L’Intelligence Artificielle dans l’industrie des procédés

25 octobre 2024

SEII

Déjeuner-Conférence

société européenne des ingénieurs et des industriels
european society for engineers and industrialists

Introduction

- L'intelligence artificielle (IA) révolutionne de nombreux secteurs, et l'industrie des procédés n ‘y fait pas exception
- Avec le machine learning (ML), sous-branche de l’IA, les entreprises améliorent la productivité, réduisent les coûts et augmentent la qualité de leurs produits.
- Les modèles entraînés sur des données historiques permet d'anticiper les pannes et d'optimiser la maintenance, réduisant ainsi les temps d'arrêt

Objectifs et plan de l’exposé

Dans cette présentation, nous allons :

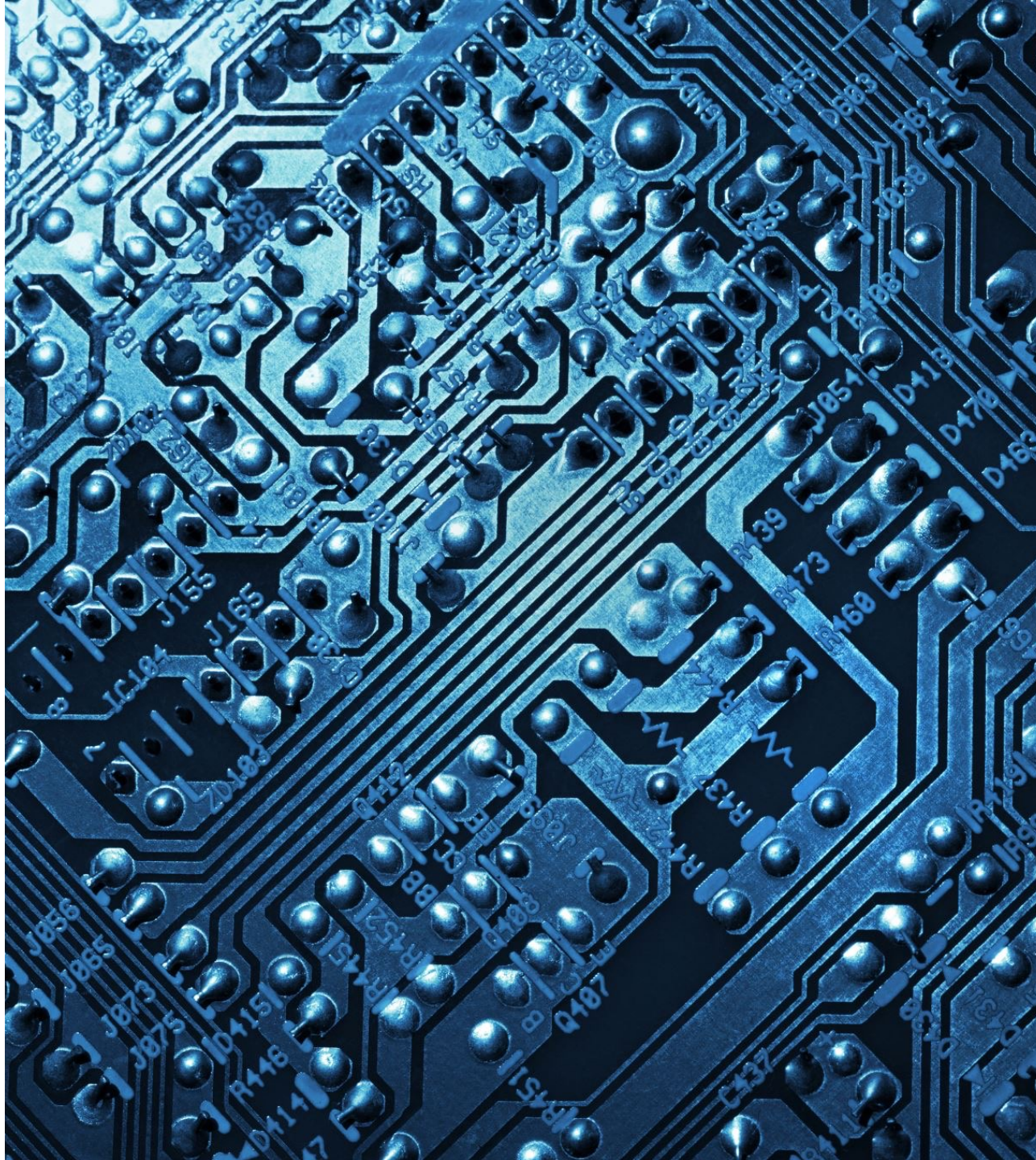
1. Introduire les concepts fondamentaux de l’IA , plus spécifiquement du ML
2. Examiner la structure des données industrielles pour la mise en œuvre de ces technologies.
3. Découvrir comment l'IA transforme les opérations à travers des exemples concrets.
4. Évaluer les limites actuelles et les perspectives d'avenir de l'IA dans ce secteur.

Concepts clés de l’IA

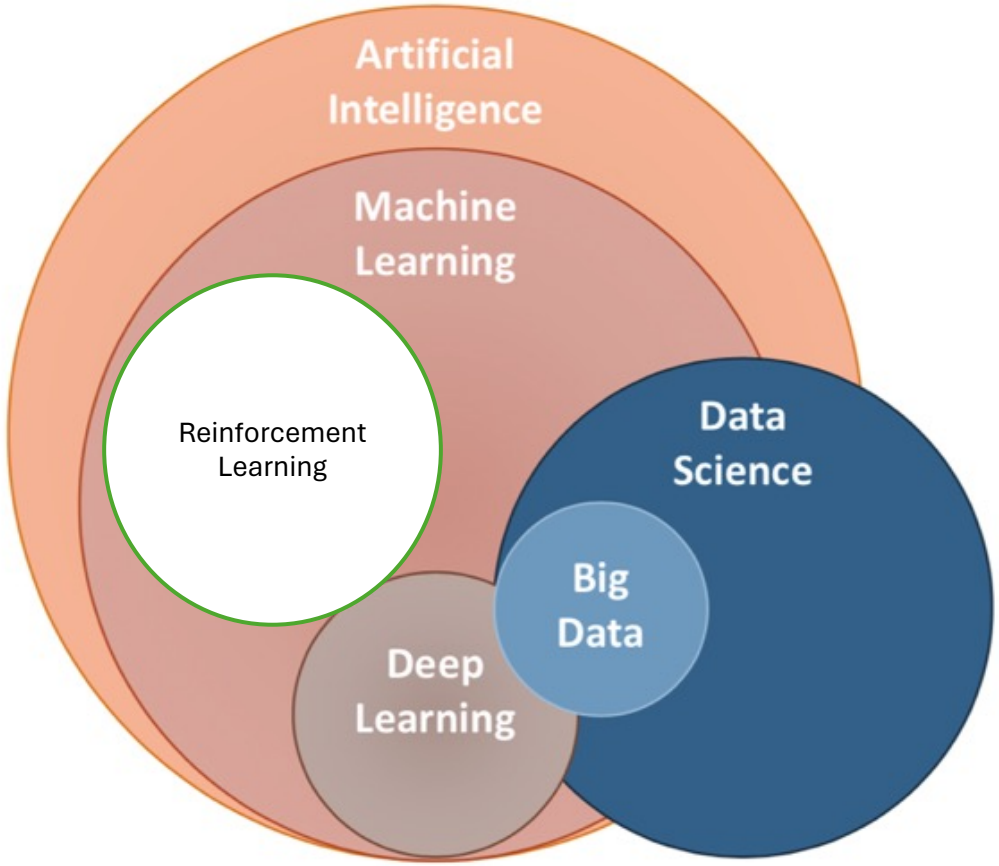
1

Définition de l’IA

L'intelligence artificielle (IA) désigne l’ensemble des techniques et des technologies qui permettent à une machine de simuler des capacités cognitives humaines telles que la compréhension, le raisonnement, l’apprentissage et la résolution de problèmes.

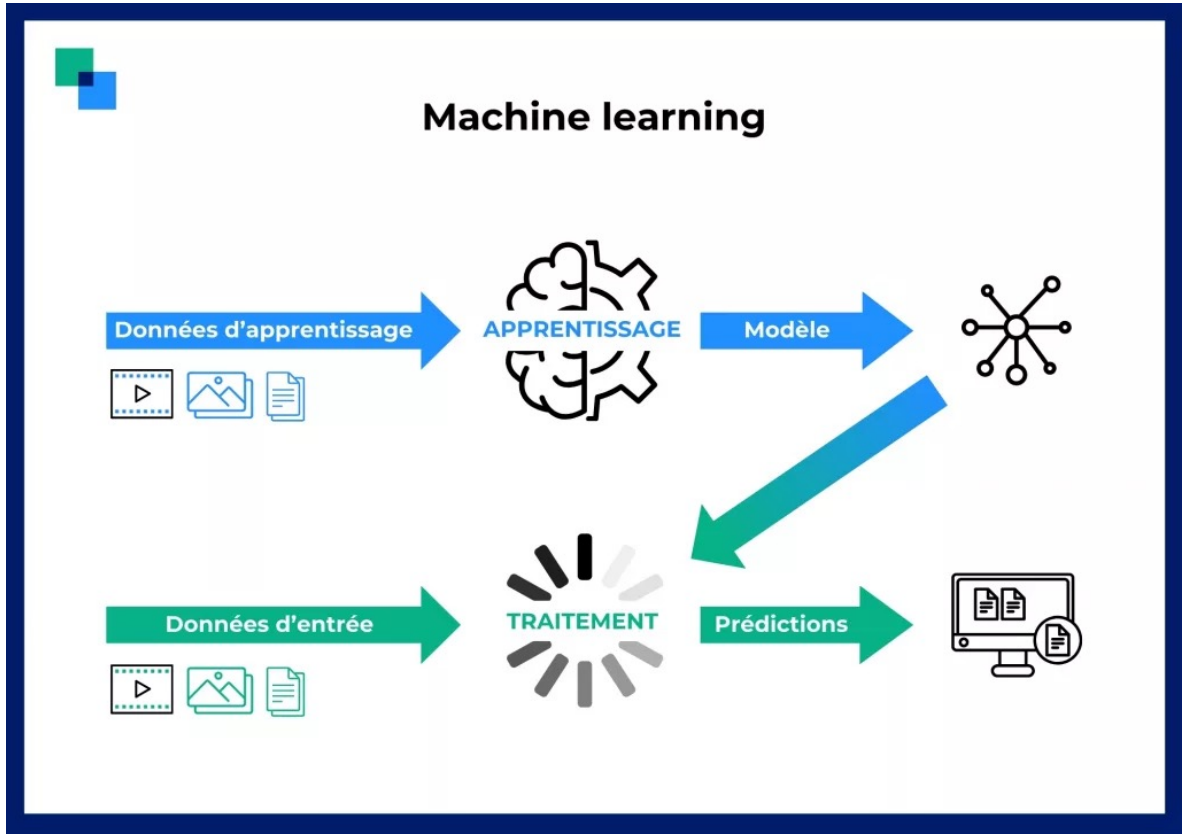


IA et concepts associés

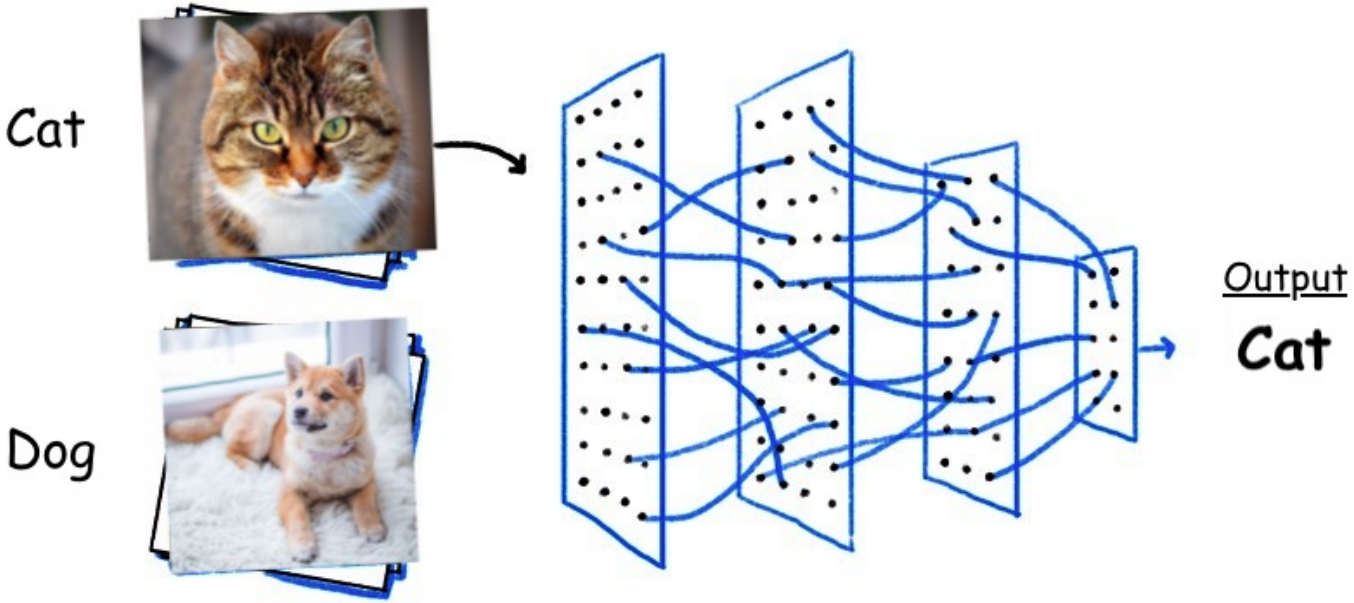


Machine Learning (ML)

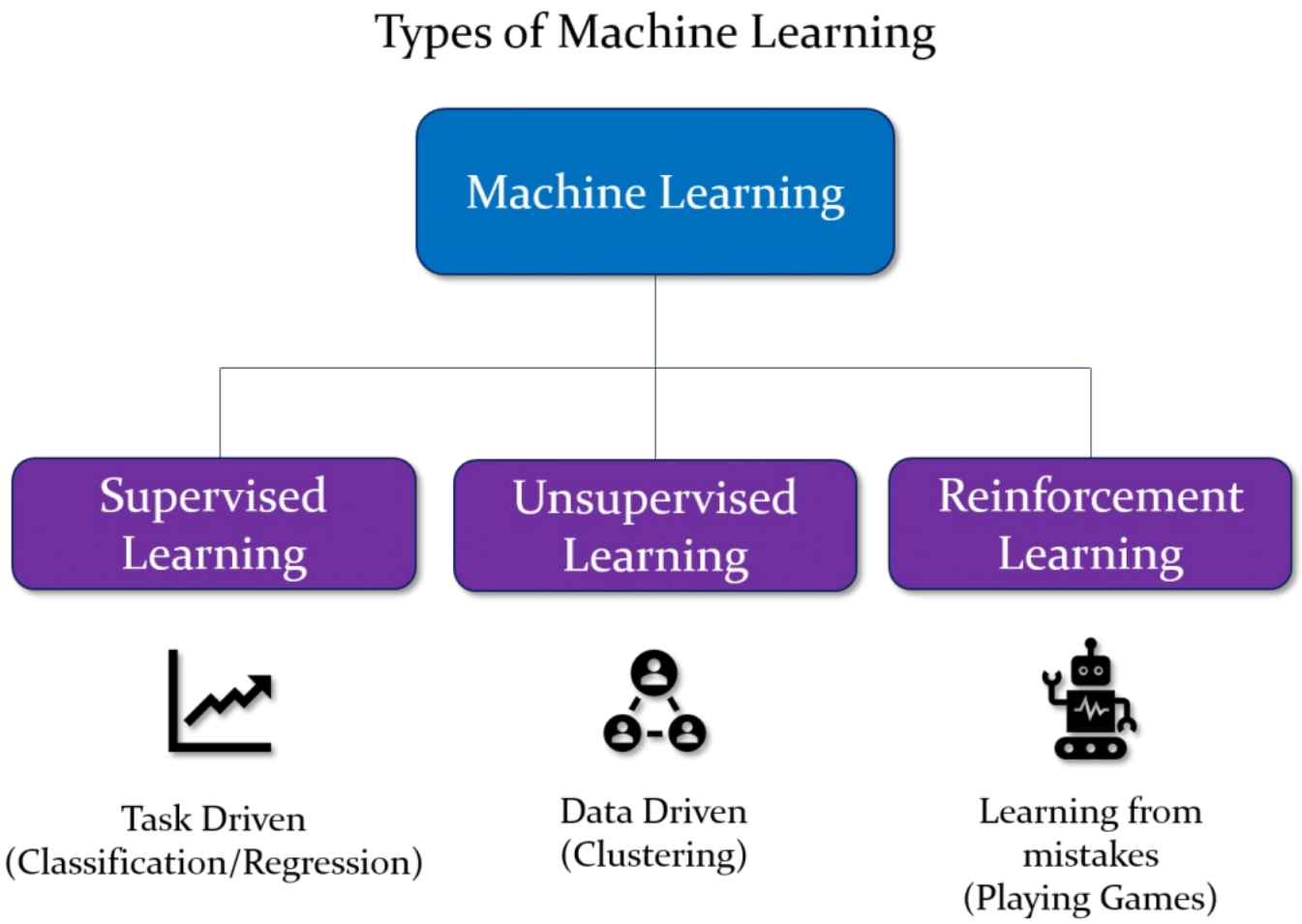
Apprendre à partir de données, pour entraîner des modèles



Exemple de modèle: reconnaissance d’images

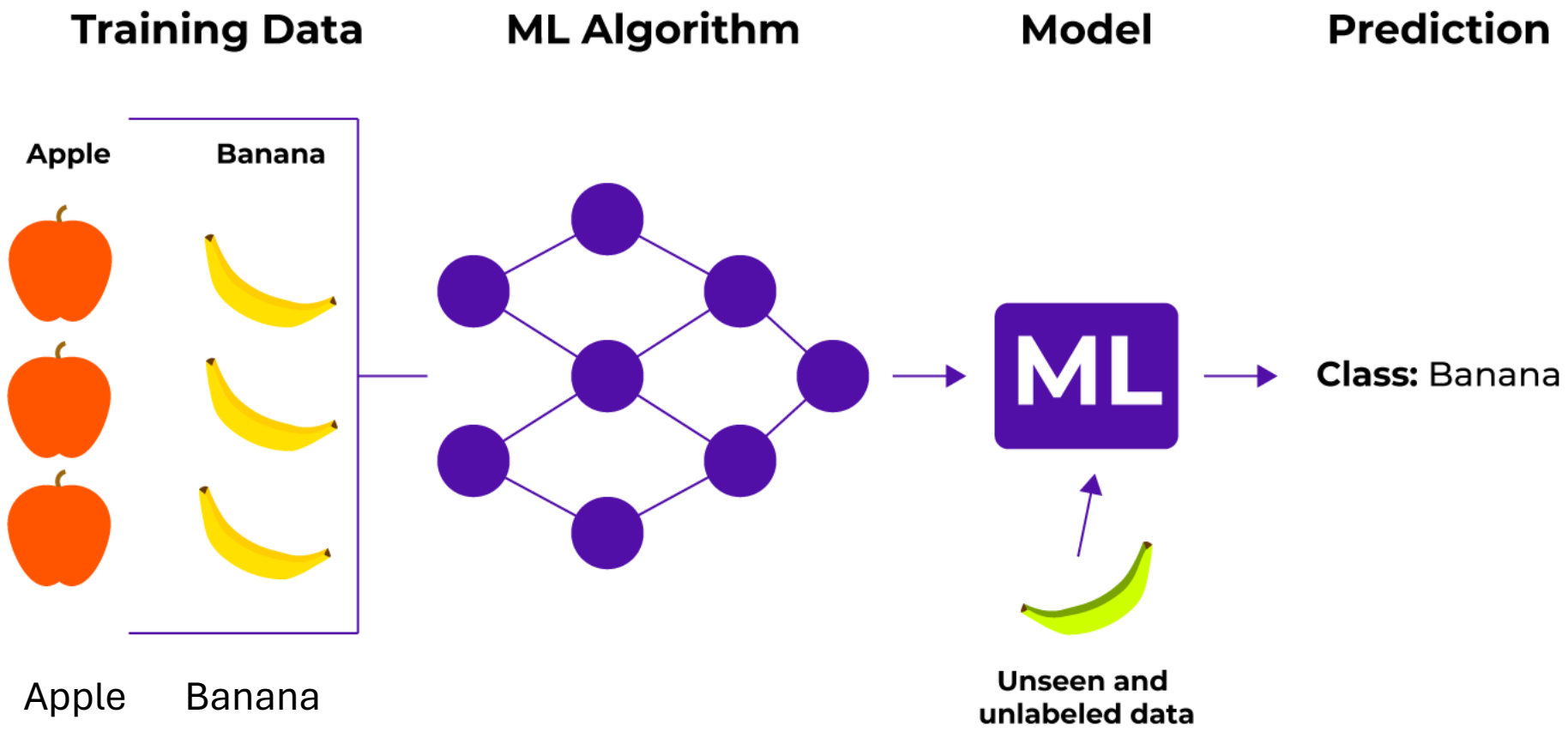


Trois types de ML



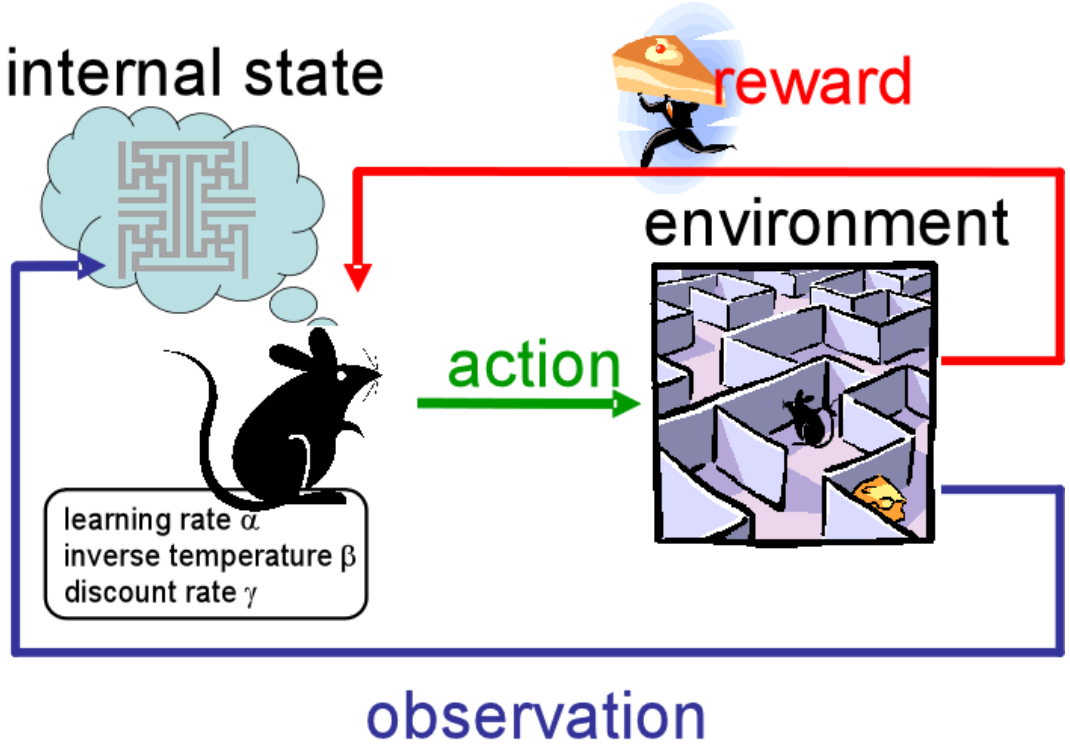
Apprentissage supervisé

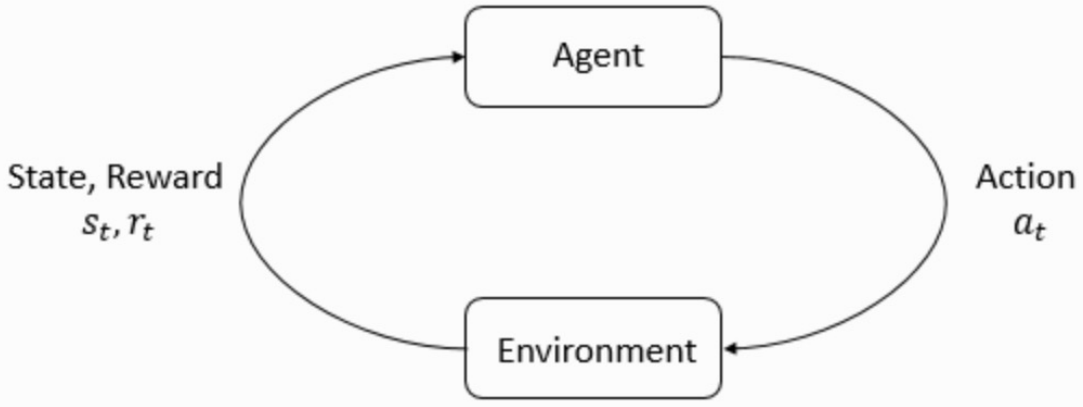
Phase d'apprentissage: données d'entrée et de sortie *connues* et *etiquetées*



Label: Apple Banana

Reinforcement Learning



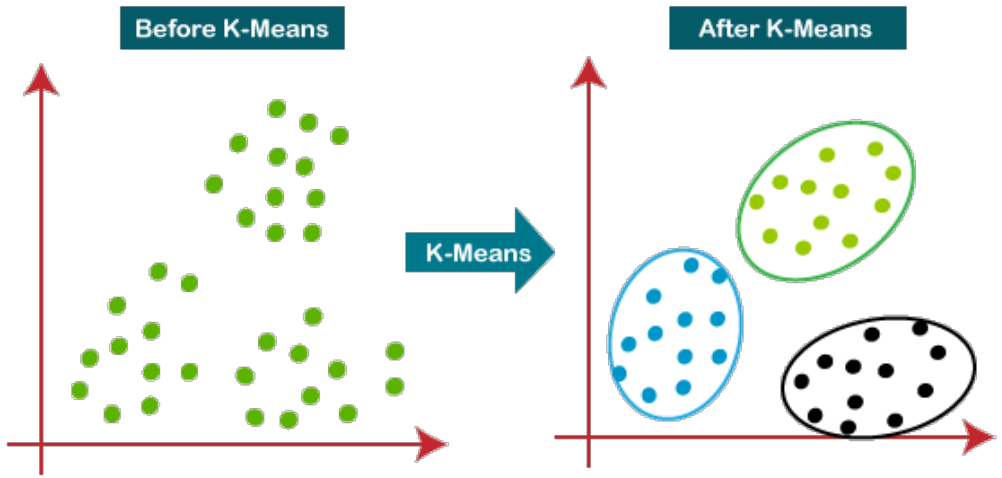


Agent-environment interaction loop.

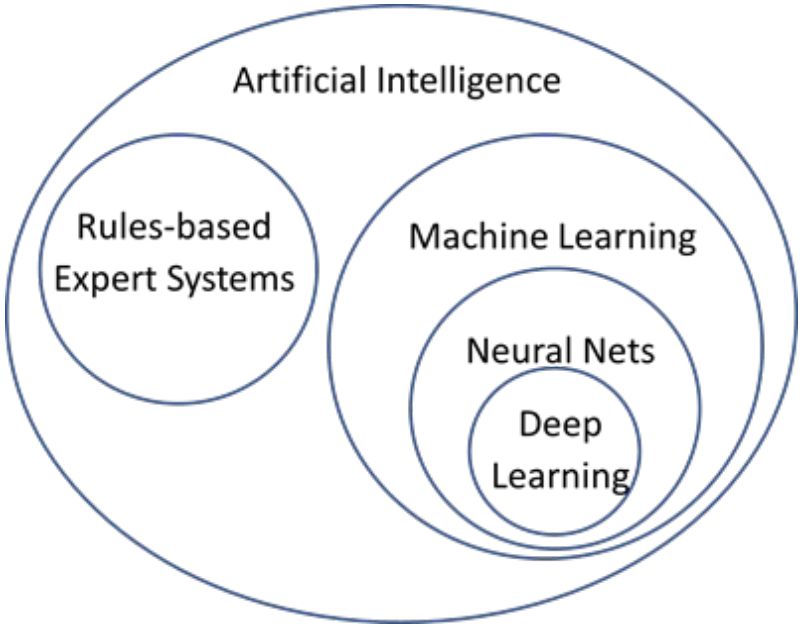
Optimisation d'une fonction stratégie/action qui maximise les récompenses sur le long terme
Pas de "bonne" réponse à chaque étape, contrairement au supervisé

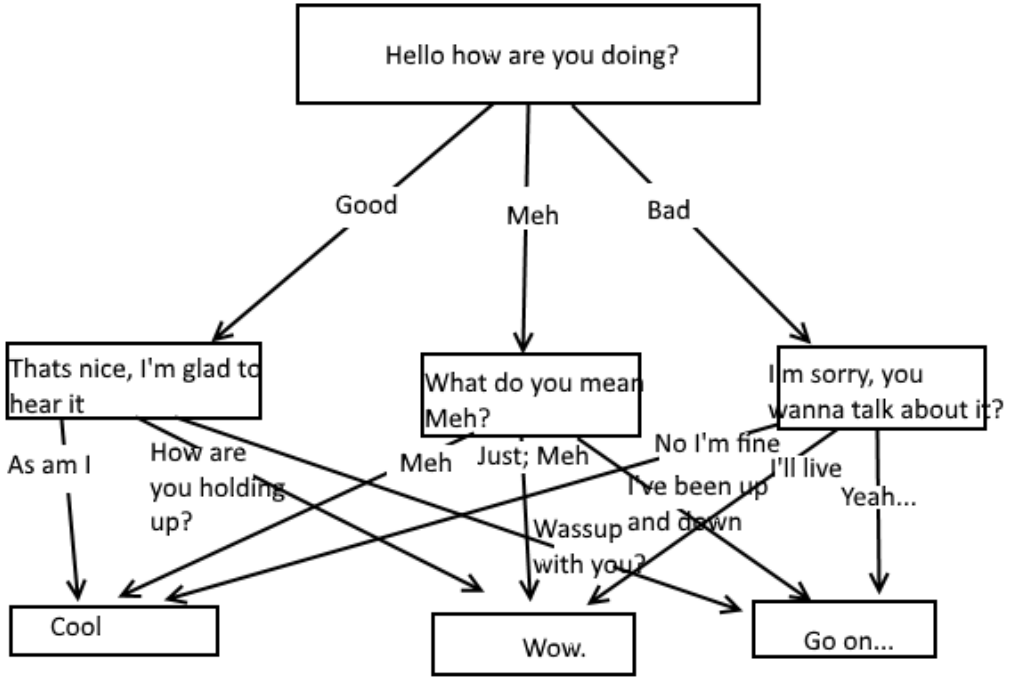
Apprentissage non supervisé

- Analyse des données sans labels
- L'objectif est de découvrir des structures cachées ou des schémas dans les données.
- Le clustering regroupe des données similaires dans des "clusters" (groupes) selon leurs caractéristiques.
- **Résultat attendu** : Classification des données en groupes distincts sans connaissance préalable des classes ou catégories.



IA qui n’est pas du ML- Rule based



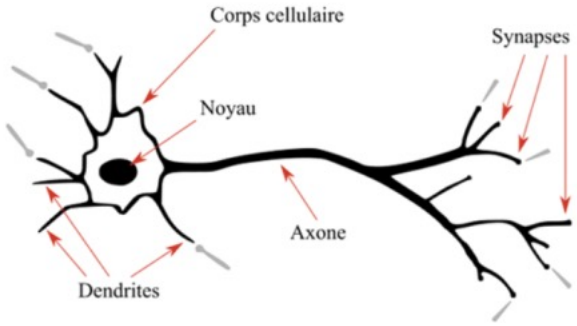


Ensemble de règles préétablies

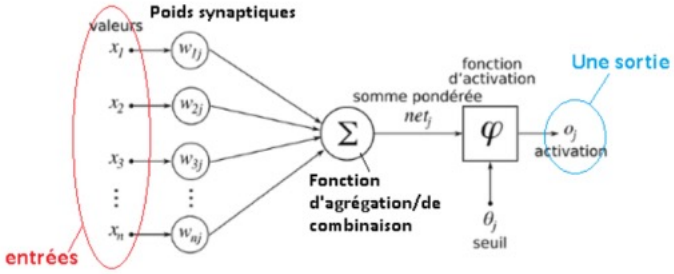
Pas d’apprentissage à partir de données

Supervisé Réseau de neurones

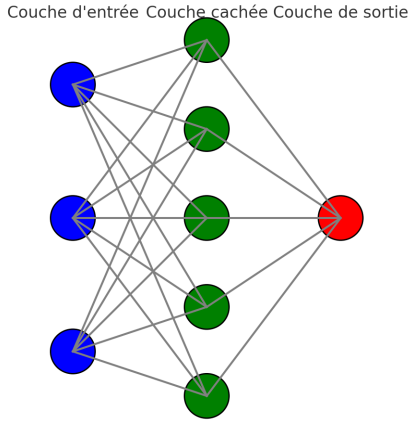
- **(RN)** modèle informatique inspiré du cerveau humain
  - Chaque neurone reçoit plusieurs entrées pondérées, applique une fonction mathématique, puis transmet un signal de sortie à d'autres neurones.
  - Le réseau ajuste des “poids” des connexions entre les neurones lors de l'entraînement, afin de minimiser l'erreur entre la sortie prédite et la sortie attendue.
  - Les **réseaux** sont organisés en **couches**



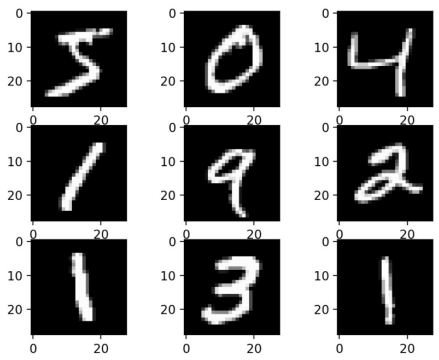
NEURONE BIOLOGIQUE



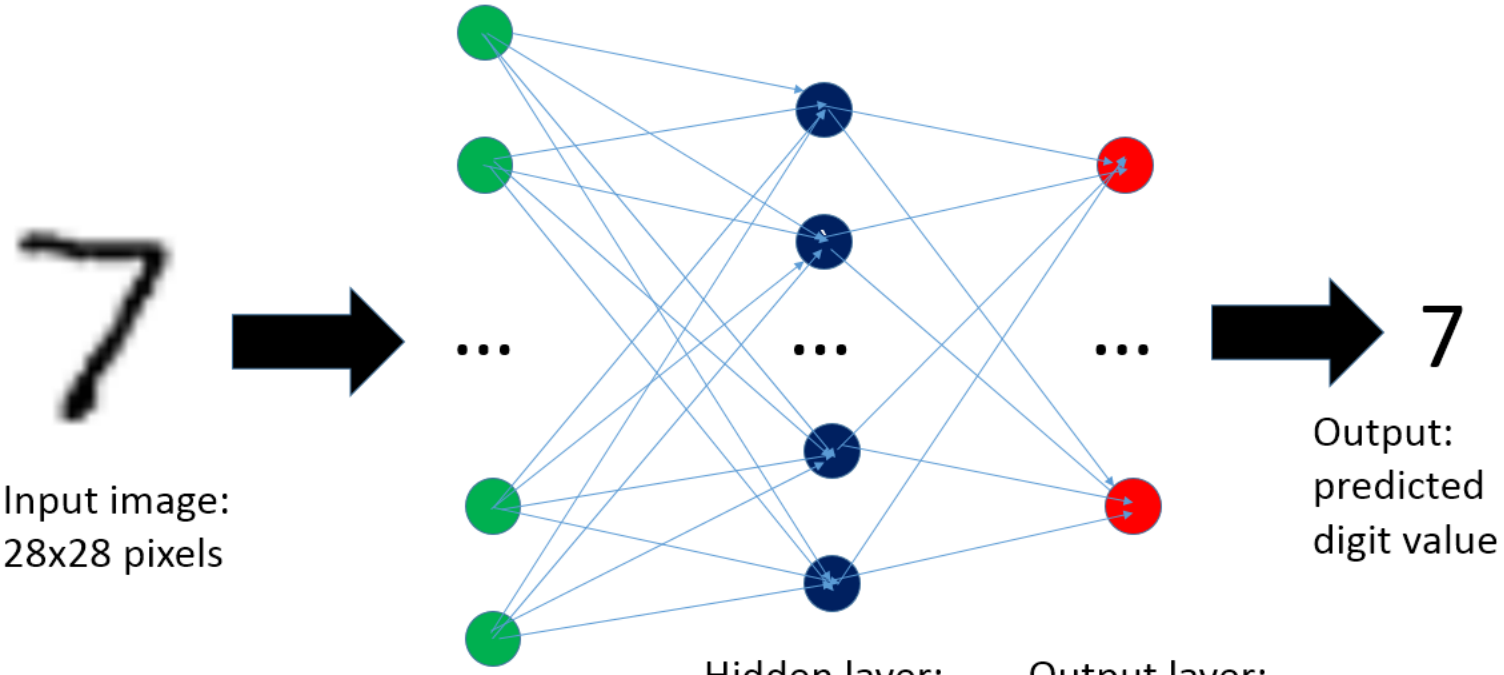
NEURONE ARTIFICIEL



Exemple de RN: reconnaissance d’images



Entraînement: 80000 images



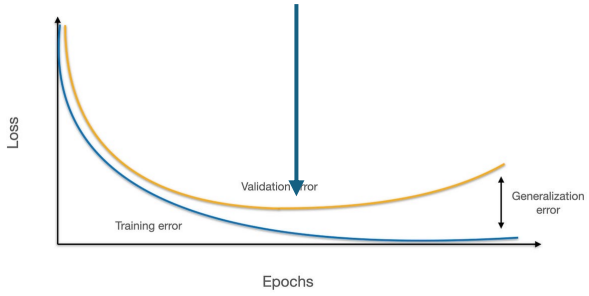
Input image:
28x28 pixels

Input layer:
784 neurons,
one per pixel

Hidden layer:
100 neurons

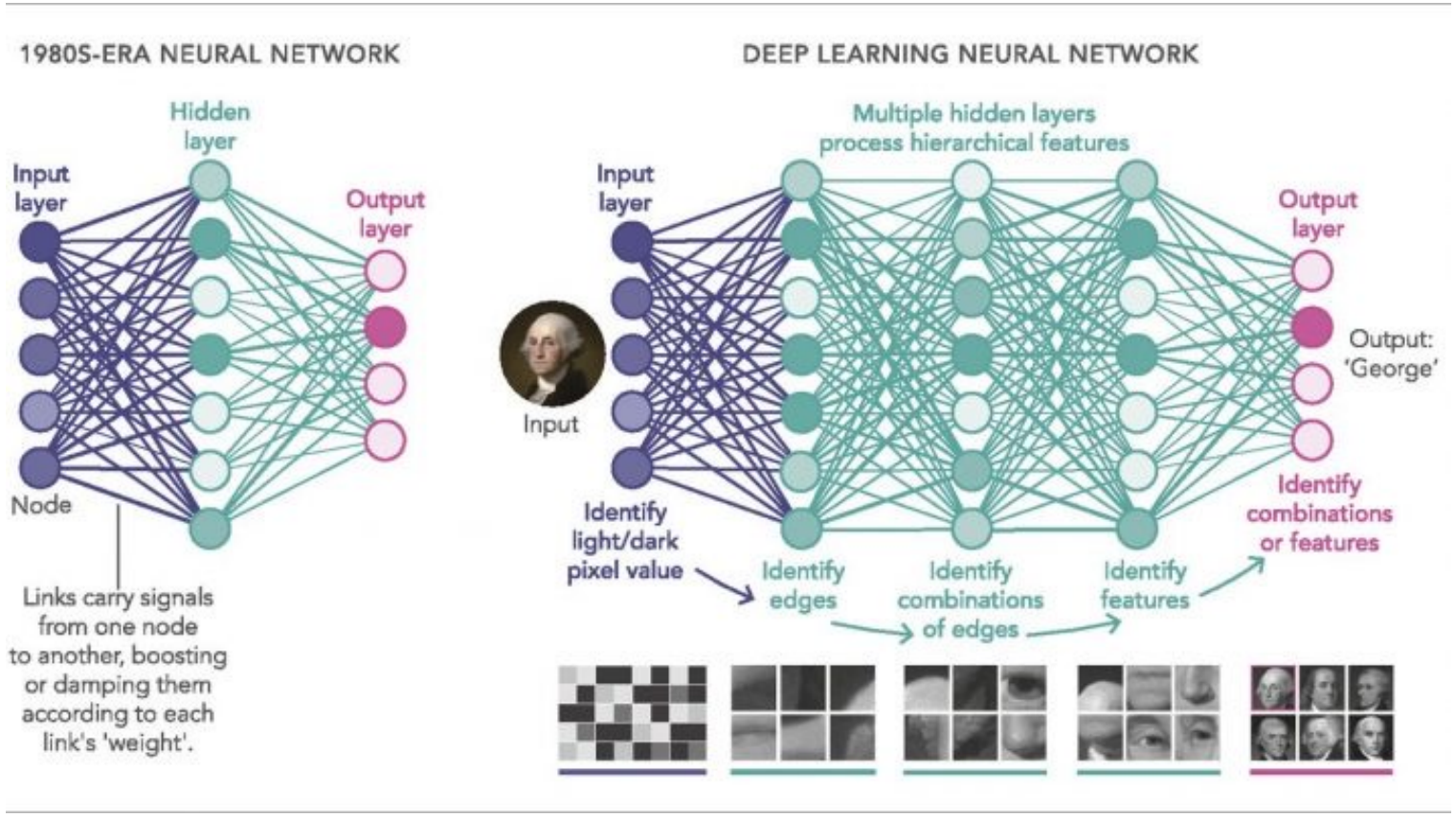
Output layer:
10 neurons

Output:
predicted
digit value



On arrête l’entraînement au “bon” moment

Deep Learning=RN profonds



Chat GPT est du deep learning

GPT 4

175 milliards de neurones

1800 milliards de paramètres

Cerveau humain

100 milliards de neurones

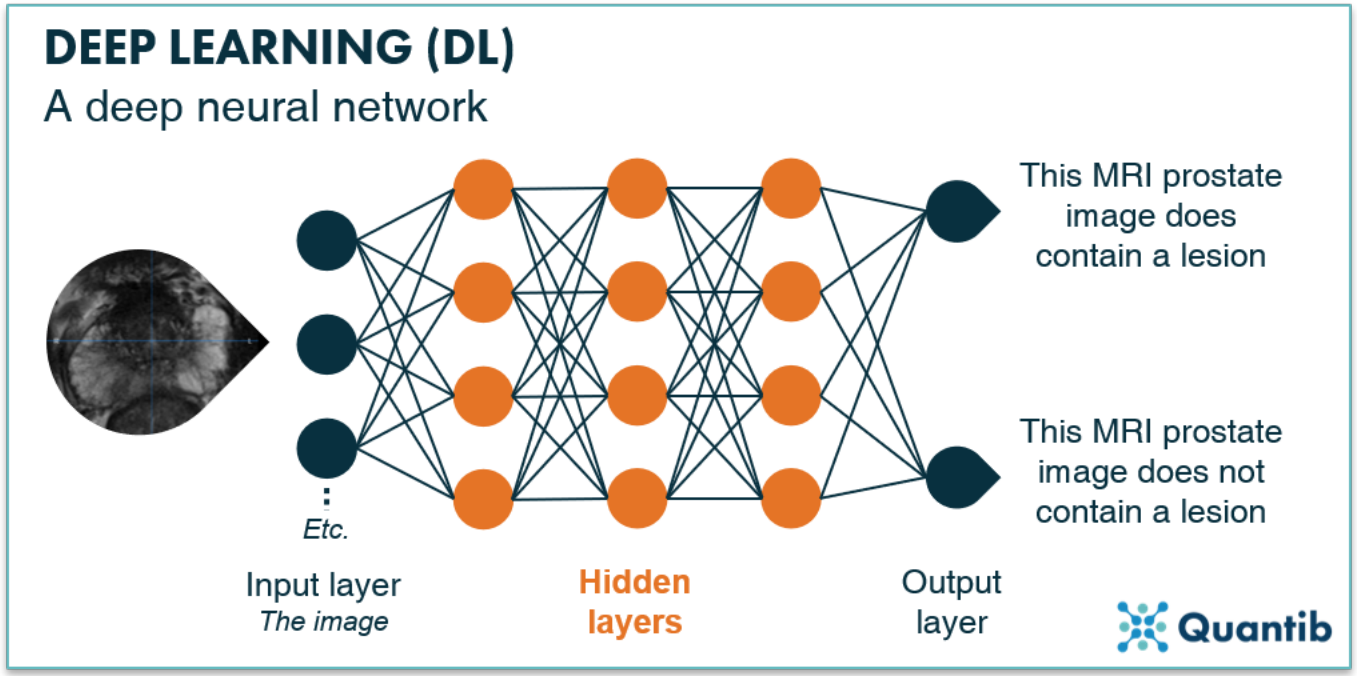
L’actualité fait bien les choses..

- Geoffrey Hinton
- Prix Nobel de physique 2024
- ”Pape” de l’IA

Principales étapes du deep learning

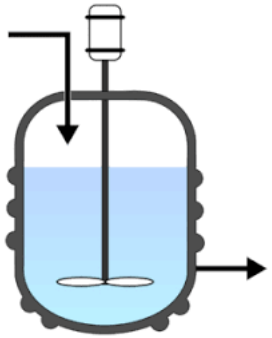
- **1950s-1980s** : Introduction des réseaux de neurones avec le perceptron (Frank Rosenblatt) et développement de l'algorithme de rétropropagation (Hinton, Rumelhart, Williams).
- **2006** : Percée avec les réseaux de croyances profondes (DBN) par Geoffrey Hinton, surmontant le problème du gradient qui disparaît.
- **2012** : Révolution avec AlexNet (Krizhevsky, Sutskever, Hinton), gagnant du défi ImageNet, popularisant les reseaux de neurones en vision par ordinateur.
- **2010s** : Explosion du deep learning grâce à l'amélioration des GPU et des grandes bases de données

Deep learning et oncologie



Deep learning et ingénieur procédé

Modèle



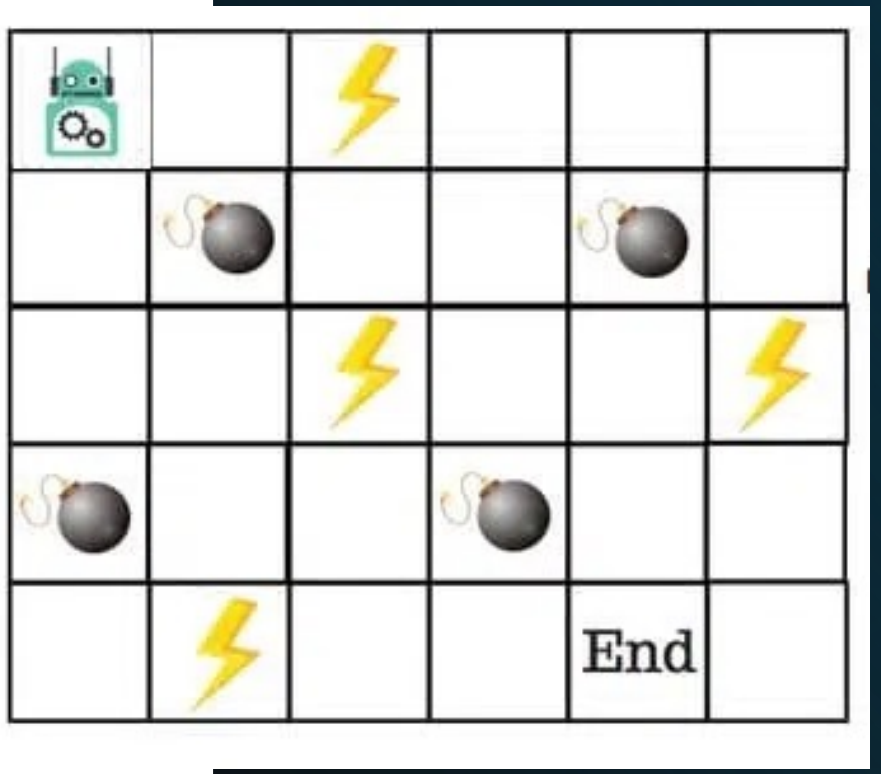
Ma réaction chimique est-elle en régime normal?

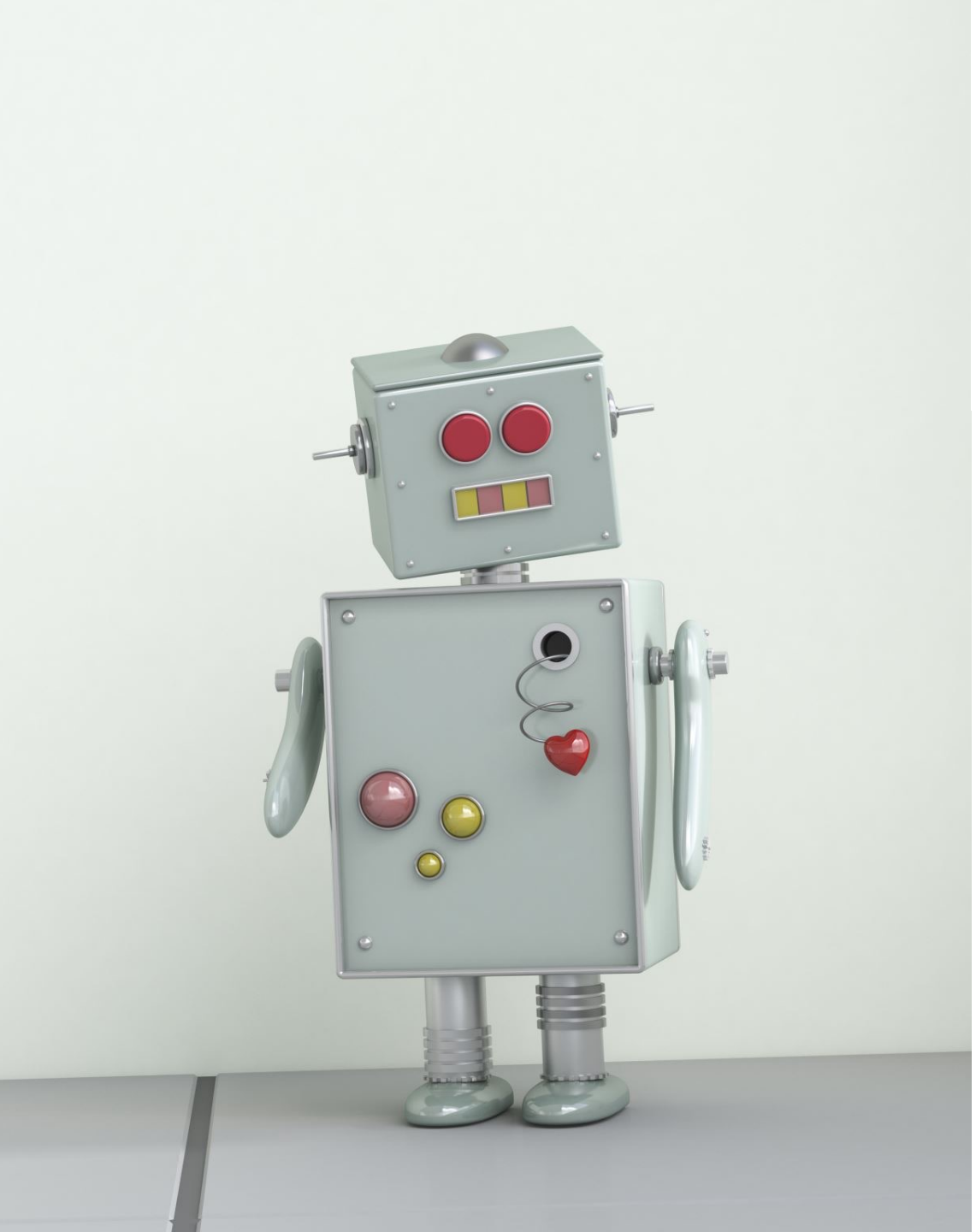
Reinforcement learning

Un robot dans une pièce apprend à se déplacer vers une zone cible (End).

Le robot ne connaît ni la disposition de la pièce ni l’emplacement de la cible au début.

Il doit éviter les obstacles (bombe), prendre le chemin le plus court, et maximiser le nombre de recompenses (éclair).





Reinforcement Learning

1. **Agent (le robot)** : L’agent explore son environnement (la pièce).
2. **Actions** : Le robot peut se déplacer dans différentes directions (gauche, droite, avant, arrière).
3. **Récompense** :
    1. **Positive** : Si le robot se rapproche de la cible(+1) ou prend une recompense(+100)
    2. **Négative** : Si le robot frappe un obstacle(-100) ou s’éloigne de la cible (-1)
4. **Apprentissage** : Le robot ajuste son comportement pour maximiser recompense sur le long terme et atteindre la cible plus rapidement.

L’agent (le robot) **apprend par essais et erreurs**, en améliorant progressivement sa stratégie pour atteindre la cible, sans instructions explicites.

Résumé intermédiare

- La terminologie de l'IA
  - Aujourd'hui, IA et ML pratiquement confondus
- Les bases du ML et ses 3 grandes familles
- L'utilisateur dispose d'un modèle préalablement entraîné sur des données (ML)
- Les données sont le carburant du ML
- Les modèles sont de plus en plus complexes, et nécessitent logiquement de plus en plus de données.

Les données dans l’industrie process

2

Secteurs concernés

Industrie pétrolière et gazière : Raffinage du pétrole, traitement du gaz naturel.

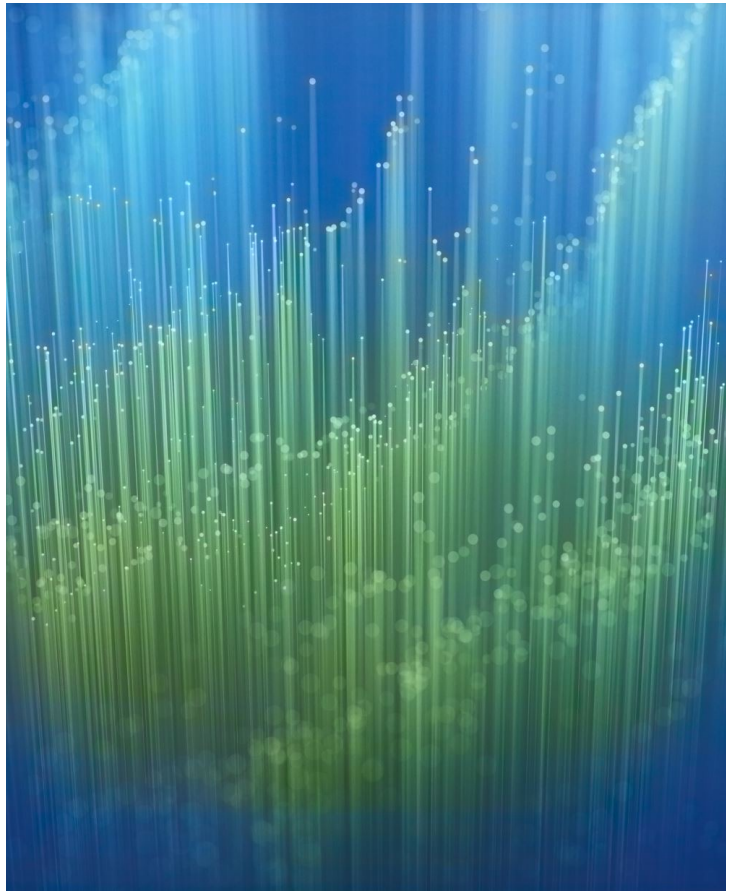
Industrie chimique : Production de produits chimiques, polymères, plastiques, peintures, etc.

Industrie pharmaceutique : Fabrication de médicaments à partir de procédés chimiques ou biologiques.

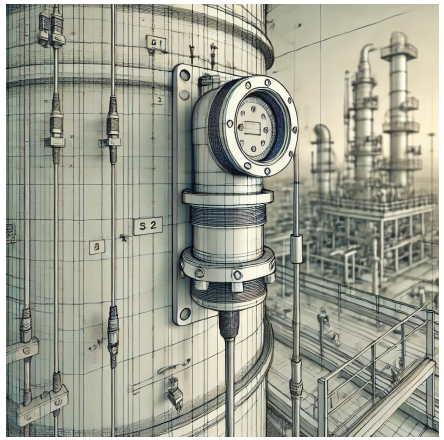
Agroalimentaire : Transformation des produits agricoles (lait, viande, céréales) en produits alimentaires.

Industrie papetière : Fabrication du papier et du carton.

Processus en continu et en batch : Certains procédés fonctionnent de manière continue (raffinage, production chimique), tandis que d'autres fonctionnent par batch ou par lots (fabrication pharmaceutique).



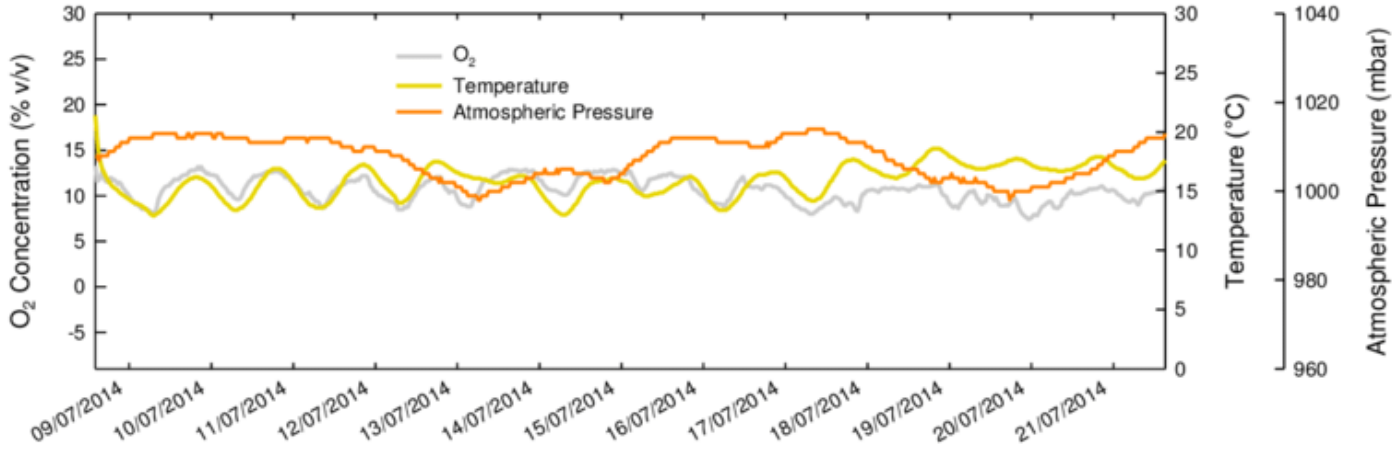
L’ingénieur procédé



Ingénieur procédé dans une raffinerie, qui visualise des *séries temporelles* de mesures de capteurs

Les séries temporelles (time series)

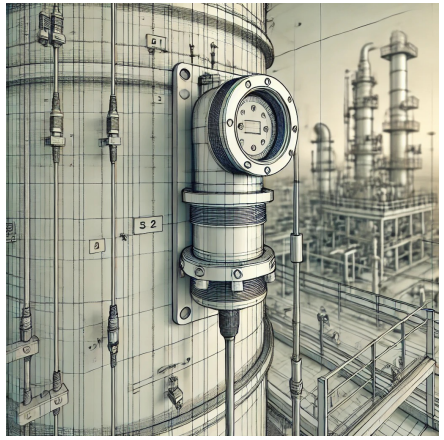
- Une **série temporelle** est une **suite de données** ou d'observations mesurées ou collectées à des **intervalles de temps réguliers ou irréguliers**. Chaque point de la série est associé à un **moment précis dans le temps**, ce qui permet de suivre l'évolution d'un phénomène au fil du temps.



Série temporelle

Tableau de bord de series temporelles des performances d'une turbine (**AVEVA™ Data Hub**).

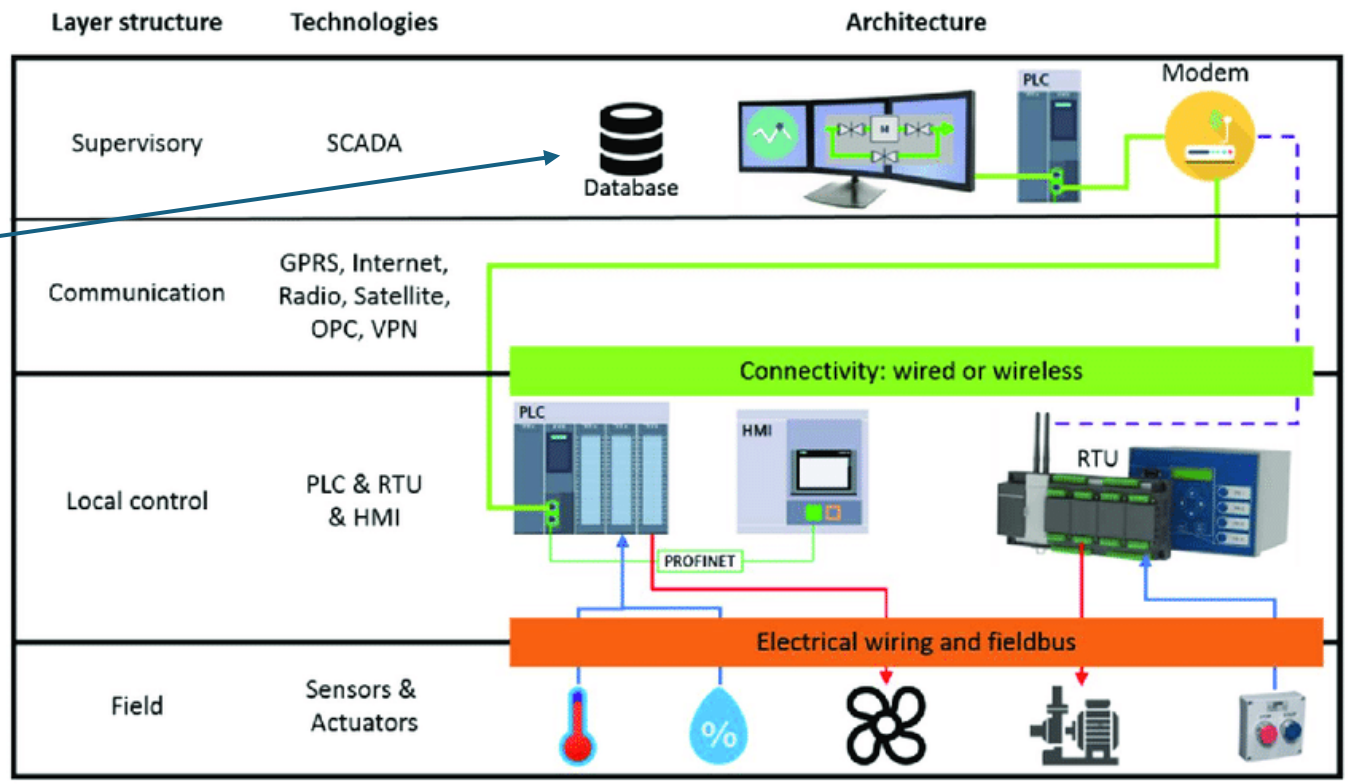
IOT (Internet of things)



L'**Internet des Objets (IoT)**, ou **Internet of Things** en anglais, désigne un réseau d'objets physiques (ou "choses") équipés de capteurs, logiciels, et autres technologies, qui leur permettent de **collecter, échanger et transmettre des données** via Internet ou d'autres réseaux de communication. Ces objets peuvent être aussi bien des appareils électroménagers, des machines industrielles, des capteurs environnementaux ou même des véhicules.

L’infrastructure SCADA (Supervisory Control and Data Acquisition)

Database=Data Historian

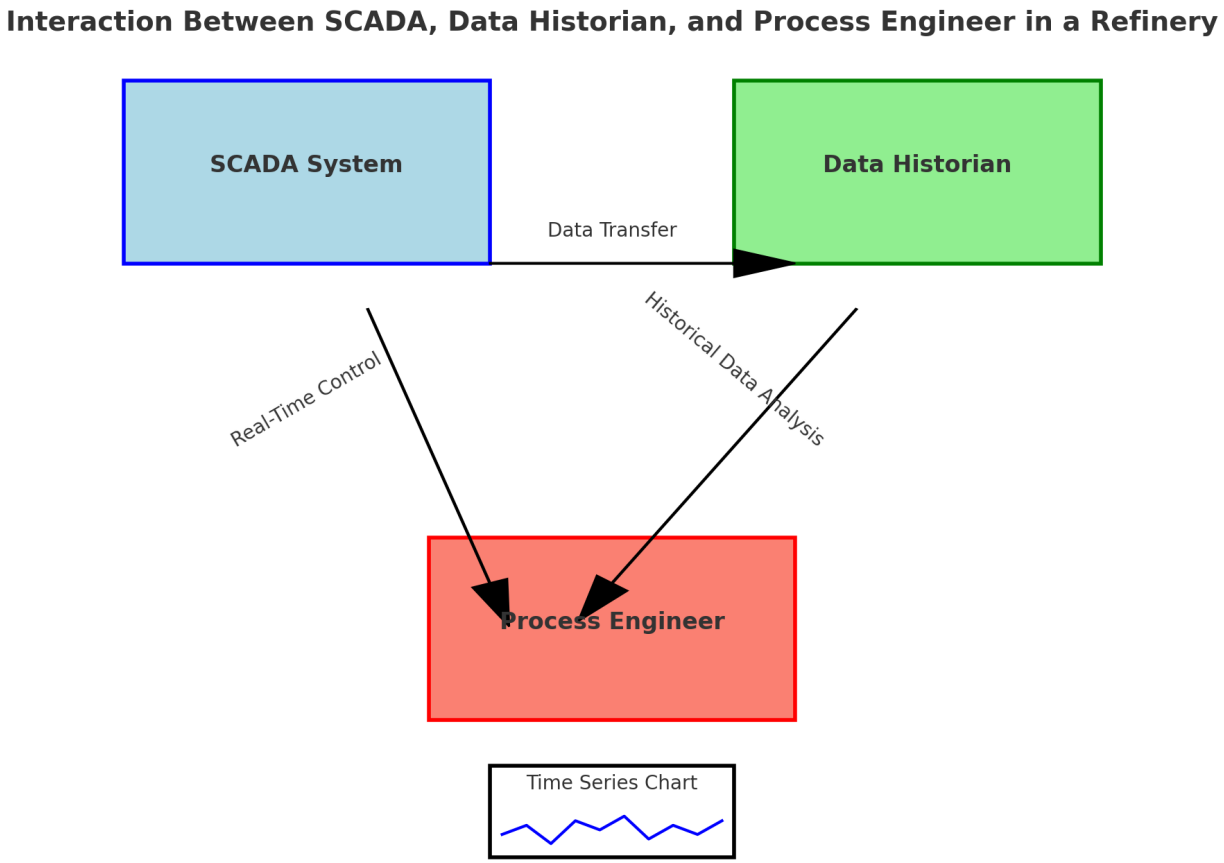


Data Historian: base de données optimisée pour séries temporelles

Un data historian est conçu pour stocker et gérer des **données de séries temporelles** (ex: température d'un réacteur mesurée toutes les seconds,la pression d'une conduite surveillée en continu).

Contrairement aux bases de données classiques (relationnelles) avec tables et jointures, un data historian est optimisé pour traiter des **flux continus de données** temporelles, généralement en provenance de capteurs industriels.

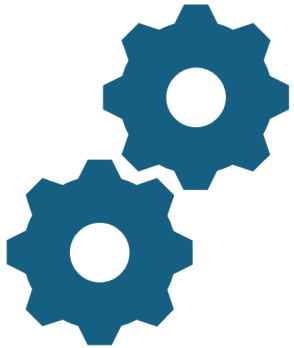
En résumé


Interaction Between SCADA, Data Historian, and Process Engineer in a Refinery
SCADA System
Data Historian
Data Transfer
Real-Time Control
Historical Data Analysis
Process Engineer
Time Series Chart

Cas d’utilisation

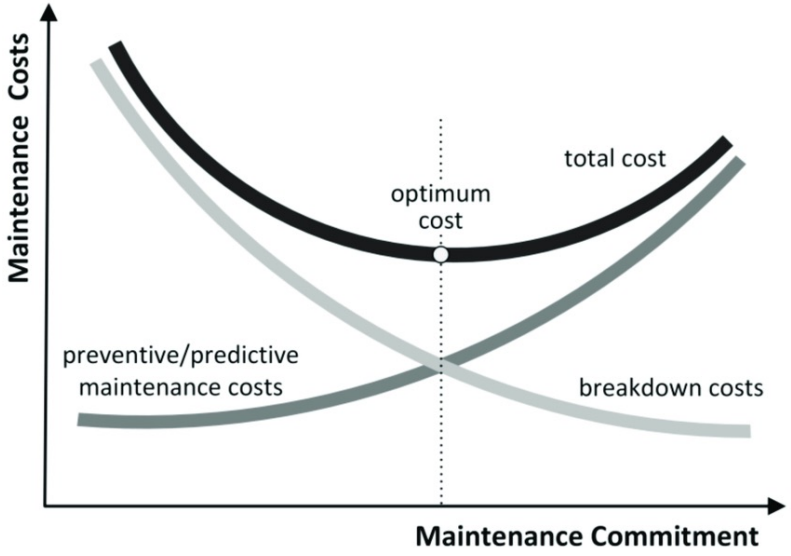
3

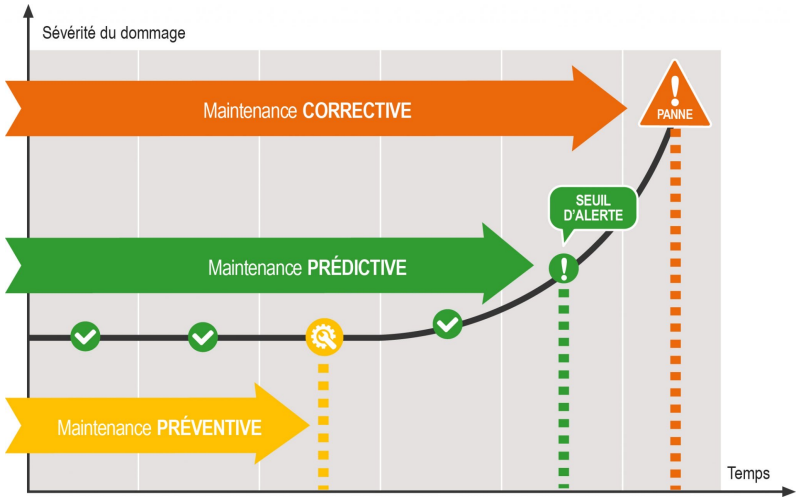
Cas d’utilisation



Maintenance prédictive

Optimisation de procédé



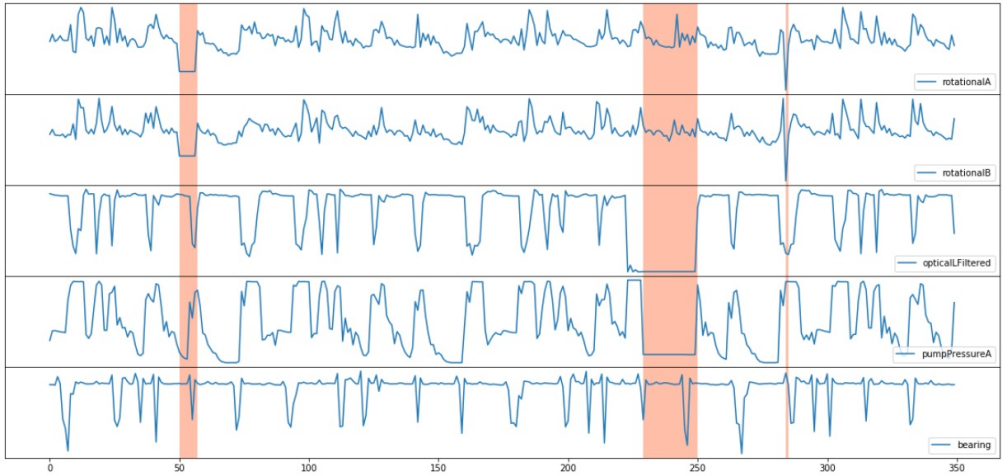


La maintenance prédictive

- La **maintenance prédictive** est une stratégie proactive qui utilise des données en temps réel des capteurs pour prévoir les pannes d'équipements avant qu'elles ne surviennent.
- Quel est le bon moment pour changer son chauffe-eau?
  - Tous les ans?
  - Attendre qu'il rende l'âme?
  - Au bon moment?

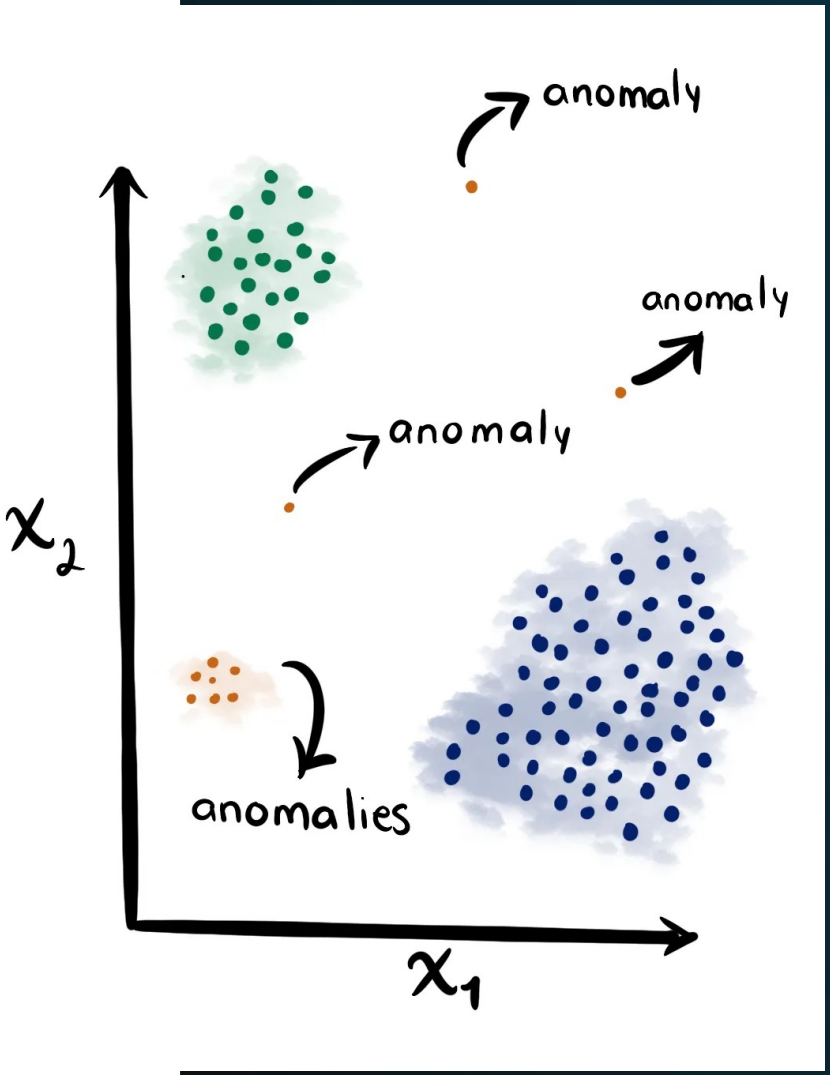
Solution ML: la détection d’anomalie

- On entraîne un réseau de neurones à distinguer les régimes de fonctionnement normal et anormal, à partir de (nombreuses) variables d’entrée (température, pression, puissance, etc…)
- Ensuite, on monitore le procédé, qui renvoie un seul indicateur de “normalité”



Solution ML: la détection d’anomalie non supervisée

- Mesure de la distance entre un point de donnée et le **cluster le plus proche**.
- Si un point se trouve à une distance significativement plus grande par rapport aux autres points du cluster, il est considéré comme une **anomalie**.
- L'idée est que les points normaux sont proches du centre de leur cluster, tandis que les anomalies se situent loin, dans les zones où la densité des données est faible.
- Ces points éloignés, sont ainsi facilement détectés en fonction de leur distance élevée aux clusters.



Bénéfices

1. **Réduction des arrêts imprévus** en anticipant les défaillances.
2. **Optimisation des coûts** en intervenant uniquement lorsque nécessaire.
3. **Prolongation de la durée de vie** des équipements grâce à la surveillance continue.
4. **Amélioration de la sécurité** en évitant les pannes critiques.

Optimisation de procédé

Dans l’industrie des procédés, le reinforcement learning (RL) optimise les équipements pour maximiser la production, réduire la consommation d'énergie ou les émissions.

Exemple: dans une raffinerie, un agent RL ajuste les paramètres d'une colonne de distillation pour améliorer la pureté du produit tout en réduisant les coûts énergétiques.

C’est mieux que l'approche classique par essai-erreur qui est coûteuse et longue.

Exemple du raffinage

Distillation fractionnée : Séparation du brut en composants (essence, diesel, naphta)

- **Quantité à optimiser** : taux de récupération des fractions légères et lourdes.

Conversion : Transformation des fractions lourdes en produits légers (craquage catalytique, hydrotraitement).

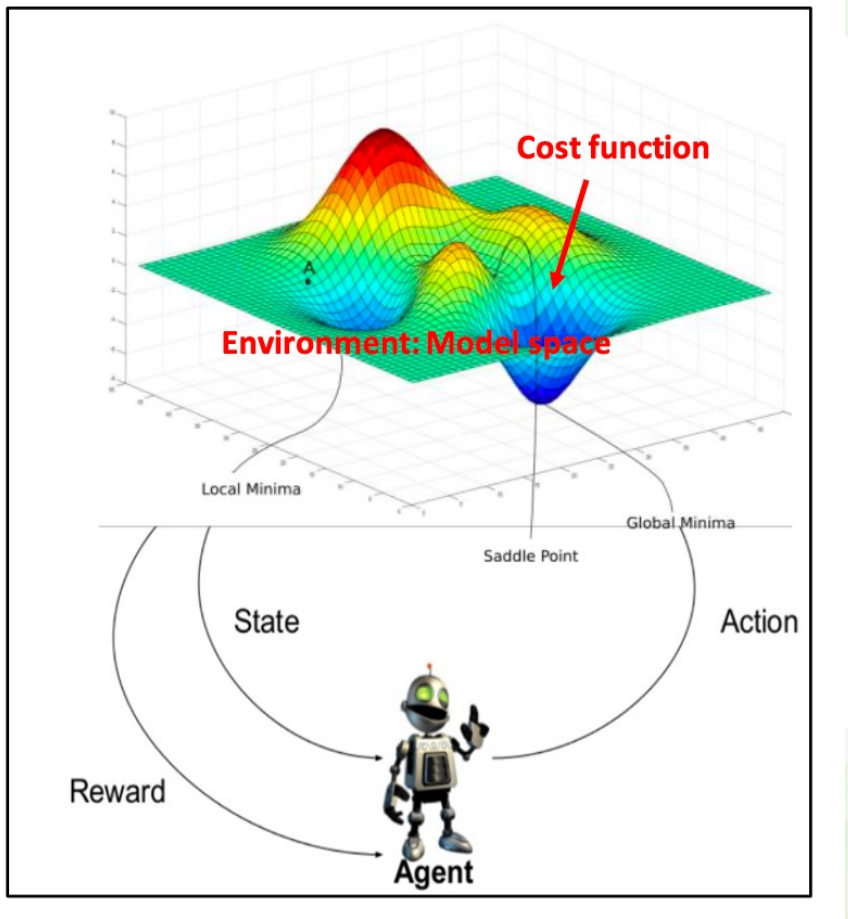
- **Quantité à optimiser** : rendement en produits légers (essence, diesel).

Traitement et purification : Élimination des impuretés (soufre, azote) pour normes environnementales.

- **Quantité à optimiser** : taux de désulfuration et qualité finale des produits.

Optimisation de procédé via le Reinforcement Learning

- **Analogie avec un Robot**
- **Processus Chimique** : Dans un procédé chimique, l'agent RL remplace le robot et doit apprendre à ajuster les paramètres du processus (température, pression, débit) pour **optimiser** la production (objectif : maximiser le rendement, minimiser les coûts).
    - **Action positive** : Un réglage améliore la production ou économise de l'énergie (récompense positive).
    - **Action négative** : Un réglage inefficace augmente les coûts ou diminue le rendement (récompense négative).



Bénéfices

- Le RL permet au système d'apprendre continuellement en s'adaptant aux changements du processus chimique, comme un robot ajustant son trajet en fonction des obstacles.
- Il automatise l'optimisation en testant et apprenant à partir des résultats, à l'image d'un robot qui s'entraîne pour atteindre une cible dans un environnement inconnu

Limites et perspectives

4

Complexité des processus industriels

Multiplicité des modèles : Les algorithmes de ML peuvent être inefficaces pour modéliser des relations complexes, surtout si les données historiques ne couvrent pas toutes les variations possibles.

Modèles peu robustes : Le sur-apprentissage (overfitting) peut rendre les modèles de ML peu adaptables aux conditions non observées.

Garbage in, garbage out

Dans l'industrie, les données peuvent être bruitées ou incomplètes à cause de capteurs défectueux. Une usine avec des capteurs mal calibrés faussera l'analyse prédictive.

Sans un nettoyage des données, le modèle de ML donnera des prévisions biaisées.

La préparation des données est donc cruciale : traitement des données manquantes, normalisation, etc.

Difficulté d'intégration dans les systèmes existants

Les systèmes industriels traditionnels ne sont pas toujours conçus pour utiliser les algorithmes de ML, nécessitant une infrastructure coûteuse et une refonte des processus de gestion de données.

La formation des équipes et le changement de culture d'entreprise représentent un défi majeur pour l'adoption à grande échelle du ML.

Une entreprise de production utilisant des technologies datées pourrait devoir remplacer l'ensemble de son infrastructure de collecte de données avant de bénéficier pleinement des modèles ML.

Sécurité et interprétabilité

Sécurité des données :

Les procédés industriels manipulent des informations sensibles. L'utilisation du ML via le cloud pose des questions de sécurité et de protection de la propriété intellectuelle.

Interprétabilité des modèles :

Les modèles de ML, comme les réseaux de neurones, sont souvent des "boîtes noires" difficiles à interpréter. Cela pose un défi pour les industries réglementées, où l'explicabilité est essentielle.

Exemple :

Dans le secteur pharmaceutique, un modèle ML qui optimise la production de médicaments sans expliquer ses décisions ne peut être utilisé sans une validation rigoureuse des autorités de régulation.

Conclusion et recommandations

Résumé des limites :

- Dépendance à la qualité des données
- Complexité des procédés
- Notion de GIGO
- Intégration difficile dans les systèmes existants
- Problèmes de sécurité et d'interprétabilité

Recommandations :

- Améliorer la collecte et la gestion des données pour limiter les effets de GIGO.
- Travailler sur l'interprétabilité des modèles pour une adoption plus large.
- Renforcer la sécurité des données afin d'assurer une utilisation en toute confiance des algorithmes de ML dans les processus critiques.

Conclusion générale

- **Transformation numérique** : L'intelligence artificielle révolutionne l'industrie des procédés en augmentant la productivité, optimisant les opérations, et améliorant la qualité des produits.
- **Bénéfices concrets** : Grâce au machine learning,on anticipe les pannes, améliore la maintenance prédictive et optimise en continu les procédés.
- **Défis et perspectives** : Malgré les avantages, des défis subsistent, notamment en termes de qualité des données, complexité des processus, et sécurité. La clé réside dans une meilleure gestion des données et l'amélioration de l'interprétabilité des modèles.
- **Recommandations** : Investir dans l'infrastructure de données et la formation des équipes pour garantir une adoption efficace des technologies IA.

Questions?