

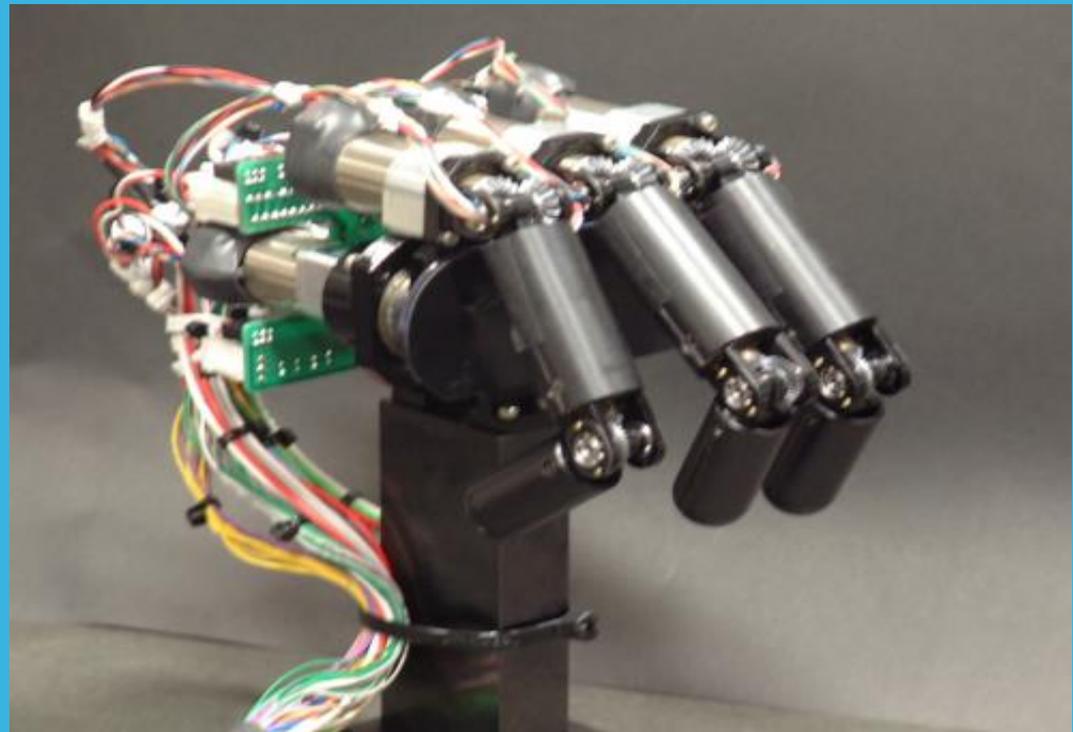


CHAPITRE 4:

LA ROBOTIQUE

LA MATIÈRE : IMAGERIE INDUSTRIELLE

Présentée par :
Mme M. Chibani



SOMMAIRE

1. Description : Différents type de robots
2. Domaine d'application
3. Capteurs et actionneurs : principalement les caméras embarquées
4. **L'apprentissage**
5. Vision 3D

L'apprentissage automatique (Machine Learning)

«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)», la définition du Machine Learning selon son inventeur Arthur Samuel, un mathématicien américain qui a développé un programme capable de s'auto-apprendre à jouer aux dames en 1959.



"Machine Learning is the science of getting computers to learn without being explicitly programmed."

Arthur Samuel, 1959.

METHODES D'APPRENTISSAGE

- Pour donner à un ordinateur la capacité d'apprendre, nous utilisons des méthodes d'apprentissage fortement inspirées de la manière dont nous, les humains, apprenons à faire les choses. Ces méthodes comprennent :
 - L'apprentissage supervisé (Supervised Learning)
 - L'apprentissage non supervisé (Unsupervised Learning)
 - L'apprentissage par renforcement (Reinforcement Learning)



METHODES D'APPRENTISSAGE – Apprentissage supervisé

- On parle **d'apprentissage supervisé** lorsqu'une machine reçoit un grand nombre d'exemples à étudier.
- Pour maîtriser l'apprentissage supervisé, il est essentiel de comprendre et de connaître les 4 concepts suivants :
 1. **Le Dataset**
 2. **Le Modèle et ses paramètres**
 3. **La Fonction Coût**
 4. **L'Algorithme d'apprentissage**



METHODES D'APPRENTISSAGE – Apprentissage supervisé

Notion 1 : Apprendre à partir d'exemples (Dataset)

- On parle d'apprentissage supervisé lorsqu'une machine reçoit un grand nombre d'exemples (\mathbf{x}, \mathbf{y}) dans le but d'apprendre la relation entre \mathbf{x} et \mathbf{y} .
- Dans l'apprentissage automatique, nous compilons ces exemples (\mathbf{x}, \mathbf{y}) dans un tableau que nous appelons Dataset (ensemble de données):



METHODES D'APPRENTISSAGE – Apprentissage supervisé

- La variable y est appelée cible. C'est la valeur que nous essayons de prédire.
- La variable x est appelée caractéristique. Un facteur influence la valeur de y , et nous avons généralement de nombreuses caractéristiques (x_1, x_2, \dots) dans notre ensemble de données que nous regroupons dans une matrice X .



METHODES D'APPRENTISSAGE – Apprentissage supervisé

- Vous trouverez ci-dessous un ensemble d'exemples d'appartements avec leur prix y et certaines de leurs caractéristiques.

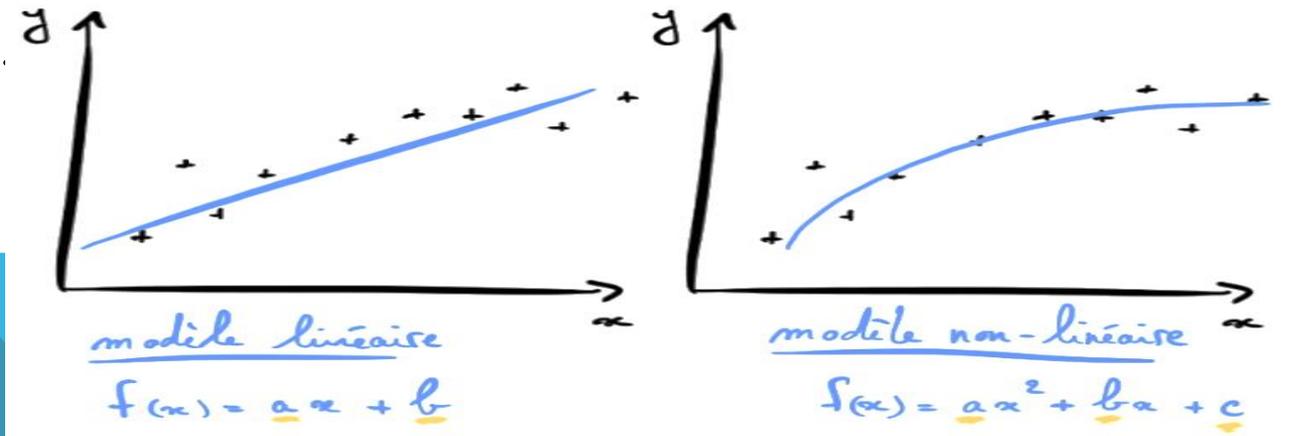
Target y	Features		
Prix	x_1 Surface m2	x_2 N chambres	x_3 Qualité
€ 313,000.00	124	3	1.5
€ 2,384,000.00	339	5	2.5
€ 342,000.00	179	3	2
€ 420,000.00	186	3	2.25
€ 550,000.00	180	4	2.5
€ 490,000.00	82	2	1
€ 335,000.00	125	2	2



METHODES D'APPRENTISSAGE – Apprentissage supervisé

Notion 2 : Développement d'un modèle à partir d'un ensemble de données

Dans le cadre de l'apprentissage automatique, nous développons un modèle à partir de cet ensemble de données. Il peut s'agir d'un modèle linéaire, comme illustré à gauche, ou d'un modèle non linéaire, comme illustré à droite.

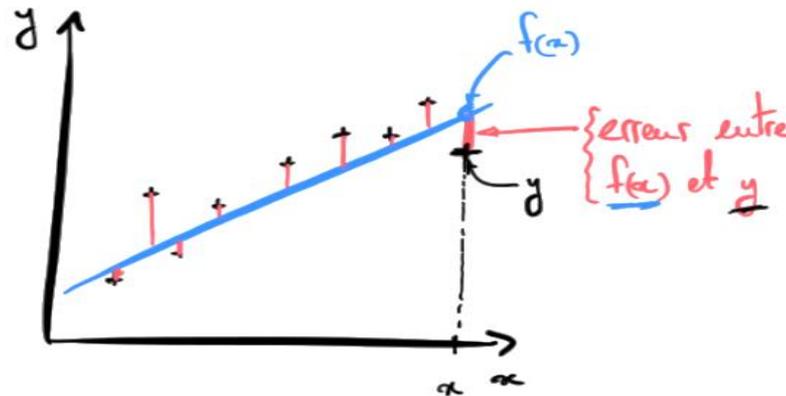


Nous définissons a, b, c comme les paramètres d'un modèle.

METHODES D'APPRENTISSAGE – Apprentissage supervisé

Notion 3 : Les erreurs de notre modèle - la Fonction Coût

Un modèle renvoie des erreurs par rapport à notre ensemble de données. Nous appelons l'ensemble des erreurs la fonction coût.

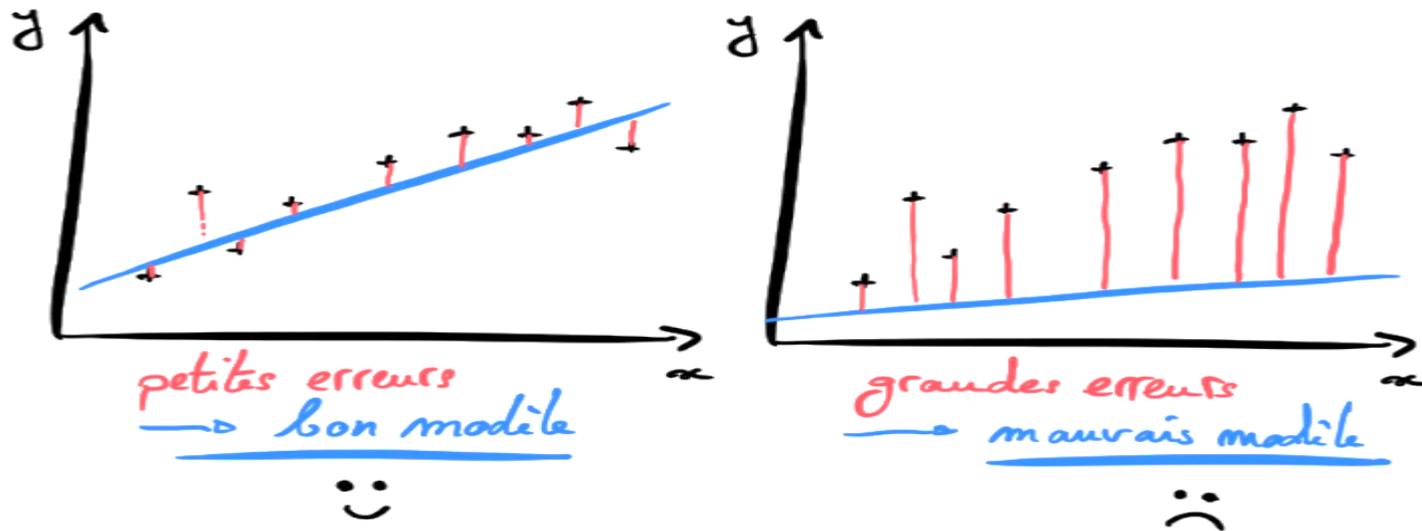


Fonction Coût = l'ensemble des erreurs.



METHODES D'APPRENTISSAGE – Apprentissage supervisé

Avoir un bon modèle signifie avoir un modèle qui nous donne de petites erreurs, et donc une petite fonction coût.



METHODES D'APPRENTISSAGE – Apprentissage supervisé

Notion 4 : Apprendre, c'est minimiser la Fonction Coût

Ainsi, l'objectif central de l'apprentissage supervisé est de trouver les paramètres du modèle qui minimisent la fonction coût. Pour ce faire, nous utilisons un algorithme d'apprentissage, l'exemple le plus courant étant l'algorithme de **Gradient Descent**.



METHODES D'APPRENTISSAGE – Apprentissage supervisé



$x = 100 \text{ m}^2$
 $y = 300,000 \text{ €}$

Inputs x
 100 m^2

Calcul

Output $f(x)$
 100 €

Mesure
l'erreur entre
 y et $f(x)$

Erreur
 $299,900 \text{ €}$

Modifie f pour
minimiser
erreur

Nouveau modele f

Data

Phase d'Entrainement

METHODES D'APPRENTISSAGE – Apprentissage supervisé

- L'apprentissage supervisé permet de développer des modèles pour résoudre deux types de problèmes :
 - Les problèmes de Régression
 - Les problèmes de Classification
- Dans les problèmes de régression, nous essayons de prédire la valeur d'une variable continue, c'est-à-dire une variable qui peut prendre un nombre infini de valeurs. Par exemple : Prédire le prix d'un appartement (y) en fonction de sa surface habitable (x).

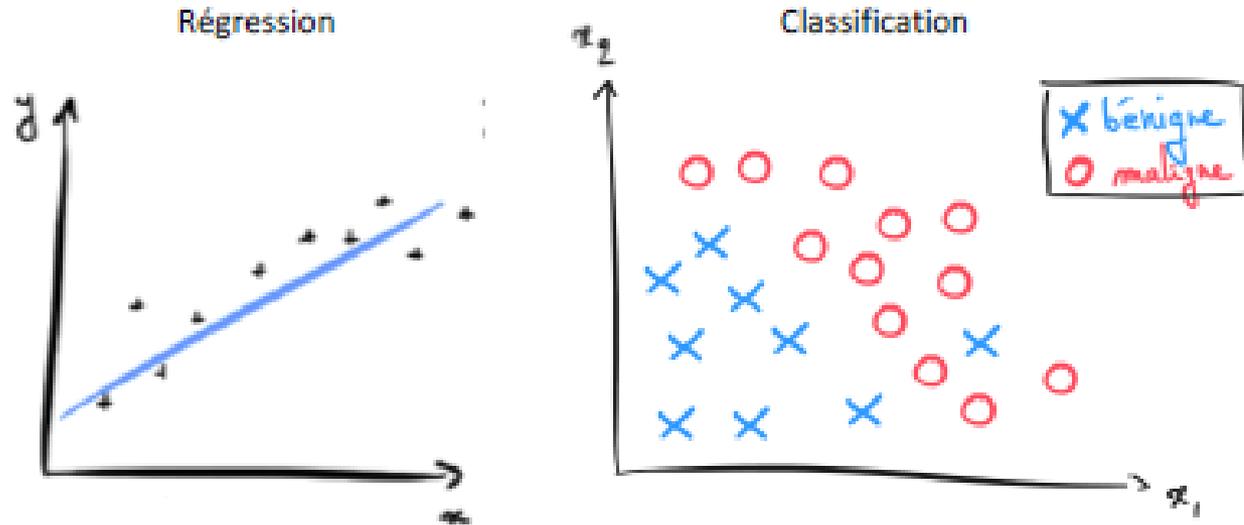


METHODES D'APPRENTISSAGE – Apprentissage supervisé

Dans un problème de classification, nous cherchons à classer un objet dans différentes classes, c'est-à-dire que nous cherchons à prédire la valeur d'une variable discrète (qui ne prend qu'un nombre fini de valeurs). Par exemple : Prédire si un courriel est un spam (*classe* $y = 1$) ou non (*classe* $y = 0$) en fonction du nombre de liens présents dans le courriel (x).



METHODES D'APPRENTISSAGE – Apprentissage supervisé



METHODES D'APPRENTISSAGE – Apprentissage supervisé

Les Algorithmes de d'apprentissage supervisé

Les algorithmes d'apprentissage supervisé sont légions: K-Nearest Neighbour, Régression Linéaire, Régression Logistique, Support Vector Machine, Neural Network, Naive Bayes, Random Forest, Adaboost, etc.



METHODES D'APPRENTISSAGE – Apprentissage non-supervisé

- L'apprentissage supervisé consiste à enseigner à la machine des choses que nous connaissons déjà, en construisant à l'avance un ensemble de données contenant X questions et y réponses.
- Que se passe-t-il si vous disposez d'un ensemble de données sans valeur y ? Et si vous voulez que la machine vous aide à enrichir vos connaissances en apprenant des choses que vous ne connaissez pas ?



METHODES D'APPRENTISSAGE – Apprentissage non-supervisé

- Dans l'apprentissage non supervisé, on nous donne un ensemble de données (x) sans valeur (y), et la machine apprend à reconnaître les structures dans les données (x) qui lui sont présentées.
- Cela permet de regrouper les données en clusters, de détecter les anomalies ou de réduire les dimensions des données en compilant les dimensions entre elles.



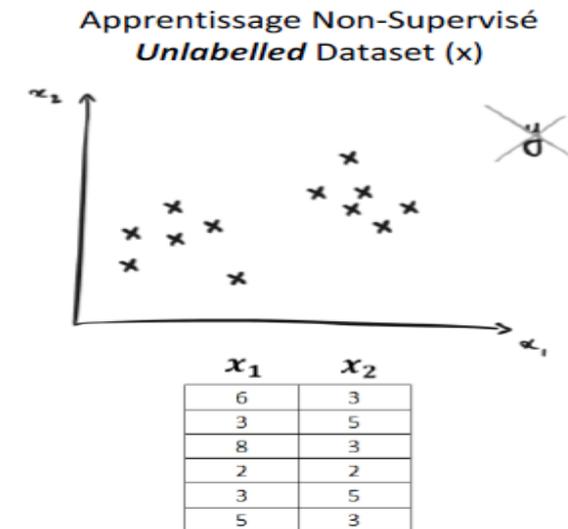
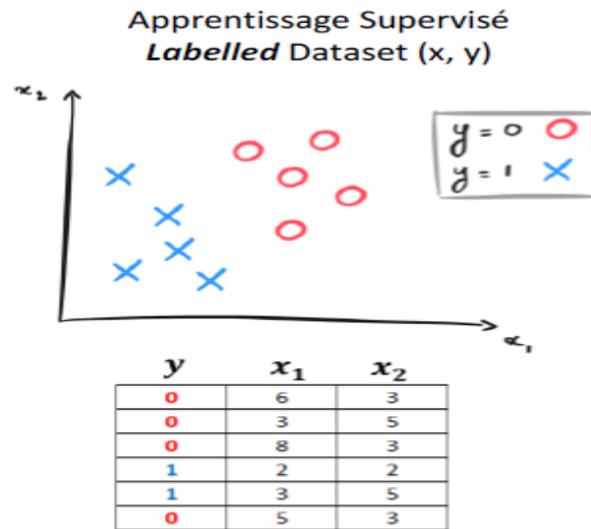
METHODES D'APPRENTISSAGE – Apprentissage non-supervisé

- Le **K-Mean Clustering** est probablement l'algorithme le plus populaire pour les problèmes de clustering (regroupement de données en fonction de leur structure commune). Il est souvent utilisé en marketing pour cibler des groupes de clients similaires pour certaines campagnes publicitaires.
- Dans l'apprentissage supervisé, la machine reçoit un ensemble de données où les exemples (x) sont étiquetés avec une valeur (y) (nous appelons cela un ensemble de données étiqueté). Il est alors possible de trouver une relation générale qui relie (x) à (y).



METHODES D'APPRENTISSAGE – Apprentissage non-supervisé

- Dans l'apprentissage non supervisé, nous ne pouvons pas faire cela, car il manque la variable (y) dans notre ensemble de données. Il est donc non étiqueté



METHODES D'APPRENTISSAGE – Apprentissage non-



Labelled Dataset
(données étiquetées)
Supervised Learning



Unlabelled Dataset
(données non-étiquetées)
Unsupervised Learning

METHODES D'APPRENTISSAGE – Apprentissage non-supervisé

Les Algorithmes de l'apprentissage non supervisé

Les algorithmes les plus couramment utilisés dans l'apprentissage non supervisé sont les suivants: K-mean Clustering, Anomaly Detection Systems, Principal Component Analysis, Auto-encoder Neural Networks, Generative Adversarial Network, Manifold Learning, etc

METHODES D'APPRENTISSAGE – Apprentissage par Renforcement

- L'apprentissage par renforcement consiste à laisser la machine apprendre à effectuer une tâche (par exemple, piloter un mini-drone) en la laissant s'entraîner seule. Lorsque la machine réussit, elle reçoit un bonus. Lorsqu'elle échoue, elle reçoit une pénalité.
- Nous développons un programme qui force la machine à maximiser ses bonus, et la machine analyse ensuite ses propres erreurs passées afin de s'améliorer au fil du temps. C'est un peu comme apprendre à faire du vélo : au début, on ne réussit pas du tout, mais au fur et à mesure qu'on apprend à rouler, on développe intrinsèquement son équilibre pour éviter les erreurs du passé qui nous ont fait chuter !



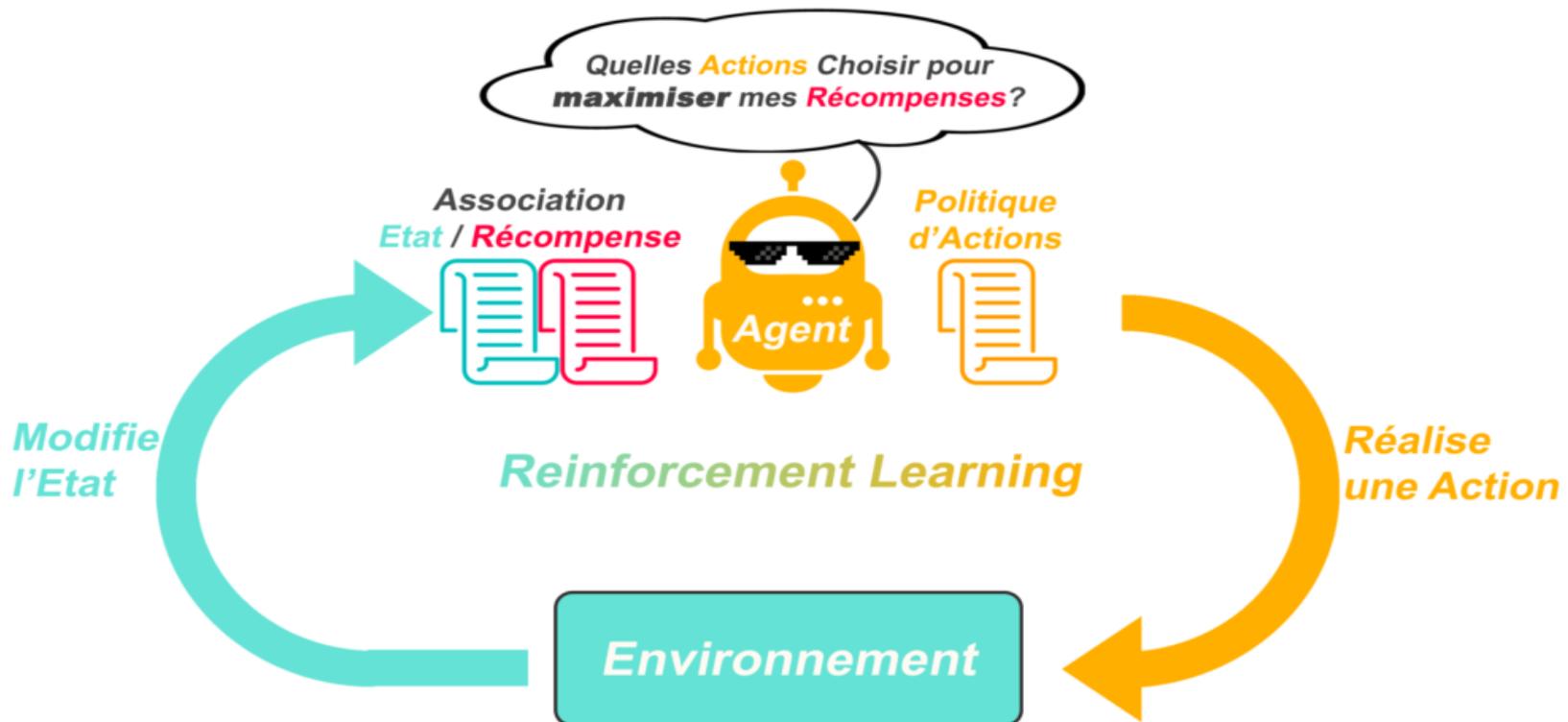
METHODES D'APPRENTISSAGE – Apprentissage par Renforcement

- C'est la discipline qui se rapproche le plus de l'idée d'un "robot intelligent" capable de faire ce que nous faisons : conduire une voiture, résoudre des problèmes, jouer aux échecs, etc
- Dans cette technique, la machine génère elle-même ses propres expériences. En tant qu'agent, elle est libre d'entreprendre des actions dans un environnement. En fonction de l'action entreprise, l'environnement modifie l'état de l'agent et lui donne une récompense positive ou négative associée à cet état. Cela constitue une nouvelle expérience.



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement

- Au sein du programme, un algorithme élabore une politique d'actions visant à maximiser le nombre de récompenses positives.



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement

Exemple pour illustrer l'apprentissage par renforcement

- Imaginons que notre agent soit une voiture. À tout moment, la voiture peut choisir d'agir sur son environnement, qui est la route : elle peut freiner, accélérer, tourner à gauche ou à droite. Si l'agent choisit de tourner à gauche alors qu'il y a un mur devant lui, l'état de la voiture passera de "Normal" à "Accidentel", et l'environnement donnera à l'agent une récompense négative. L'algorithme, qui cherche à maximiser les récompenses, modifiera donc sa politique d'action afin de ne pas répéter un tel "accident".
- Le programme apprend à conduire : il améliore sa performance sans accident grâce à une nouvelle expérience.



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement

Les algorithmes d'apprentissage par renforcement

Les algorithmes d'apprentissage par renforcement les plus répandus sont les suivants: La recherche de Monte Carlo, Temporal Difference Learning, SARSA, et le Q-Learning.



METHODES D'APPRENTISSAGE-Structure d'apprentissage par renforcement

Agent

L'agent est la partie la plus importante de l'apprentissage par renforcement, car il contient l'intelligence nécessaire pour prendre des décisions et recommander l'action optimale dans une situation donnée, et il interagit avec l'environnement dans lequel il est placé. Ces interactions se produisent de manière séquentielle au fil du temps.



METHODES D'APPRENTISSAGE-Structure d'apprentissage par renforcement

Politique (policy)

Une politique est une fonction qui associe un état donné aux probabilités de sélection de chaque action possible dans cet état. Nous utiliserons le symbole Π pour désigner une politique. En ce qui concerne les politiques, nous disons formellement qu'un agent "suit une politique".



METHODES D'APPRENTISSAGE-Structure d'apprentissage par renforcement

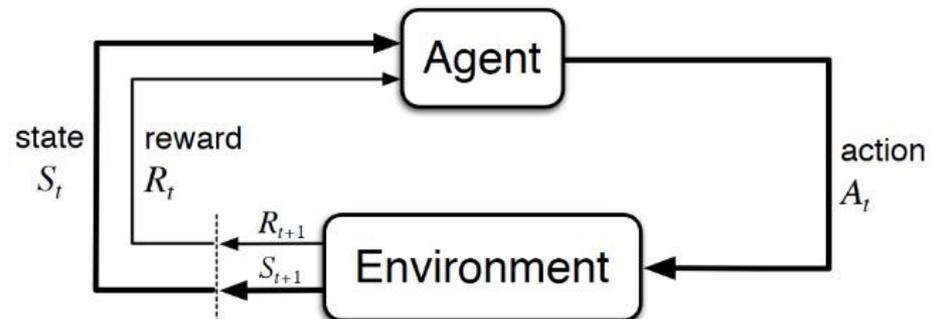
La séquence état-action-récompense (state-action-reward)

À chaque étape, l'agent obtient une représentation de l'état de l'environnement. Sur la base de cette représentation, l'agent choisit une action à exécuter, l'environnement entre alors dans un nouvel état et l'agent reçoit une récompense (ou une pénalité est une récompense négative) en conséquence de l'action précédente.



METHODES D'APPRENTISSAGE-Structure d'apprentissage par renforcement

- Au moment t , l'environnement est dans l'état S_t .
- L'agent observe l'état actuel et sélectionne l'action A_t .
- L'environnement passe à l'état S_{t+1} et accorde à l'agent la récompense R_{t+1} .
- Ce processus recommence ensuite pour l'étape suivante $t+1$. Notez que $t+1$ n'est plus dans le futur, mais est maintenant le présent. Lorsque nous traversons la ligne pointillée en bas à gauche, le diagramme montre que l'instant j se transforme en instant t , de sorte que S_{t+1} et R_{t+1} sont maintenant S_t et R_t .



METHODES D'APPRENTISSAGE–La récompense

La récompense est une valeur numérique reçue par l'agent de l'environnement en réponse directe à ses actions. Toutes les actions génèrent des récompenses, qui peuvent être divisées en trois types : les récompenses positives, qui soulignent une action souhaitée; les récompenses négatives, qui soulignent une action dont l'agent devrait s'écarter ; et zéro, qui signifie que l'agent n'a rien fait de spécial.



METHODES D'APPRENTISSAGE–La récompense

Rendement

Pour que l'agent atteigne son but, il doit choisir les meilleures actions pour maximiser ses récompenses, pour cela non seulement la récompense immédiate est prise en compte, mais aussi les récompenses futures cumulées à chaque instant t . Nous définissons le rendement G à l'étape t comme. $G_t = R_{t+1} + R_{t+2} + R_{t+3} + \dots + R_T$ Où T est la dernière étape.



METHODES D'APPRENTISSAGE–La récompense

Rendement attendu

Dans certains cas l'état de l'environnement peut-être une fonction continue, et comme chaque état donne une récompense, le rendement risque de prendre des valeurs grandes, pour contrôler l'importance des récompenses futures nous introduisons le concept d'actualisation des récompenses. Ainsi, nous définissons le

rendement actualisé comme : $G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$

le taux d'actualisation γ (discount rate) est un nombre compris entre 0 et 1.



METHODES D'APPRENTISSAGE–La récompense

Épisode

L'interaction entre l'agent et l'environnement se décompose naturellement en sous-séquences, appelées épisodes, chaque épisode étant un ensemble d'étapes.

L'épisode se termine dans un état terminal T , qui est suivi par la réinitialisation de l'environnement à un état de départ standard ou à un échantillon aléatoire à partir d'une distribution d'états de départ possibles. L'épisode suivant commence alors indépendamment de la façon dont l'épisode précédent s'est terminé.



METHODES D'APPRENTISSAGE–La fonction action–valeur (la fonction–Q)

la fonction action-valeur pour la politique π , dénommée Q_π , nous indique à quel point il est bon pour l'agent de prendre une action donnée à partir d'un état donné tout en suivant la politique π (la valeur d'une action sous π). La valeur de l'action a dans le cadre de la politique π est le rendement attendu à partir de l'état s au moment t , en prenant l'action a et en suivant la politique par la suite.



METHODES D'APPRENTISSAGE–Le dilemme exploration/exploitation

Pour former une valeur optimale pour la fonction action-valeur (fonction-Q), nous pouvons commencer par des valeurs initialisées de façon aléatoire ou une initialisation fixe/heuristique par défaut. Les valeurs de la fonction-Q sont améliorées par rapport aux valeurs par défaut en effectuant des expériences sur les données, scénarios et/ou épisodes d'entraînement.



METHODES D'APPRENTISSAGE–Le dilemme exploration/exploitation

- L'agent essaie de nombreuses actions différentes dans de nombreux états différents, afin d'apprendre toutes les possibilités disponibles et de trouver la trajectoire qui maximisera sa récompense totale, c'est ce que nous appelons «**exploration**», car l'agent explore son environnement.
- En revanche, si l'agent ne fait qu'explorer, il ne maximisera jamais la récompense totale. Il doit également utiliser les informations apprises, c'est ce que nous appelons «**exploitation**», car l'agent exploite ses connaissances pour maximiser les récompenses qu'il reçoit.



METHODES D'APPRENTISSAGE–Le dilemme exploration/exploitation

- Le compromis entre l'exploration et l'exploitation est l'un des plus grands défis des problèmes d'apprentissage par renforcement, car les deux doivent être équilibrés afin de permettre à l'agent à la fois d'explorer suffisamment l'environnement, mais aussi d'exploiter ce qu'il a appris et de répéter la trajectoire la plus profitable trouvée.
- Pour obtenir un équilibre entre l'exploitation et l'exploration, nous utilisons ce que nous appelons la stratégie Epsilon Greedy. Avec cette stratégie, nous définissons un taux d'exploration ϵ que nous fixons au départ à 1. Ce taux d'exploration est la probabilité que notre agent explore l'environnement plutôt que de l'exploiter.



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement : Q-Learning

- Q-Learning est un algorithme d'apprentissage par renforcement, considéré comme l'un des plus fondamentaux. Dans sa forme la plus simplifiée, il utilise un tableau appelé **tableau-Q** pour stocker toutes **les valeurs-Q** de toutes les paires **état-action** possibles. Il actualise ce tableau en utilisant **l'équation de Bellman**, tandis que la sélection des actions est généralement effectuée avec une politique **ϵ -greedy**.



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement : Q-Learning

Itération de la valeur

L'algorithme Q-Learning actualise les valeurs-Q pour chaque paire état-action en utilisant l'équation de Bellman jusqu'à ce que la fonction-Q converge vers la fonction-Q optimale Q_* . Cette approche est appelée itération des valeurs. Comme l'agent ne sait rien de l'environnement ou des rendements attendues pour une paire état-action, toutes les valeurs-Q du tableau sont d'abord initialisées à zéro. Au fil du temps, lorsque l'agent effectue plusieurs épisodes, les valeurs-Q produites pour les paires état-action que l'agent expérimente seront utilisées pour mettre à jour les valeurs-Q stockées dans le tableau-Q.



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement : Q-Learning

La forme du tableau-Q

		Les actions			
		Action 1	Action 2	Action 3	...
Les états	S_0	0	0	0	0
	S_1	0	0	0	0
	S_2	0	0	0	0
	⋮	0	0	0	0



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement : Q-Learning

Le taux d'apprentissage

- Le taux d'apprentissage est un nombre compris entre 0 et 1 , qui peut être considéré comme la vitesse à laquelle l'agent abandonne la précédente valeur-Q dans le tableau-Q pour la nouvelle valeur-Q.
- Nous indiquerons le taux d'apprentissage avec le symbole α , plus le taux d'apprentissage est élevé, plus l'agent adoptera rapidement la nouvelle valeur-Q.



METHODES D'APPRENTISSAGE – Apprentissage par Renforcement : Q-Learning

Actualisation de la valeur-Q

- Pour mettre à jour la valeur-Q de toute action effectuée à partir de l'état précédent, nous utilisons l'équation d'optimalité de Bellman.
- La formule de calcul de la nouvelle valeur-Q pour la paire état-action (s, a) à l'étape t est la suivante :
- Où s' est l'état suivant et a' toute action qui peut être exécuter à partir de cet état.

$$Q^{new}(s, a) = (1 - \alpha) \underbrace{Q(s, a)}_{\substack{\text{ancienne} \\ \text{valeur}}} + \alpha \overbrace{\left(R_{t+1} + \gamma \max_{a'} Q_*(s', a') \right)}^{\substack{\text{nouvelle} \\ \text{valeur}}}$$

METHODES D'APPRENTISSAGE – Apprentissage par Renforcement : L'algorithme Q-learning

- 1 Initialisation : nombre des épisodes, nombre des étapes par épisode T , taux d'actualisation γ , taux d'apprentissage α , taux d'exploration ε ;
- 2 Créer un tableau-Q de taille nombre d'actions par nombre d'états, et initialisé toutes ses valeurs à zéro;
- 3 **tant que** *Nombre d'épisode* < *nombre d'épisode maximal* **faire**
 - 4 Commencer par l'état initial;
 - 5 **tant que** *Nombre d'étape* t < *nombre d'étape maximal* T **faire**
 - 6 Sélectionnez une action par exploration (action aléatoire) ou exploitation (l'action de la valeur-Q maximale) en suivant la stratégie Epsilon Greedy;
 - 7 Exécutez l'action et observez la récompense et l'état suivant;
 - 8 Actualiser la valeur-Q : $Q^{new}(s, a) = (1 - \alpha) Q(s, a) + \alpha Q_*(s, a)$;
 - 9 Passer à l'état suivant;
 - 10 Actualiser le taux d'exploration ε ;
 - 11 Sortir de la boucle si le but est atteint;

SOMMAIRE

1. Description : Différents type de robots
2. Domaine d'application
3. Capteurs et actionneurs : principalement les caméras embarquées
4. L'apprentissage
- 5. Vision 3D**

1. Vision 3D-INTRODUCTION

- Les systèmes de vision robotiques doivent reproduire le système de vision humain.
- La méthode, fidèle à la vision humaine, est la vision **stéréoscopique** ou **stéréovision**.
- La stéréovision est l'ensemble de tous les procédés utilisés pour déterminer les informations de profondeur d'une scène à partir de plusieurs images (deux ou plus), prises sous des angles de vue différents.



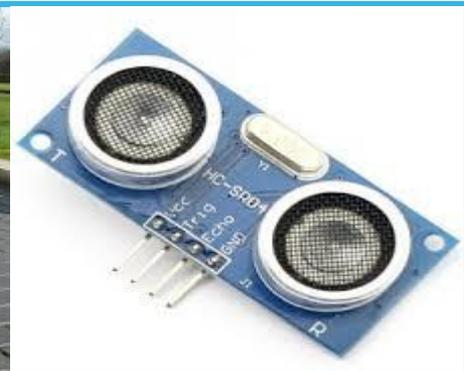
1. Vision 3D-INTRODUCTION

- Le système stéréoscopique le plus répandu et le plus simple est le système stéréoscopique binoculaire, qui n'utilise que deux images.
- La stéréovision trinoculaire est l'utilisation de trois caméras pour éviter certaines ambiguïtés dans le cas d'occultations.

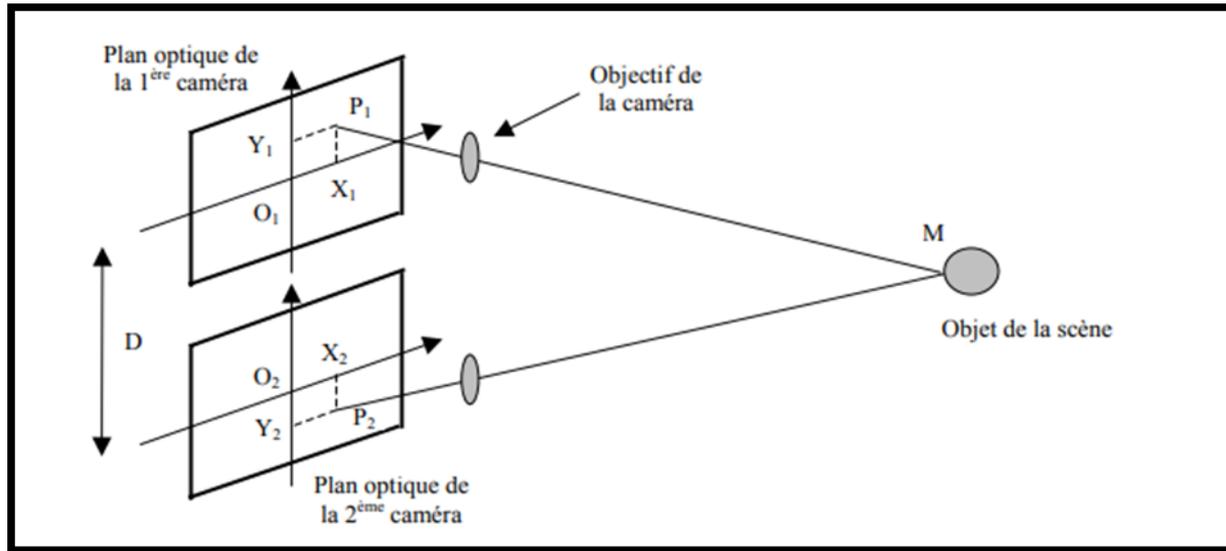


1. Vision 3D–Principe

- La stéréovision est une méthode de vision passive inspirée de la vision humaine basée sur l'acquisition de deux images, elle permet d'obtenir l'information de relief d'une scène à partir de deux projections bidimensionnelles de la même scène.
- La stéréovision se déroule en trois étapes successives :
 - **Calibrage**
 - **Appariement (mise en correspondance)**
 - **La reconstruction 3D**



1. Vision 3D-Principe



Principe de la stéréovision



1. Vision 3D–Le calibrage

- Il permet de trouver les paramètres des capteurs. Tout d'abord, il faut savoir comment former les images prises par les caméras du système stéréoscopique. En d'autres termes, il faut connaître le modèle géométrique des caméras.
- L'étalonnage de la caméra consiste à déterminer analytiquement la relation entre un point dans l'espace tridimensionnel et sa projection sur l'image de la caméra. Il s'agit de calibrer les deux caméras pour obtenir deux matrices : la matrice des paramètres intrinsèques et la matrice des paramètres extrinsèques, qui seront utilisées ultérieurement pour la rectification et la reconstruction.



1. Vision 3D–La mise en correspondance

- C'est la partie la plus difficile du processus de stéréovision.
- L'objectif de cette étape est de trouver les primitives homologues dans les deux images. En d'autres termes, l'appariement peut se faire soit entre les primitives extraites des images, soit entre tous les points des images. Il s'agit donc de trouver, pour un point donné dans une image, son point correspondant dans l'autre image.



1. Vision 3D–La reconstruction 3D

- La reconstruction 3D est l'étape finale du processus et constitue un sujet de recherche très actif dans le domaine de la vision par ordinateur.
- Grâce aux paramètres et aux correspondances des capteurs, un modèle 3D de la scène est reproduit par triangulation.

