

Introduction au Machine Learning avec Python

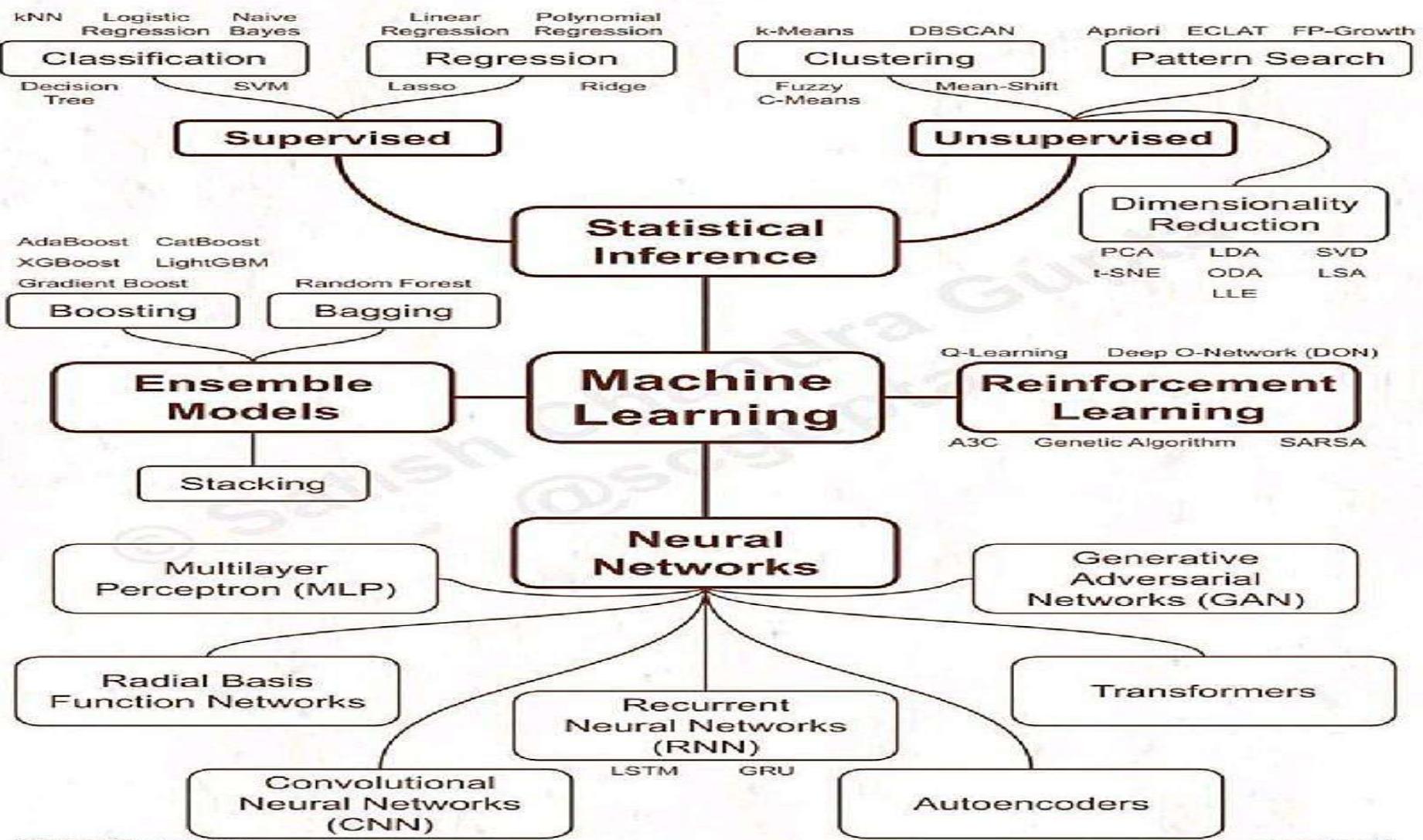
Niveau: Ing2_GL

AU: 2025/2026

Présenté par Dr.Manel SEKMA

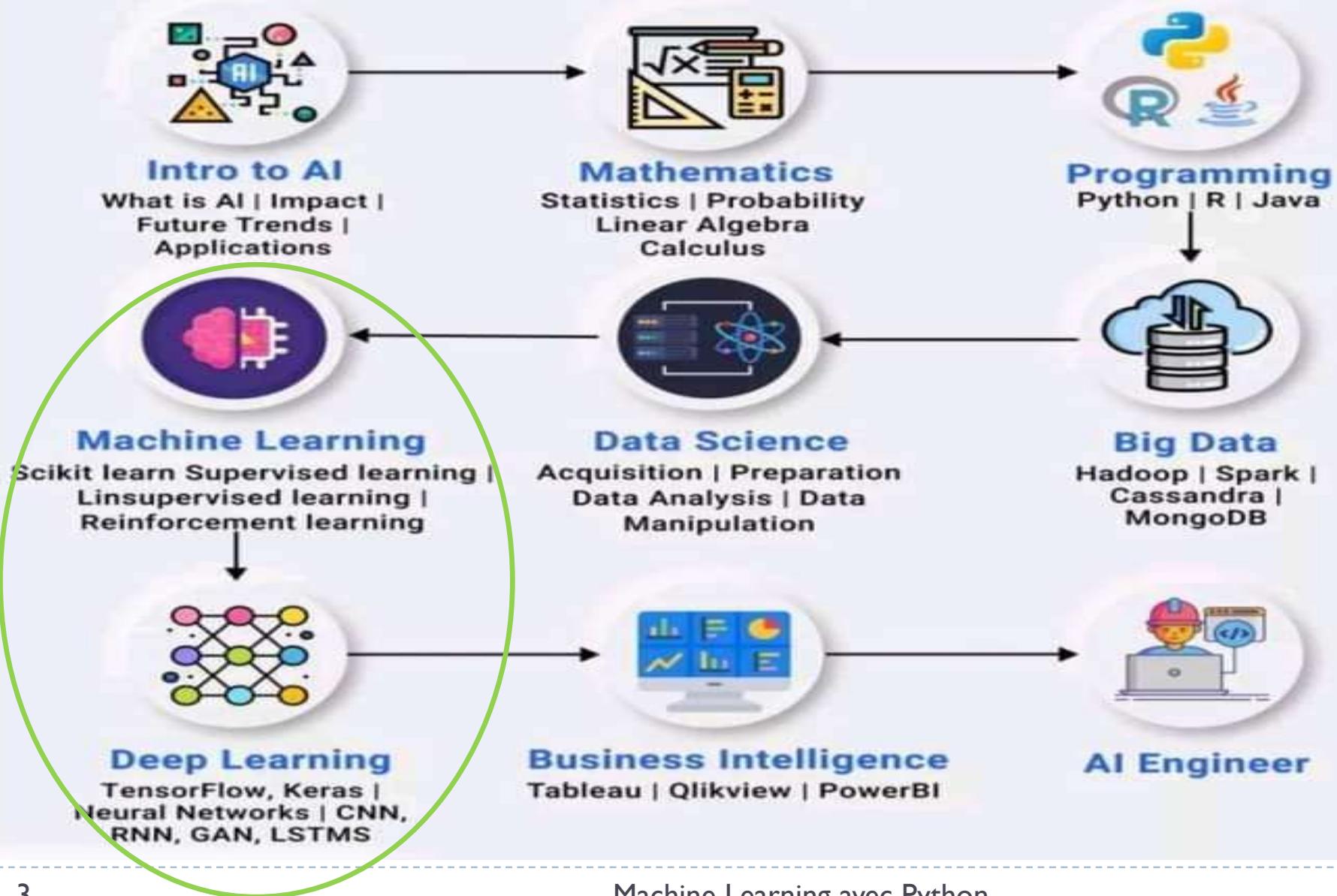
Popular Machine Learning Models

ml4devs.com/ml-intro 



or

AI ENGINEER ROADMAP



Introduction: C'est quoi le Machine Learning ?

Qu'est-ce que le Machine Learning, en quelques mots ?

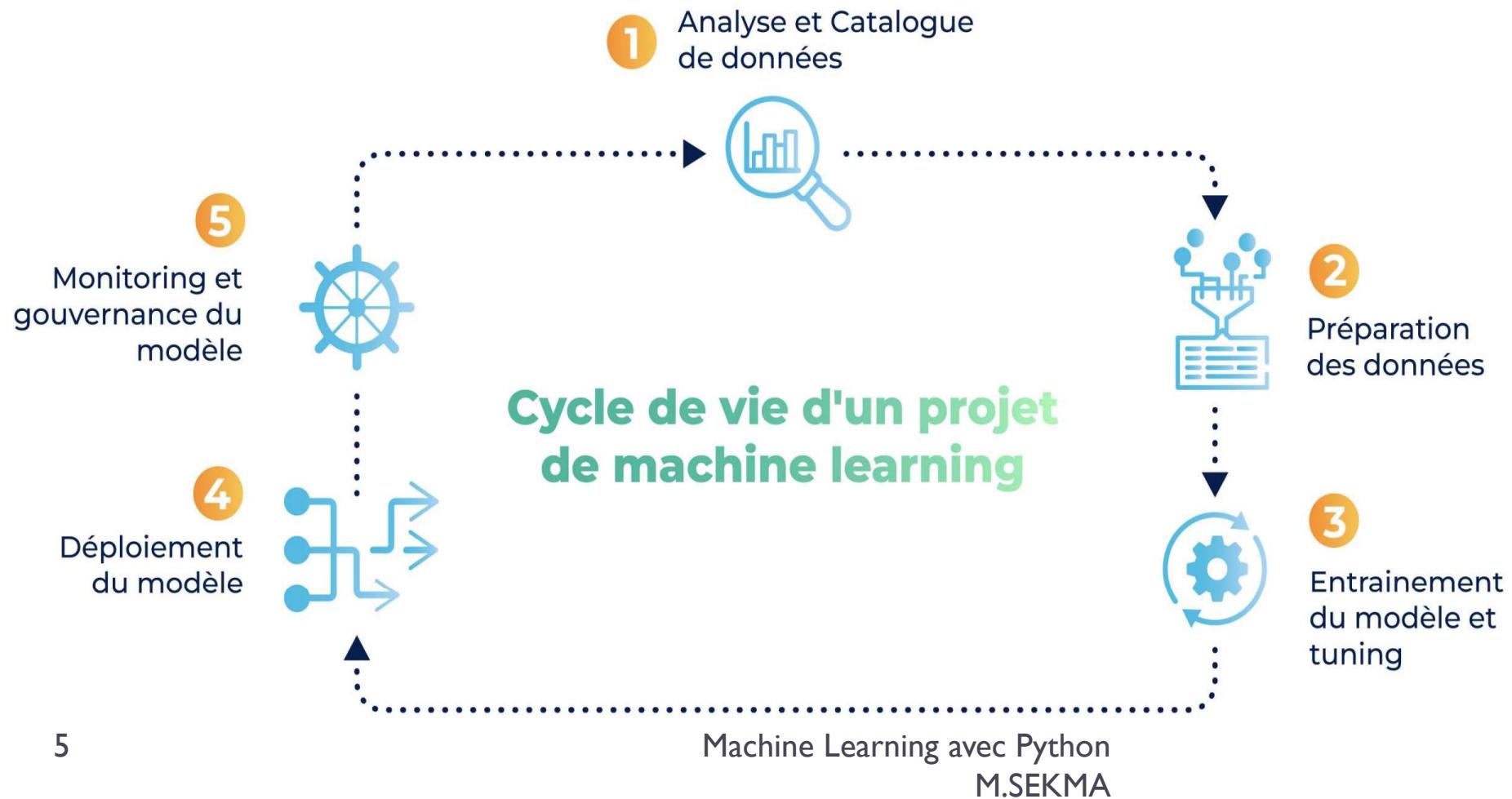
→ ML est une branche de l'intelligence artificielle qui se concentre sur le développement de systèmes informatiques capables d'apprendre à partir de données passées.

Quelle est la principale différence entre le Machine Learning et la programmation traditionnelle ?

→ Dans la programmation traditionnelle, les développeurs écrivent explicitement des instructions et des règles pour que l'ordinateur exécute des tâches spécifiques. En revanche, en ML, les algorithmes sont conçus pour apprendre à partir de données, plutôt que de suivre des instructions strictes.

Introduction: C'est quoi le Machine Learning ?

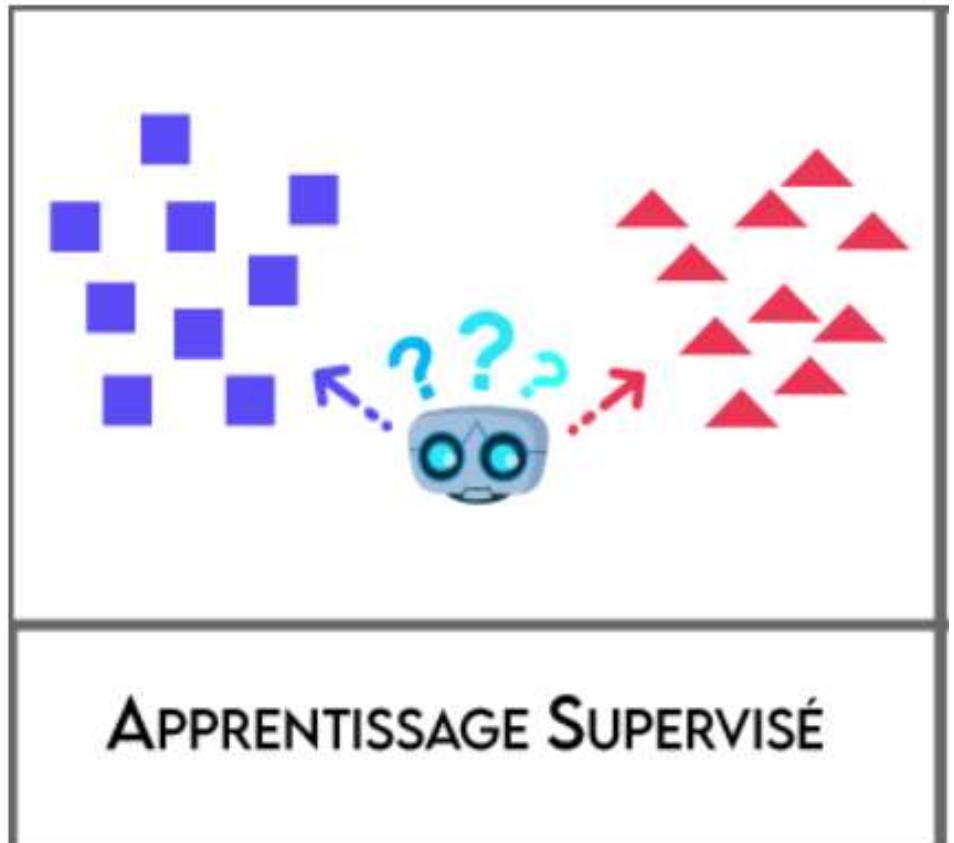
Quelles sont les étapes principales d'un projet de machine Learning, de la collecte des données à la mise en production du modèle ?



Introduction

I. Apprentissage supervisé

- ▶ La tâche est **d'expliquer** ou de **prédire** la valeur d'une certaine variable réponse Y en fonction de variables explicatives X.
- ▶ On cherche à classifier, prédire, estimer.



Apprentissage supervisé

- ▶ les données d'entraînement fournies aux machines fonctionnent comme le **superviseur** qui apprend aux machines à **prédirer correctement la sortie**.
- ▶ On a des variables d'entrée x et une variable de sortie y et on utilise un algorithme pour apprendre la fonction de mappage de l'entrée à la sortie.

$$y = f(x)$$

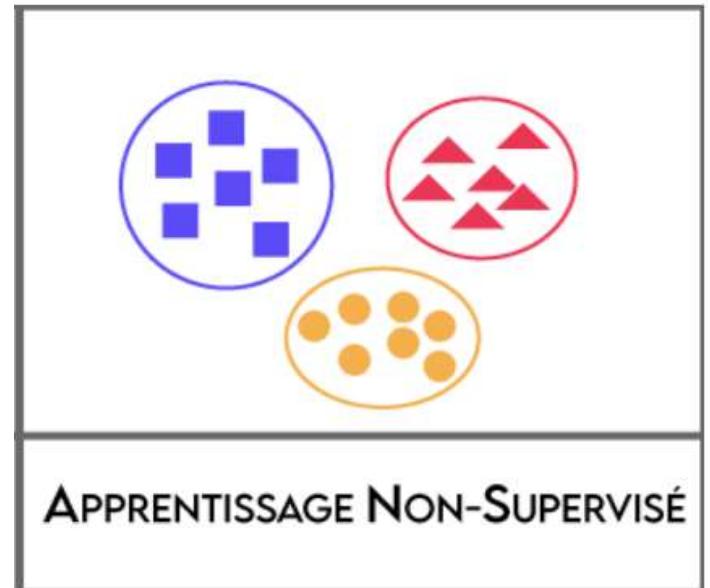
Dans l'apprentissage supervisé, on a deux types d'algorithmes :

- ▶ **Les algorithmes de régression**, qui cherchent à prédirer une valeur continue, une quantité.
- ▶ **Les algorithmes de classification**, qui cherchent à prédirer une classe/catégorie.

Introduction

2. Apprentissage non supervisé

- ▶ On s'intéresse à **identifier** des relations, associations ou groupements significatifs.
- ▶ On s'intéresse à la **distribution conjointe** des X.



- ▶ Il n'y a pas de variable réponse.
- ▶ On cherche à faire des **groupements** (clustering), à **segmenter** ou à **décrire**

Apprentissage non supervisé

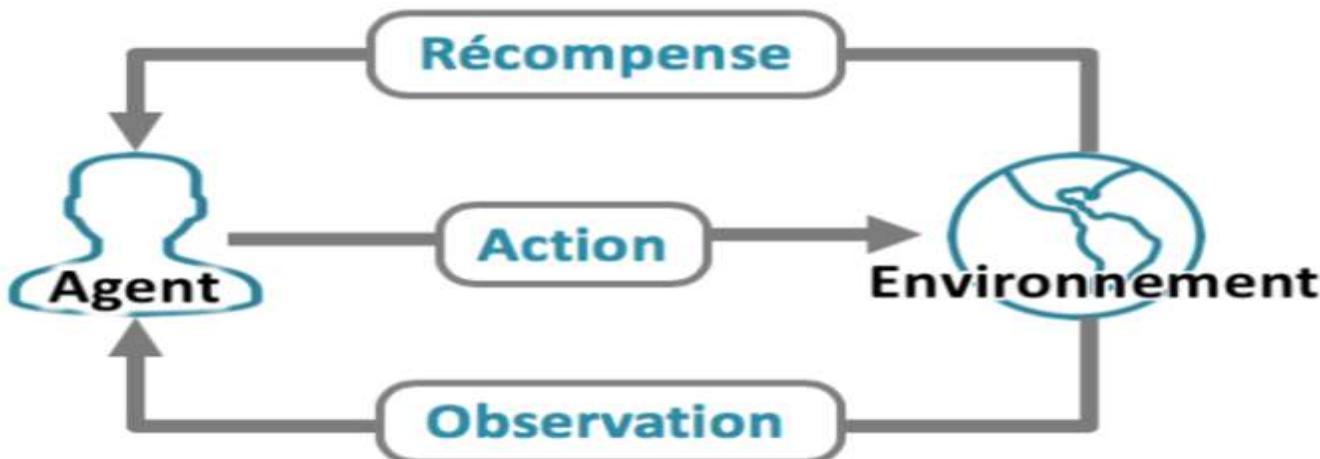
Les algorithmes de ce type d'apprentissage peuvent être utilisés pour trois types en problèmes.

- ▶ **La réduction de dimension** : on vise à réduire le nombre de variables à prendre en compte dans l'analyse.
- ▶ **Association** : un problème où on désire découvrir des règles qui décrivent de grandes portions de ses données
 - ▶ par exemple, dans un contexte d'une étude de comportement d'achat d'un groupe de clients, les personnes qui achètent tel produit ont également tendance à acheter un autre produit spécifique.
- ▶ **Regroupement** : un problème où on veut découvrir les groupements inhérents aux données, comme le regroupement des clients par le comportement d'achat.

Introduction

3. Apprentissage par renforcement

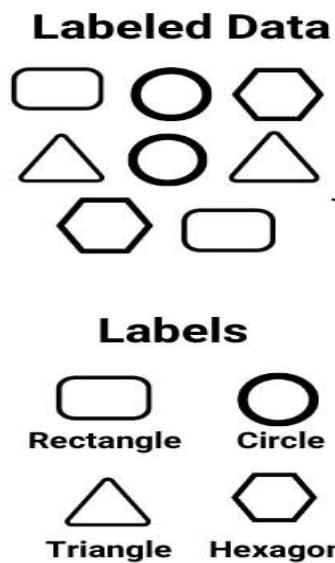
- ▶ l'apprentissage se fait par **essai et erreur**.



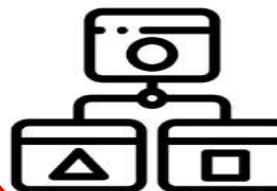
$$\max_{\pi} \sum_t R_t$$

- ▶ Lors de l'entraînement de ce type de modèle, une **réussite** obtient une **récompense** et, pour simplifier, le modèle cherche la solution qui maximise les chances d'obtenir une récompense

Type d'apprentissage?



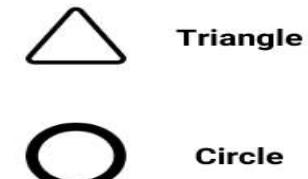
ML Model



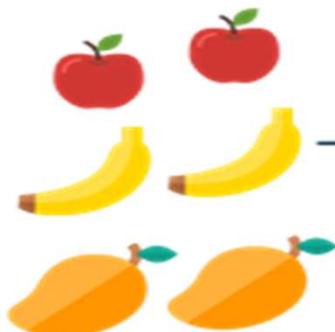
Test Data
Triangle Circle

Apprentissage supervisé,

Predictions



INPUT RAW DATA



- Unknown Output
- No Training Data Set

Interpretation

Algorithm



Processing

Model Training

Apprentissage Non supervisé,

OUTPUT



Model Trained

C'est quoi le Machine learning

Qu'est-ce qu'un jeu de données d'apprentissage (training dataset) et à quoi sert-il dans le machine Learning ?

- est un utilisé pour entraîner un modèle de machine Learning.
- Il est composé de paires de données d'entrée et de leurs valeurs cibles (ou étiquettes) associées.
- Le modèle apprend à partir de ces exemples en identifiant des schémas et des relations entre les données d'entrée et les valeurs cibles.

C'est quoi le Machine learning

- ▶ Qu'est-ce qu'un **modèle** de machine Learning, et comment est-il utilisé pour effectuer des prédictions ?
 - est un algorithme ou un ensemble d'algorithmes qui apprend à partir de données pour effectuer des prédictions ou prendre des décisions.
 - Il est basé sur des paramètres ajustables qui sont optimisés pendant le processus d'entraînement pour rendre les prédictions aussi précises que possible.

Quiz: déterminer si une application demande un apprentissage supervisé ou non supervisé.

1. Vous avez un ensemble de photos d'animaux et vous souhaitez classer automatiquement chaque photo en fonction du type d'animal (chien, chat, oiseau, etc.).
2. Vous avez un ensemble de commentaires de clients sur votre site web et vous souhaitez identifier les sujets ou les thèmes les plus fréquemment abordés sans avoir de catégories prédéfinies.
3. Vous collectez des données sur le comportement des utilisateurs d'une application mobile et vous voulez prédire si un utilisateur est susceptible de se désabonner de votre service.
4. Vous avez un grand ensemble de données clients contenant des informations telles que l'âge, le revenu, le lieu de résidence, etc., et vous souhaitez regrouper les clients en segments similaires pour des campagnes marketing ciblées.

Quiz: déterminer si une application demande un apprentissage supervisé ou non supervisé.

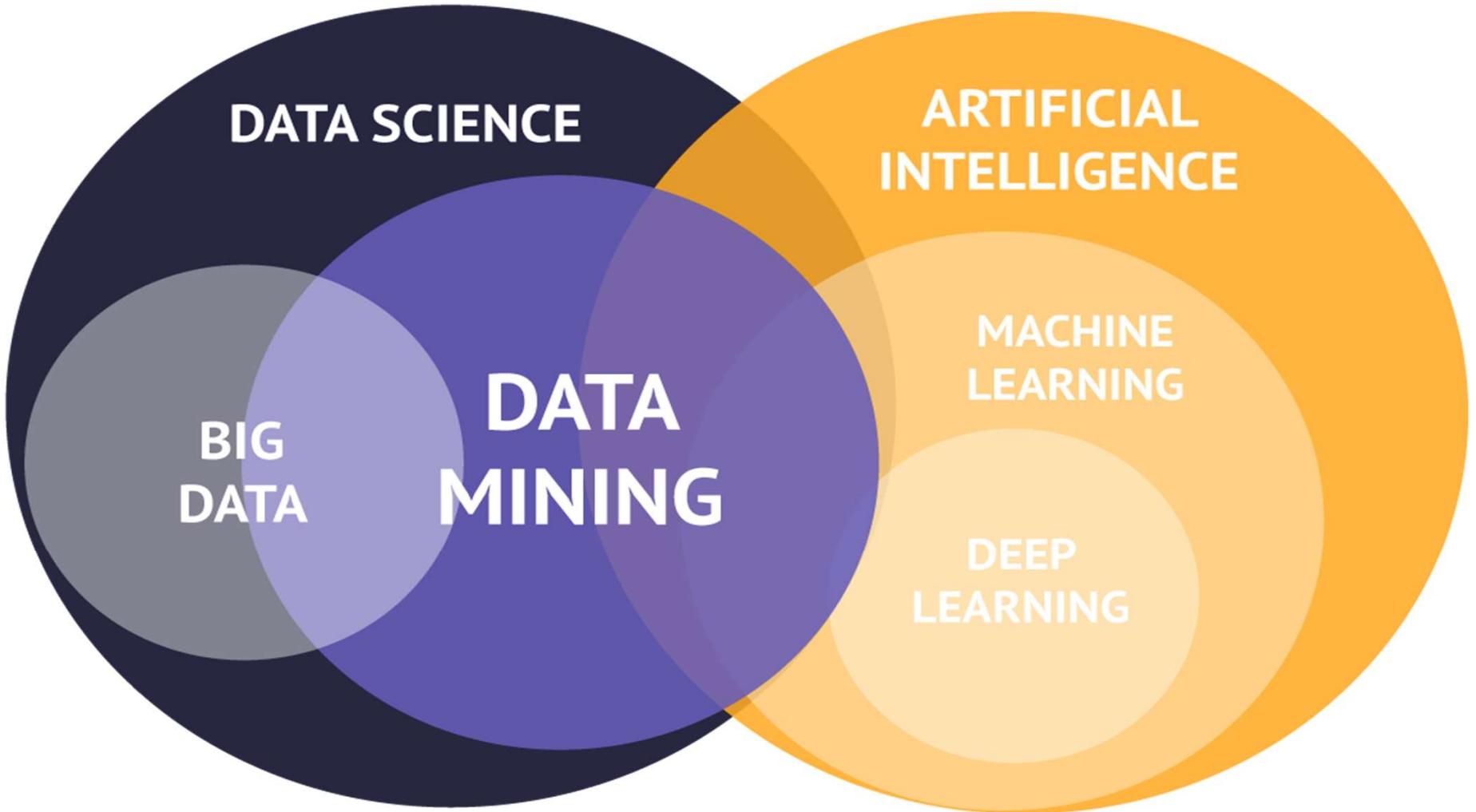
5. Vous voulez construire un système de recommandation pour un site de streaming vidéo qui suggère des films aux utilisateurs en fonction de leurs préférences passées et du comportement de visionnage.
6. Vous avez un ensemble de documents juridiques et vous souhaitez identifier automatiquement les catégories juridiques auxquelles ils appartiennent (par exemple, contrats, litiges, brevets, etc.)
7. Vous analysez des données de vente pour détecter les tendances saisonnières ou les modèles de consommation sans avoir de catégories prédéfinies.
8. Vous voulez créer un *chatbot* qui peut répondre aux questions des clients en utilisant une base de connaissances préexistante et en apprenant à partir des interactions passées.

Introduction

Quel est le principe du machine learning alors ?

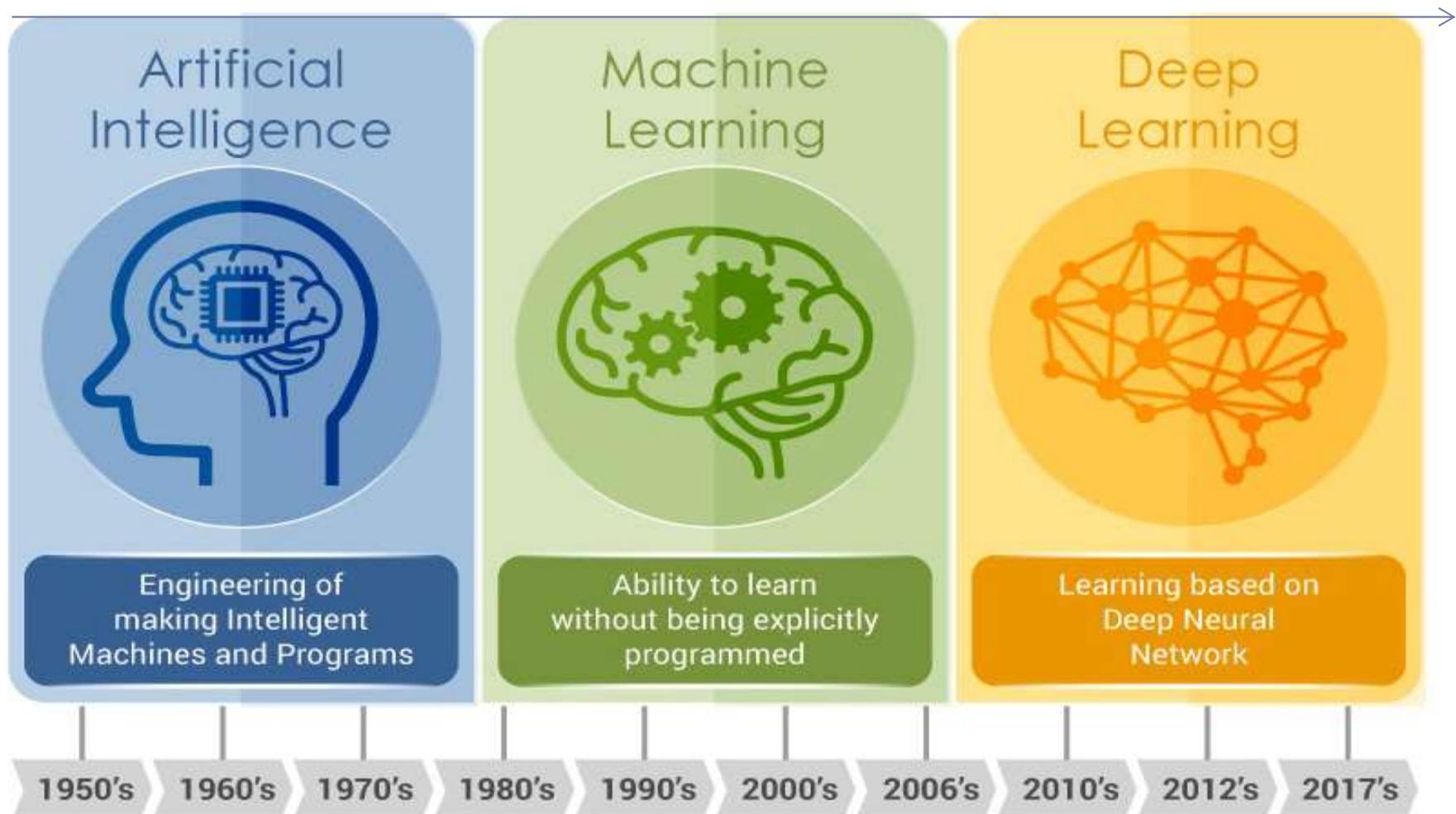
- ▶ Le machine Learning (ML) est une forme d'intelligence artificielle (IA) qui est axée sur la création de systèmes qui apprennent, ou améliorent leurs performances, en fonction des données qu'ils traitent.
- ▶ Le ML se base sur le forage de données (data mining), permettant la reconnaissance de patterns pour fournir des analyses prédictives.
- ▶ Le ML révèle tout son potentiel dans les situations où des insights (tendances) doivent être repérés à partir de vastes ensembles de données diverses et variées, appelés le Big Data.

Introduction



Introduction

Les dates clés



Rappel

Régression vs classification

- ▶ Un algorithme de **régression** peut prédire une **valeur discrète** qui se présente sous la forme d'une **quantité entière**.
- ▶ Un algorithme de **classification** peut prédire **une valeur continue** si elle se présente sous la forme d'une **probabilité d'étiquette de classe**.



Regression

What is the temperature going to be tomorrow?



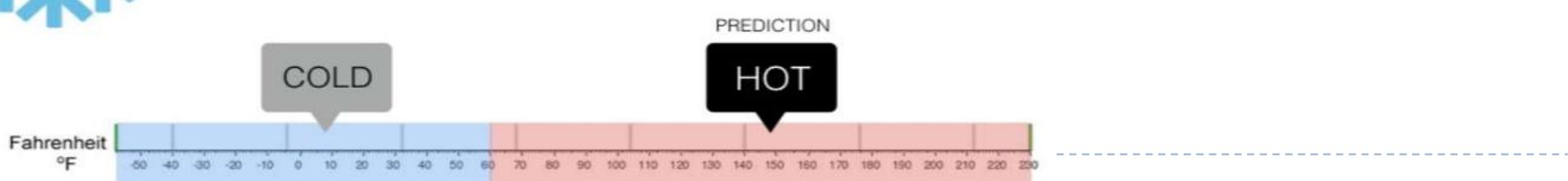
Classification

Will it be Cold or Hot tomorrow?



COLD

PREDICTION
HOT



Quiz: deviner en fonction des informations fournies si ces applications nécessitent une classification ou une régression

1. Prédiction de la note d'un film:

L'objectif est de prédire la note d'un film en fonction de diverses caractéristiques telles que le budget, le genre, la durée, etc.

2. Détection de spam dans les e-mails L'objectif est de déterminer si un e-mail est du spam (oui) ou non (non) en se basant sur son contenu et ses caractéristiques.

3. Prédiction du prix de vente d'une maison

L'objectif est de prédire le prix de vente d'une maison en fonction de ses caractéristiques, telles que la superficie, le nombre de chambres, l'emplacement, etc

4. Classification des espèces d'oiseaux L'objectif est de classer les espèces d'oiseaux en fonction de leurs caractéristiques physiques, telles que la taille, la couleur, la forme du bec, etc

Quiz (suite):

5. Prédiction du nombre de ventes d'un produit L'objectif est de prédire le nombre de ventes d'un produit en fonction de la publicité dépensée, de la saison, du prix, etc.

6. Détection de fraude par carte de crédit L'objectif est de déterminer si une transaction par carte de crédit est frauduleuse (fraude) ou non (non) en fonction des caractéristiques de la transaction.

7. Diagnostic de diabète chez les patients L'objectif est de déterminer si un patient est atteint de diabète (oui) ou non (non) en se basant sur des mesures telles que la glycémie, l'âge, l'indice de masse corporelle (IMC), l'hérédité, etc.

8. Recommandation de produits cosmétiques à des clients L'objectif est de recommander des produits cosmétiques à des clients en fonction de leur type de peau, de leurs préférences, de leurs besoins spécifiques (par exemple, hydratation, anti-âge, anti-acné), et des caractéristiques des produits (ingrédients, type de produit, etc.).

Réponses

- ▶ Exemple 1 : **Régression** (prédiction d'une valeur numérique, la note d'un film).
- ▶ Exemple 2 : **Classification** (détection de spam, une tâche de classement binaire).
- ▶ Exemple 3 : **Régression** (prédiction d'une valeur numérique, le prix d'une maison).
- ▶ Exemple 4 : **Classification** (classification en catégories d'espèces d'oiseaux).
- ▶ Exemple 5 : **Régression** (prédiction d'une valeur numérique, le nombre de ventes).
- ▶ Exemple 6 : **Classification** (détection de fraude, une tâche de classement binaire).
- ▶ Exemple 7 : **Classification** (diagnostic binaire du diabète, une tâche de classement binaire)
- ▶ Exemple 8 : **Classification** (recommandation de produits cosmétiques, une tâche de classement ou de catégorisation des produits en fonction des besoins et des préférences du client).