



# Introduction à l'IA et au machine learning

# Introduction IA & ML

## **Quelle est la différence entre Machine Learning et Intelligence Artificielle ?**

Le machine learning et l'intelligence artificielle sont souvent utilisés de manière interchangeable. Dans les deux cas, il s'agit de traiter d'énormes volumes de données, d'améliorer la prise de décision, de réaliser des analyses prédictives, etc. Et pourtant, ces deux notions ne sont pas synonymes.

L'intelligence artificielle fait référence à l'utilisation de diverses techniques permettant aux machines de développer des fonctions cognitives propres à l'intelligence humaine. Les principales étant la compréhension, la communication, la mémorisation, le raisonnement, l'adaptation et l'apprentissage autonome.

Or, en permettant aux algorithmes d'apprendre de manière autonome grâce à des ensembles de données, le machine learning participe justement au développement de cette IA. Mais ce n'est pas la seule technique utilisée. Il y a aussi le deep learning, la robotique, les systèmes experts, le traitement du langage naturel, etc.

=> La machine learning n'est donc qu'une branche de l'intelligence artificielle.

# Introduction IA & ML

Machine learning	Intelligence artificielle
L'apprentissage automatique est une technique permettant de développer l'IA.	L'IA est un concept visant à imiter l'intelligence humaine.
Le ML permet aux machines d'apprendre par elles-mêmes (sans raisonnement propre).	L'IA cherche à créer des modèles capables de raisonner par eux-mêmes.
Le ML permet aux machines de résoudre des problématiques spécifiques.	L'IA créer des modèles capables de résoudre des tâches complexes comme un humain.
Le ML n'utilise que des données structurées ou semi-structurées.	L'IA peut utiliser tous types de données.

# Machine Learning

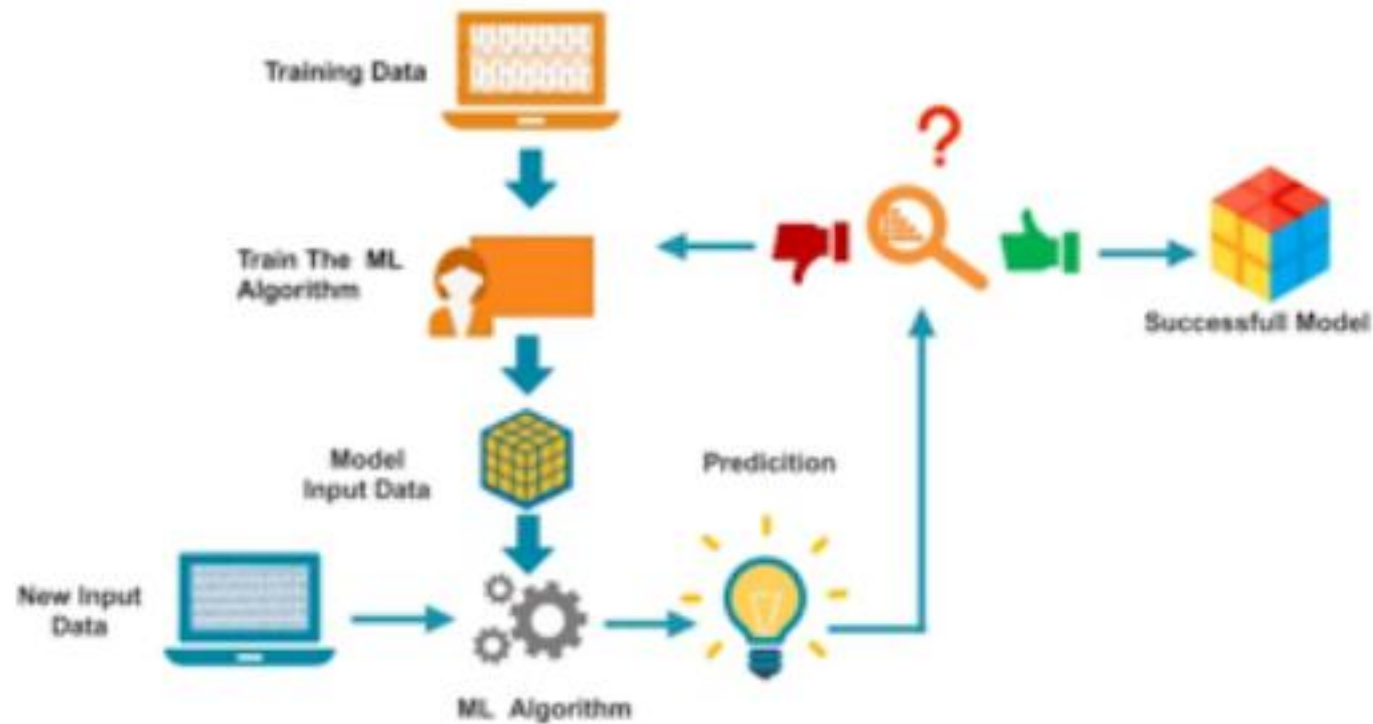
Le Machine Learning ou apprentissage automatique est un domaine scientifique, et plus particulièrement une sous-catégorie de l'intelligence artificielle. Elle consiste à laisser des algorithmes découvrir des « patterns », à savoir des motifs récurrents, dans les ensembles de données. Ces données peuvent être des chiffres, des mots, des images, des statistiques...

Tout ce qui peut être stocké numériquement peut servir de données pour le Machine Learning. En décelant les patterns dans ces données, les algorithmes apprennent et améliorent leurs performances dans l'exécution d'une tâche spécifique.

=> les algorithmes de Machine Learning apprennent de manière autonome à effectuer une tâche ou à réaliser des prédictions à partir de données et améliorent leurs performances au fil du temps. Une fois entraîné, l'algorithme pourra retrouver les patterns dans de nouvelles données.

# How does Machine Learning Work?

edureka!



# Comment fonctionne le Machine Learning ?

**Etape 1-** sélectionner et à préparer un ensemble de données d'entraînement. Ces données seront utilisées pour nourrir le modèle de Machine Learning pour apprendre à résoudre le problème pour lequel il est conçu.

Les données peuvent être labellisées, afin d'indiquer au modèle les caractéristiques qu'il devra identifier, ou pas et le modèle devra repérer et extraire les caractéristiques récurrentes de lui-même.

=> Dans les deux cas, les données doivent être soigneusement préparées, organisées et nettoyées. Dans le cas contraire, l'entraînement du modèle de Machine Learning risque d'être biaisé. Les résultats de ses futures prédictions seront directement impactés.

Etape 2: sélectionner un algorithme à exécuter sur l'ensemble de données d'entraînement. Le type d'algorithme à utiliser dépend du type et du volume de données d'entraînement et du type de problème à résoudre.

Etape3: Entraîner l'algorithme. Il s'agit d'un processus itératif. Des variables sont exécutées à travers l'algorithme, et les résultats sont comparés avec ceux qu'il aurait du produire. Les « poids » et le biais peuvent ensuite être ajustés pour accroître la précision du résultat.

=> On exécute ensuite de nouveau les variables jusqu'à ce que l'algorithme produise le résultat correct la plupart du temps. L'algorithme, ainsi entraîné, est le modèle de Machine Learning.

Etape4: l'utilisation et l'amélioration du modèle. On utilise le modèle sur de nouvelles données, dont la provenance dépend du problème à résoudre.

# Quels sont les principaux algorithmes de Machine Learning ?

Il existe une large variété d'algorithmes de Machine Learning. Certains sont toutefois plus couramment utilisés que d'autres.

On distingue trois techniques de Machine Learning : l'apprentissage supervisé, l'apprentissage non-supervisé, et l'apprentissage par renforcement

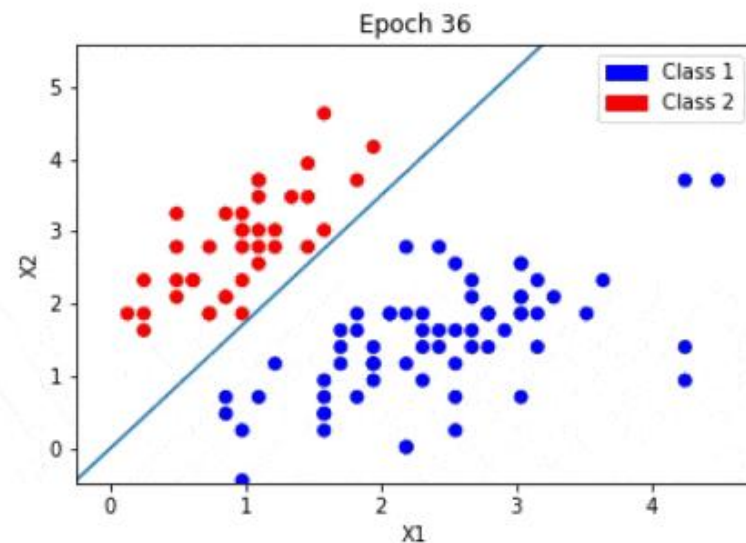
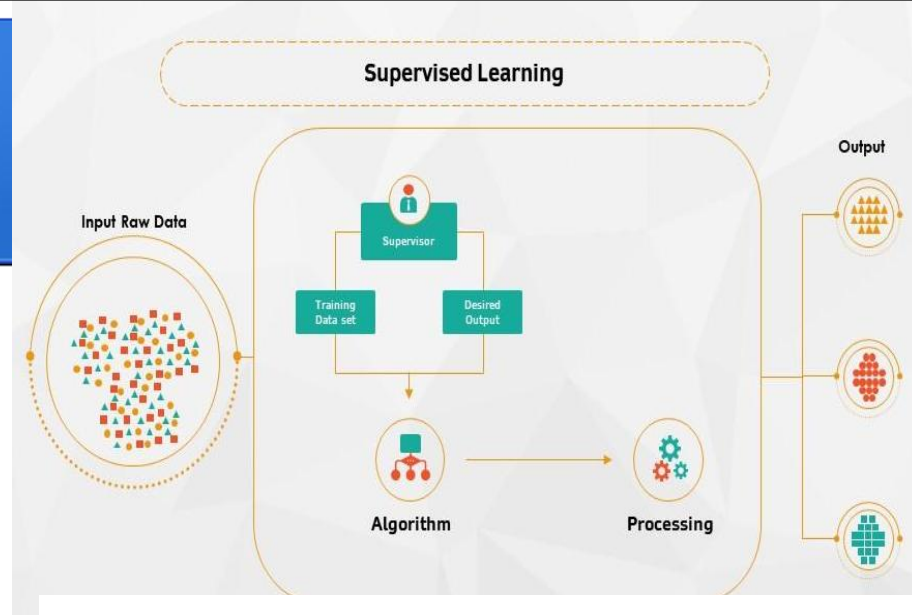


# L'apprentissage supervisé

Dans le cas de l'apprentissage supervisé, le plus courant, les données sont étiquetées afin d'indiquer à la machine quelles patterns elle doit rechercher.

Le système s'entraîne sur un ensemble de données étiquetées, avec les informations qu'il est censé déterminer. Les données peuvent même être déjà classifiées de la manière dont le système est supposé le faire.

Cette méthode nécessite moins de données d'entraînement que les autres, et facilite le processus d'entraînement puisque les résultats du modèle peuvent être comparés avec les données déjà étiquetées. Cependant, l'étiquetage des données peut se révéler onéreux. Un modèle peut aussi être biaisé à cause des données d'entraînement, ce qui impactera ses performances par la suite lors du traitement de nouvelles données.



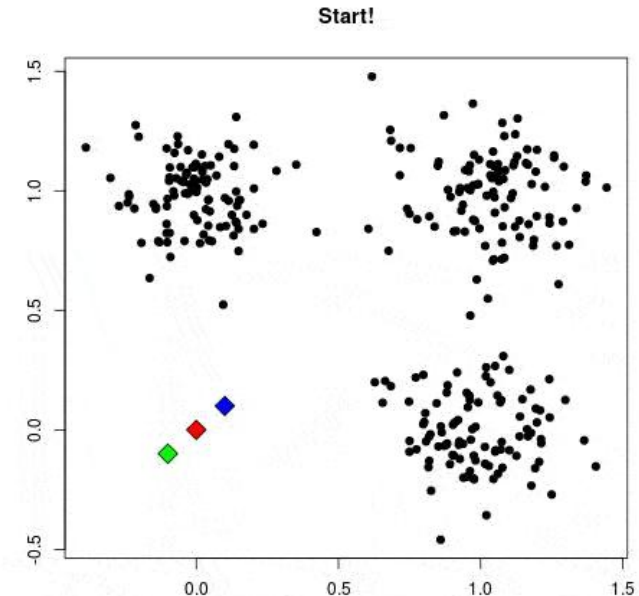
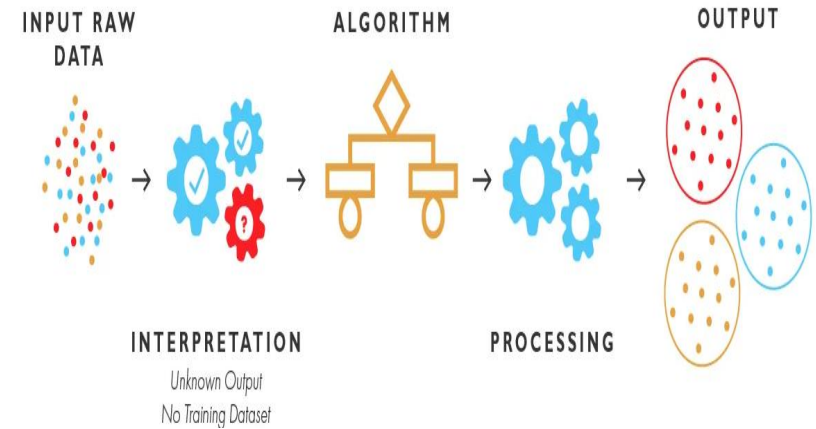


# L'apprentissage NON supervisé

Dans le cas de l'apprentissage non supervisé, les données n'ont pas d'étiquettes.

La machine se contente d'explorer les données à la recherche d'éventuelles patterns. Elle ingère de vastes quantités de données, et utilise des algorithmes pour en extraire des caractéristiques pertinentes requises pour étiqueter, trier et classifier les données en temps réel sans intervention humaine.

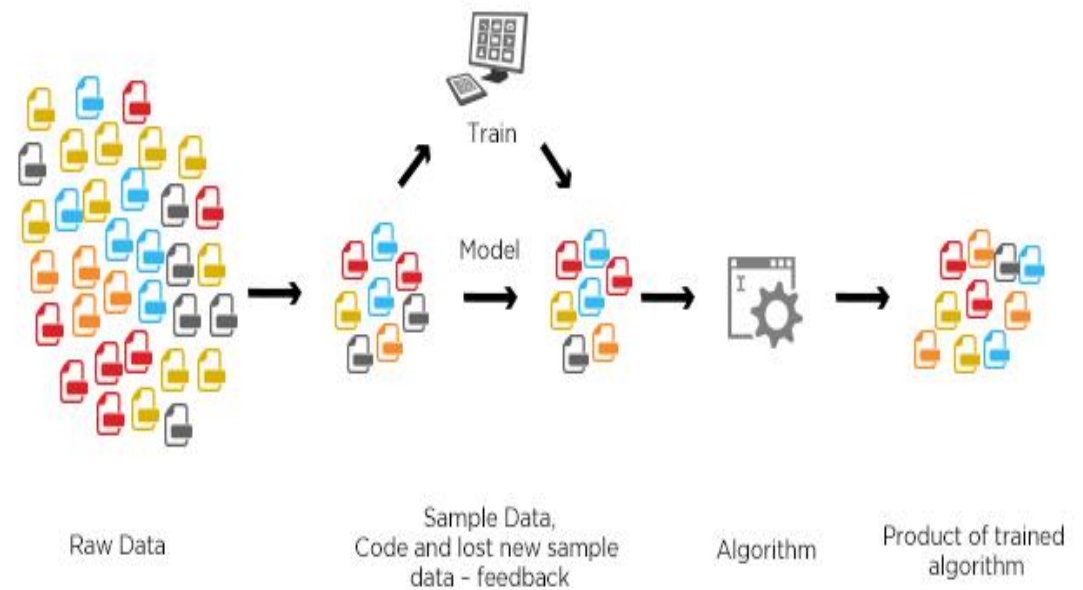
Plutôt que d'automatiser les décisions et les prédictions, cette approche permet d'identifier les patterns et les relations que les humains risquent de ne pas identifier dans les données. Cette technique n'est pas très populaire, car moins simple à appliquer. Elle est toutefois de plus en plus populaire dans le domaine de la cybersécurité.



# L'apprentissage SEMI supervisé

L'apprentissage « semi-supervisé » se situe entre les deux et offre un compromis entre apprentissage supervisé et non-supervisé. Pendant l'entraînement, un ensemble de données étiqueté de moindre envergure est utilisé pour guider la classification et l'extraction de caractéristiques à partir d'un ensemble plus large de données non étiquetées.

Cette approche s'avère utile dans les situations où le nombre de données étiquetées est insuffisant pour l'entraînement d'un algorithme supervisé. Elle permet de contourner le problème.



# L'apprentissage par Renforcement

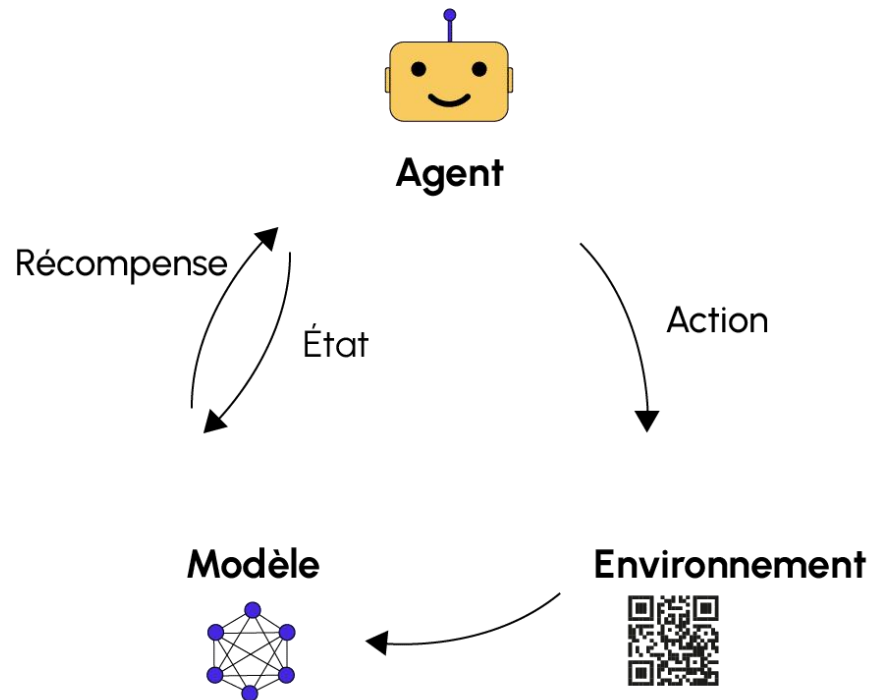
- \* l'apprentissage par renforcement consiste à laisser un algorithme apprendre de ses erreurs pour atteindre un objectif. L'algorithme essayera de nombreuses approches différentes pour tenter d'atteindre son but.
- \* En fonction de ses performances, il sera récompensé ou pénalisé pour l'inciter à poursuivre dans une voie ou à changer d'approche. Cette technique est notamment utilisée pour permettre à une IA de surpasser les humains dans les jeux.
- \* L'apprentissage par renforcement nécessite d'introduire un certain nombre de concepts et de métriques, dont les principaux sont les suivants :

**Agent** : système ou robot qui interagit et agit dans l'environnement ;

**Action  $a$**  : une action parmi l'ensemble des actions possible ;

**État  $s$**  : situation particulière dans laquelle l'agent se trouve ;

**Politique  $\pi$**  : stratégie qui définit le comportement l'agent



	APPRENTISSAGE SUPERVISÉ	APPRENTISSAGE NON-SUPERVISÉ	APPRENTISSAGE PAR RENFORCEMENT
DÉFINITION	L'algorithme apprend à partir de données labellisées	L'algorithme est entraîné à partir de données non labellisées sans indications particulières	L'algorithme interagit avec son environnement en réalisant des actions et en apprenant de ses erreurs et succès
TYPE DE PROBLÈMES	Régression et classification	Association et Clustering	Basés sur un système de récompense
TYPE DE DONNÉES	Données labellisées	Données non labellisées	Pas de données fournies au préalable
APPROCHE	Étudie les relations sous-jacentes qui lient les données en entrée aux labels	Découvre les motifs communs au sein des données d'entrée	Apprend une stratégie de comportement en fonction d'expériences passées et des récompenses perçues

# Deep Learning

**Le Machine Learning** : c'est un ensemble de techniques donnant la capacité aux machines d'apprendre automatiquement un ensemble de règles à partir de données. Contrairement à la programmation qui consiste en l'exécution de règles prédéterminées.

**Le Deep Learning ou apprentissage profond** : c'est une technique de machine learning reposant sur le modèle des réseaux neurones: des dizaines voire des centaines de couches de neurones sont empilées pour apporter une plus grande complexité à l'établissement des règles.

