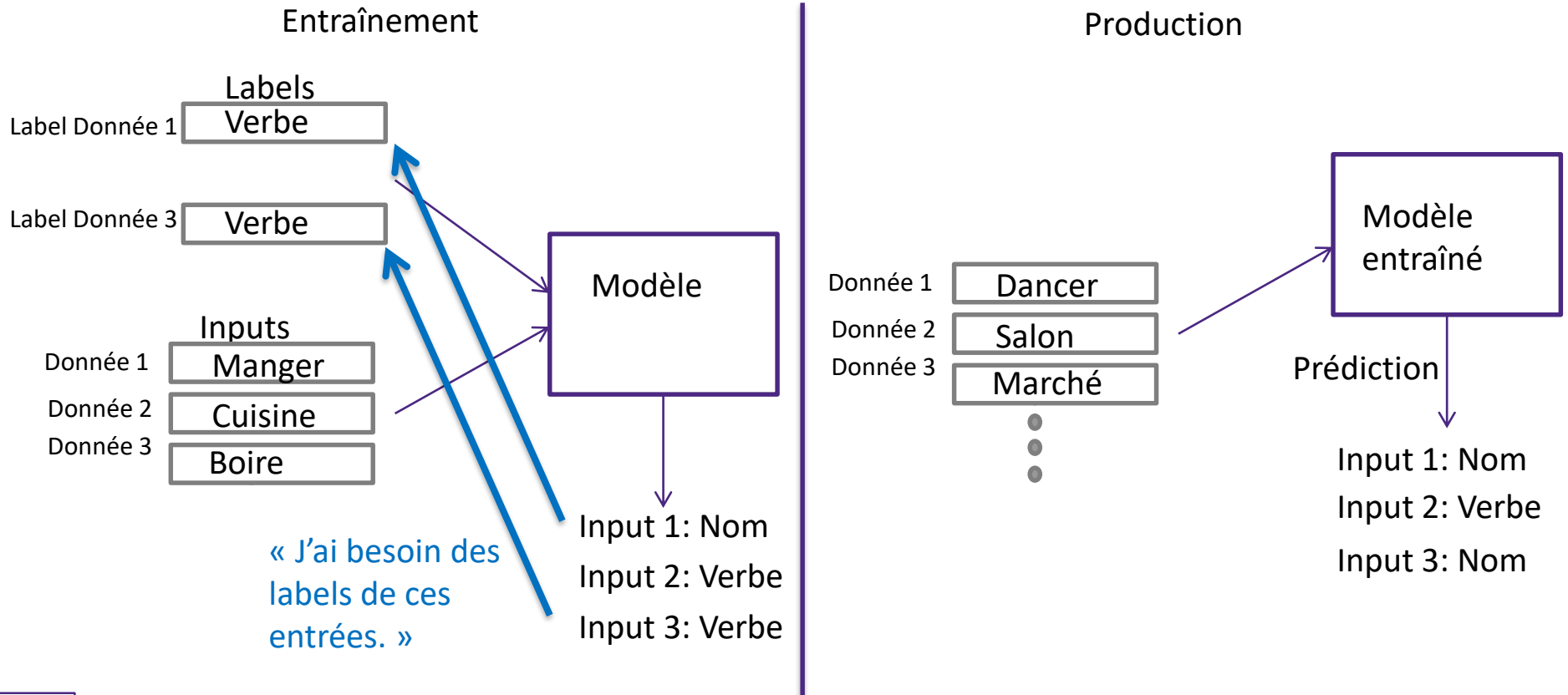
- Le modèle est entraîné à partir des labels d'une partie de la population :



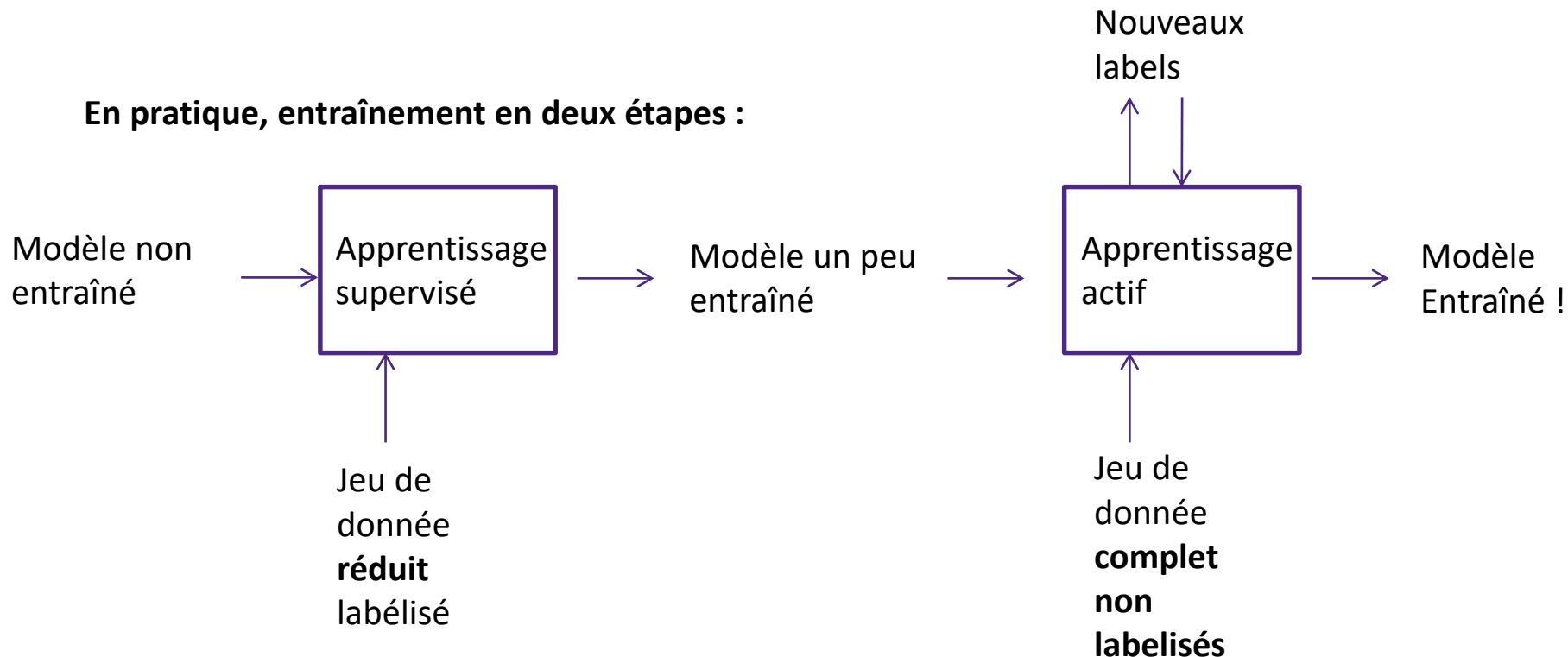
- **Apprentissage à mi chemin entre l'apprentissage supervisé et non supervisé.**
- **Utilisé sur des applications similaires à l'apprentissage non supervisés où l'on connait ou veut imposer la cible de certaines entrées.**
- **Ex : Grouper les utilisateurs de profils similaires en connaissant certains profils type**
- **Utilisé dans les tâches de classification où l'on n'est pas sur de connaître toutes les classes type.**



- Le modèle demande sur quelles entrées fournir des labels :

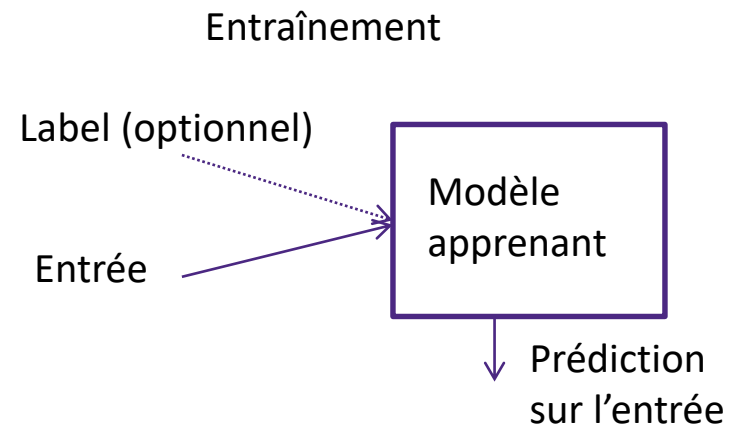


En pratique, entraînement en deux étapes :

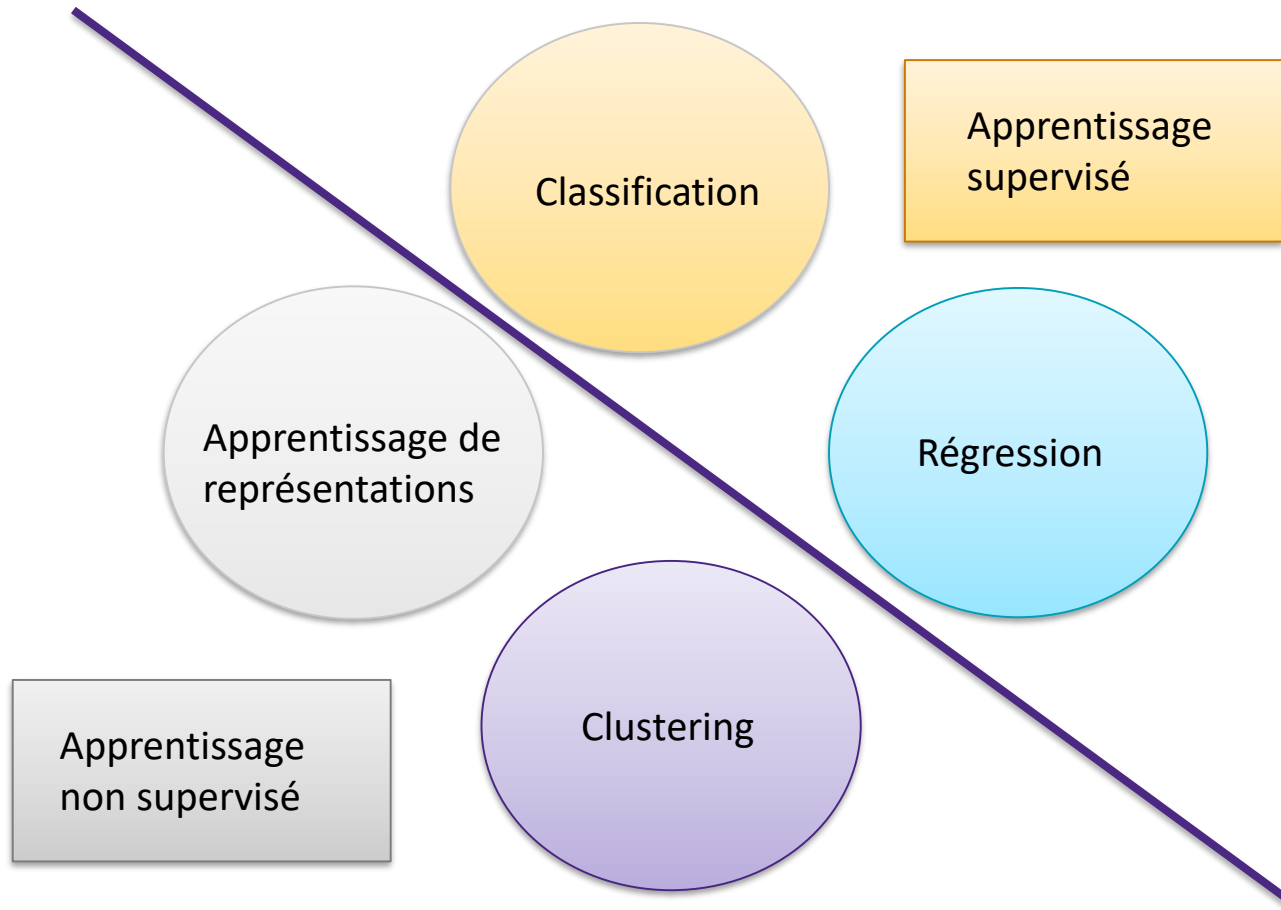


Permet d'améliorer le compromis taille du jeu de donnée/performance (90% des projets)

- **Fait parti de l'apprentissage semi-supervisé.**
- **Cherche à obtenir les labels des entrées les plus informatives pour maximiser l'apprentissage.**
- **Méthode récente et encore peu utilisée**
- **Utilisé quand accéder au label des entrées est trop coûteux pour être fait sur chaque entrée.**



CONCLUSION : DÉCOUPAGE SIMPLIFIÉ DU ML



SUPERVISÉ / NON SUPERVISÉ

QUIZZ

Supervisé

Non-Supervisé

Connaître l'âge de mon prochain acheteur

Identifier les relations entre mes clients

Connaître le genre (masculin féminin) de mon prochain acheteur

Regrouper mes utilisateurs par leur comportement

Détecter un achat à un prix anormal

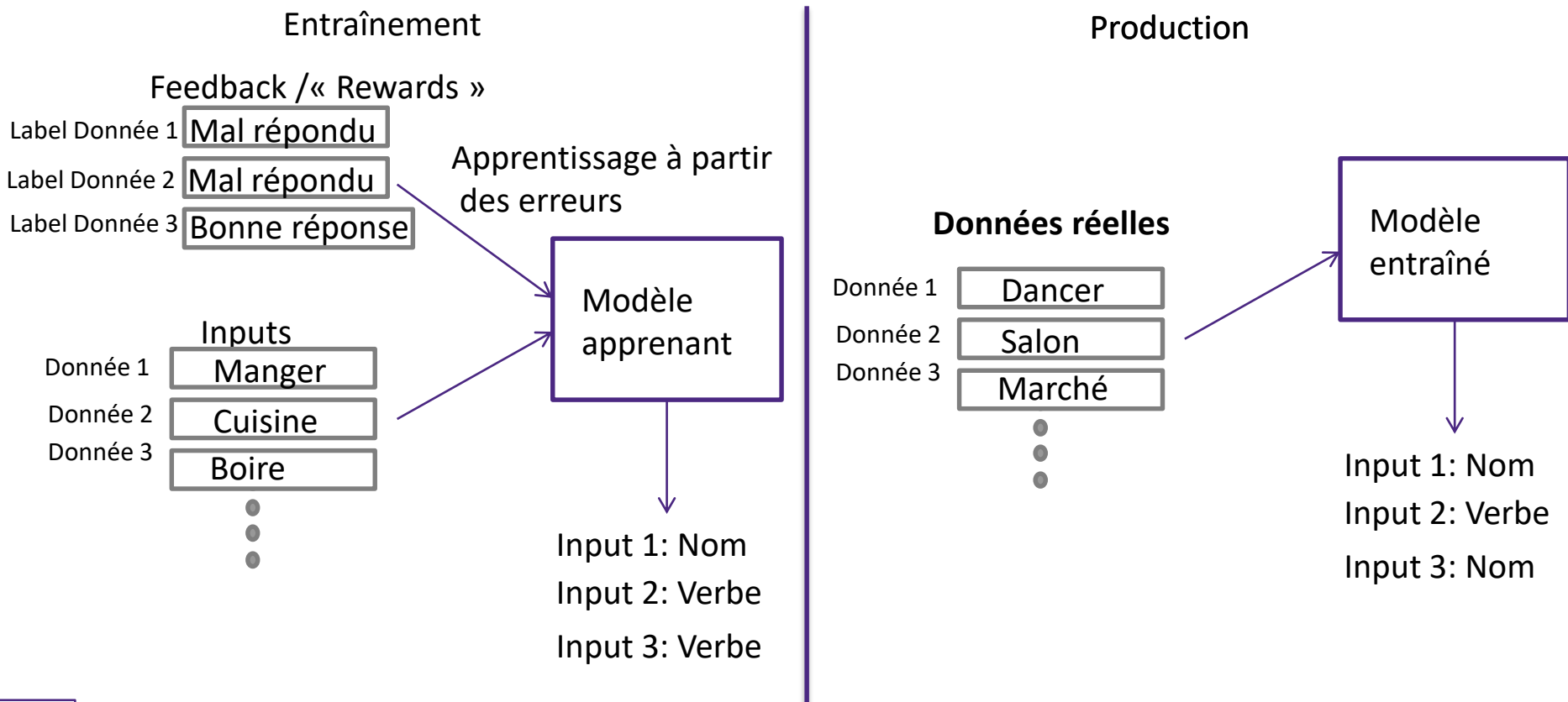
SUPERVISÉ / NON SUPERVISÉ

QUIZZ

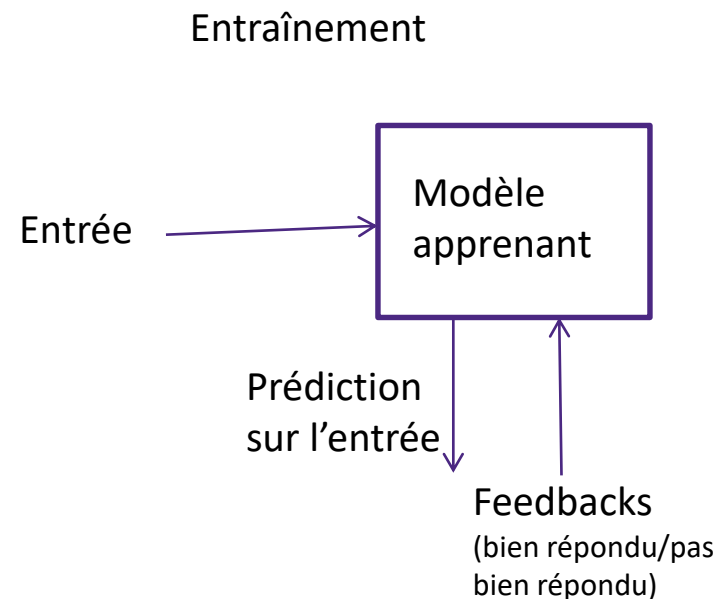
	Supervisé	Non-Supervisé
Connaître l'âge de mon prochain acheteur	✓	
Identifier les relations entre mes clients		✓
Connaître le genre (masculin féminin) de mon prochain acheteur	✓	
Regrouper mes utilisateurs par leur comportement		✓
Détecter un achat à un prix anormal	✓	✓

L'APPRENTISSAGE PAR RENFORCEMENT

- Le modèle est entraîné en connaissant s'il a bien répondu ou pas:



- **Apprentissage compliqué : le modèle ne connaît pas la bonne réponse lorsqu'il se trompe dans son entraînement.**
- **Souvent utilisé pour des systèmes apprenant en continu ou en interaction avec les utilisateurs**
- **Dès que l'on ne connaît pas la solution optimale, mais qu'on est capable d'évaluer sa qualité.**
- **Exemple : apprendre à jouer aux échecs !**
- **Le Feedback ou « Reward » peut incorporer de la connaissance d'experts pour guider le modèle.**



OU UTILISER LE RENFORCEMENT ?

QUIZZ

Renforcement

Prédire si un utilisateur va cliquer sur un bouton

**Apprendre à prédire quand le cours de la bourse
baisse**

Apprendre à un robot à se déplacer

Prédire en temps réel la météo

Choisir les publicités à proposer sur un site web

OU UTILISER LE RENFORCEMENT ?

QUIZZ

Renforcement

Prédire si un utilisateur va cliquer sur un bouton



Apprendre à prédire quand le cours de la bourse baisse



Apprendre à un robot à se déplacer



Prédire en temps réel la météo



Choisir les publicités à proposer sur un site web



OÙ EST UTILISÉ LE RENFORCEMENT ?

- On va faire le choix d'utiliser le renforcement quand **on ne sait pas donner la réponse optimale lors de l'entraînement**

En **robotique** :

- **Faire marcher/déplacer/effectuer une action à un robot**
- On sait dire si une action est réussie ou pas, mais pas comment l'effectuer
- Si une action est réussie, on ne sait pas si elle était la plus optimale ou pas

OÙ EST UTILISÉ LE RENFORCEMENT ?

- On va faire le choix d'utiliser le renforcement quand **on ne sait pas donner la réponse optimale lors de l'entraînement**

En interagissant avec le monde extérieur :

- **Ex: proposer la meilleur publicité pour faire acheter un utilisateur.**
- Lorsque la réponse du modèle doit susciter un changement de l'environnement ou une action de l'utilisateur
- On ne sait pas modéliser la réponse de l'environnement/ de l'utilisateur à l'action.
- On sait dire quand la réponse de l'environnement/ de l'utilisateur est positive.

OÙ EST UTILISÉ LE RENFORCEMENT ?

- On va faire le choix d'utiliser le renforcement quand on ne sait pas donner la réponse optimale lors de l'entraînement **en un temps suffisant**

Jeux et problèmes complexes :

- **Jeu de go : il y a trop de combinaisons pour être toutes calculées.**
- **On ne sait pas dire si un coup est optimal.**
- **Si on perd, on ne sait généralement pas dire qu'est-ce qu'il aurait fallu jouer pour gagner.**
- Par contre on sait si on gagne ou pas le match à la fin !

OÙ EST UTILISÉ LE RENFORCEMENT ?

- On va faire le choix d'utiliser le renforcement quand on ne sait pas donner la réponse optimale lors de l'entraînement **en un temps suffisant**

Pour résoudre un problème **NP-complet**:

- **NP-complet** : « que l'on ne sait pas résoudre sans tester toutes les possibilités » (ce qui est exponentiel en temps).
- **Souvent des problèmes d'optimisation.**
- **Ex de problème NP complet :**
 - *Le voyageur de commerce : quel est l'itinéraire le plus optimal pour visiter un ensemble de villes ?*
 - *Le problème du sac à dos : comment remplir de manière optimale un sac à dos ?*
 - *Dire si une formule logique est satisfaisable (admet une solution)*
 - *Clique maximale dans un graphe*
 - *Colorier un graphe tel que deux sommets n'aient pas la même couleur*
 - ...
- **Tout nos services de sécurité (bancaire, chiffrement, ...) se basent sur le fait que chiffrer/déchiffrer avec la clef est facile, mais décrypter sans la clef est NP-complet.**

OÙ EST UTILISÉ LE RENFORCEMENT ?

- On va faire le choix d'utiliser le renforcement quand on ne sait pas donner la réponse optimale lors de l'entraînement **en un temps suffisant**

Pour résoudre un problème **NP-complet**:

- **NP-complet** : « que l'on ne sait pas résoudre sans tester toutes les possibilités ».
- **A votre avis, quel est le problème NP-complet le plus critique chez Google ?**

OÙ EST UTILISÉ LE RENFORCEMENT ?

- On va faire le choix d'utiliser le renforcement quand on ne sait pas donner la réponse optimale lors de l'entraînement **en un temps suffisant**

Pour résoudre un problème **NP-complet**:

- **NP-complet** : « que l'on ne sait pas résoudre sans tester toutes les possibilités ».
- **A votre avis, quel est le problème NP-complet le plus critique chez Google ?**

Réponse : Le problème du sac à dos !

Comment répartir toute la charge de travail dans les datacenters et machine de la façon la plus optimale ?

Même un tout petit gain sur l'efficacité de la gestion peut apporter d'énormes économies vu le volume.

OÙ EST UTILISÉ LE RENFORCEMENT ?

- On va faire le choix d'utiliser le renforcement quand on ne sait pas donner la réponse optimale lors de l'entraînement **en un temps suffisant**

En informatique :

- **Pour fournir une solution performante à un problème NP-complet**
- **On saurait calculer la solution optimale mais pas en un temps raisonnable.**
- Par contre évaluer la qualité de la réponse est faisable.
- Concurrence avec la Recherche Opérationnelle, qui propose des méthodes et heuristique pour donner des solutions performantes en un temps raisonnable.
- Google : RO et RL ne sont pas forcément incompatibles mais pourraient se compléter dans des approches mixtes à l'avenir.

Pour vous rendre à Ynov ce matin, vous avez probablement choisi le « meilleur » itinéraire.

- Comment êtes-vous sûrs qu'un autre itinéraire n'est pas meilleur, ou ne l'est pas devenu depuis la dernière fois ?
- Que faites vous lorsque la qualité de votre « meilleur » itinéraire se détériore ?

C'est ce que l'on appelle le **dilemme exploration/exploitation**. Comment choisir entre :

- Privilégier la meilleure stratégie connue : **exploitation**
- Chercher une stratégie encore meilleure (qui peut ne pas exister) ou vérifier que les solutions alternatives ne se sont pas améliorées : **exploration**

Il s'agit d'un des problèmes les plus importants en RL

Lorsque ce que vous **gagnez une partie** (échecs, go, starcraft, lol, ...) , comment savez vous quel(les) coup(s) vous ont fait gagné la partie ?

- En RL, la récompense est souvent **différée** : elle n'arrive qu'après un nombre d'action qui peut être important.
- Comment déterminer quel(s) sont les actions qui vous ont permis de toucher votre récompense ?
- **L'attribution de la récompense** est une problématique récurrente en RL.

L'espace de recherche :

- Au échecs, il n'y a qu'un nombre finit d'action possibles chaque tour et de positions des pièces sur le plateau. On parle alors d'**espace d'état et d'action discrets**.
- Mais pour déplacer un robot dans une salle: il y a virtuellement un nombre d'action et de position infini. On parle alors d'**espace d'action et d'état continus**.

Ce ne sont généralement **pas les mêmes algorithmes** qui sont employés dans les deux cas.

Une dernière problématique commune au ML, **l'équilibre de l'apprentissage**.

- En ML, un jeu de donnée **déséquilibré ou non mélangé** peut mener à un mauvais entraînement.
- Ce problème est **encore plus important** en RL.
- Comment ne pas oublier les parties jouées il y a longtemps ?
- Comment ne pas être biaisé par le fait de jouer versus le même adversaire plusieurs fois de suite ?