

Apprentissage Automatique Numérique

Introduction

Loïc Barrault

loic.barrault@univ-lemans.fr
Le Mans Université

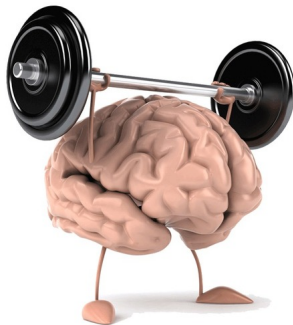
Organisation

- Intervenants : Nathalie Camelin, Loïc Barrault et Sylvain Meignier
- Séances:
 - 9 CM de 2h
 - 2 TD de 2h
 - 6 TPs de 3h
- Modalités d'évaluation
 - Session 1
 - Un examen de 2h coefficient 3
 - Une note de TP coefficient 2
 - Session 2
 - Un examen de 2h coefficient 3
 - Note de TP reportée (non rattrapable)
 - Langage de programmation : Python

Apprentissage et Intelligence artificielle

Intelligence ?

La capacités mentales et cognitives d'un individu à résoudre un problème ou à s'adapter à son environnement.



Un ordinateur peut-il être intelligent ?

Construire une machine intelligente :

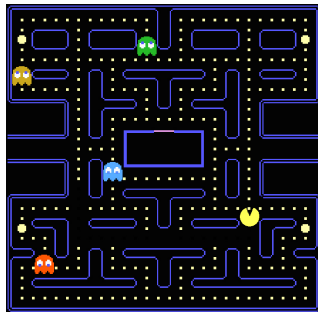
- Un vieux rêve de l'humanité
 - Machines capables d'interagir de façon naturelle avec les humains
- A.I. Intelligence Artificielle - Steven Spielberg 2001
- Domaine de recherche : **Intelligence Artificielle**



Apprentissage ?

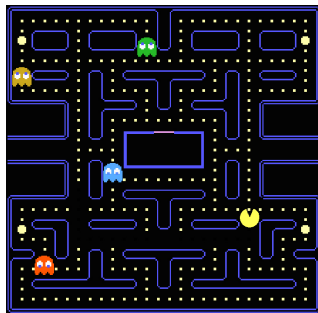
- L'acquisition d'un **savoir-faire**
- L'acquisition ou la modification de **la représentation** d'un environnement pour réaliser des **interactions efficaces**
- La capacité de **prise de décision** dans une **situation nouvelle**
- Les humains effectuent ces tâches de façon inconsciente, en temps réel et sans effort explicite
- Dans le monde animal : apprentissage par **imitation**

Pacman



- Est-ce que les adversaires sont intelligents ?

Pacman



- Est-ce que les adversaires sont intelligents ?
- Stratégie de déplacement, travail collaboratif, ...

Intelligence Collective



- Plusieurs **individus indépendants** collaborent pour résoudre **conjointement** une tâche complexe sans supervision

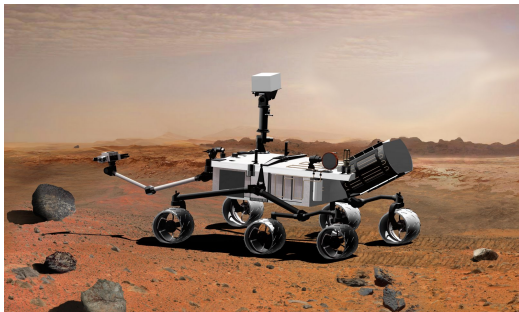
Intelligence Collective



- Plusieurs **individus indépendants** collaborent pour résoudre **conjointement** une tâche complexe sans supervision

⇒ **swarm intelligence**

Exemples d'Applications en IA : NASA / ESA



- Robots autonomes, par exemple « mars explorer »
- Durée d'une transmission terre/mars : 11 minutes
- Le robot doit être capable de reconnaître des objets, de prendre des décisions de façon **autonome**, etc

DARPA Grand Challenge



Chronicle 2 Scott Rogers



Chronicle 2 Scott Rogers

- Concours organisé par la DARPA depuis 2004 (prix de 2 millions dollars)
- Un véhicule doit parcourir une distance avec des obstacles de façon complètement autonome
- Parcours de 240 km dans le désert de Mojave avec de nombreuses courbes, descente sinueuses en montagne
- 2004 : aucun véhicule n'allait plus loin que 12kms
- 2005 : le gagnant a parcouru la distance en 6h et 54 min

DARPA Urban Challenge 2007



- Parcours urbain de 96 km sur une base militaire
- Le véhicule doit rouler avec du trafic
- Gérer des intersections avec des règles de priorité
- Doubler d'autres véhicules, gérer les rétrécissements de voie, se garer, ...
- Gagnant CMU : 4 heures et 10 minutes

Robots Autonomes

- Exploration de l'espace (NASA, ESA, ...)
- Exploration d'endroits dangereux, hostiles pour l'homme
 - sous-marin, centrales nucléaires, ...
- Chaînes de montage industrielles : réaction intelligente à des événements imprévus
- Applications de surveillance et militaires
 - détection automatique de comportement inhabituels, ...
- Aspirateurs autonomes intelligents
 - évitement d'obstacles, planification et mémorisation d'itinéraires, ...
- Robots domestiques intelligents

IA : Traitement du langage naturel

Le langage est par définition quelque chose inhérente aux êtres humains (intelligents)

Quelques exemples :

- Reconnaissance de la parole
 - Transcription de la suite de mots prononcée dans un signal de parole (pas de compréhension)
- Reconnaissance automatique du locuteur
 - Identification d'individus par la parole
- Traduction automatique
 - de texte (par exemple DeepL ou Google Translate)
 - de la parole (par exemple Skype Translate)



IA : Traitement du langage naturel

Le langage est par définition quelque chose inhérente aux êtres humains (intelligents)

Quelques exemples :

- Analyse syntaxique et sémantique
- Résumé automatique
- Recherche d'information
- Systèmes de dialogue (chatbots), question-réponse, ...

IA : Vision par machine

- Identification d'objets (caractères, visages, empreintes, objets industriels, ...)

IA : Vision par machine

- Identification d'objets (caractères, visages, empreintes, objets industriels, ...)



IA : Vision par machine

- Identification d'objets (caractères, visages, empreintes, objets industriels, ...)
- *Reconnaissance de visages*



IA : Vision par machine

- Identification d'objets (caractères, visages, empreintes, objets industriels, ...)
- *Reconnaissance de visages*
- Vision 3D



IA : Vision par machine

- Identification d'objets (caractères, visages, empreintes, objets industriels, ...)
- *Reconnaissance de visages*
- Vision 3D



→ **Reconnaissance de formes**

Exemples d'Applications de l'IA : les jeux

- Morpion ?
- Dames (premier programme dans les années 50)
- Échecs (IBM Deep Blue a battu G. Kasparov en 1997)
- Pacman (1980, déplacement intelligent dans un labyrinthe)
- Jeux de combat en 3D (rendre l'adversaire imprévisible, planification d'itinéraires et d'actions, *instinct de survie*, ...)
- Jouets qui s'adaptent au comportement du joueur (Nintendogs, Tamagotchi, ...)
- Plus récemment par Deepmind:
 - AlphaGo / AlphaZero / MuZero
 - <https://deepmind.com/research/case-studies/alphago-the-story-so-far>
 - AlphaStar
 - Résultats en vidéo: <https://youtu.be/xP7LwZxq0ss>

D'autres exemples d'Applications de l'IA

- Finance
 - Détection automatique de transactions anormales pour étude par un être humain
 - Conseil d'investissement
- Systèmes experts
 - Programmes pour répondre aux questions sur un domaine très spécifique
 - Utilise une base de données d'expertise humaine
 - Le but est de faciliter la consultation par un expert humain
- Médecine
 - Aide au diagnostic
 - Aide à la décision

Intelligence = Recherche ?

- Une approche possible
 - Représenter les connaissances
 - Rechercher une solution parmi un grand nombre de possibilités
- Exemple : les échecs
 - En principe, il est possible de calculer toutes les actions possibles et leur conséquence
 - Cependant, le nombre de possibilités explose rapidement (de façon exponentielle)

⇒ Il faut des techniques pour

 - éviter de faire des calculs redondants
 - éviter d'explorer des directions peu prometteuses
 - diminuer l'espace de recherche avec des heuristiques
 - reprendre la recherche après une erreur

→ cf. Algorithme Min-Max, élagage alpha-beta (vu en L2)

Intelligence = Comparaison ?

- Une approche possible
 - Mémoriser des exemples typiques du problèmes (caractères, images, paroles, mots, ...)
 - Comparer un objet inconnu avec tout ce qui était vu et trouver l'objet le plus proche
 - Exemple : identification de chiffres
 - ① enregistrer les images de quelques exemples de chiffres
 - ② comparer l'image d'un chiffre inconnu avec celles vues
 - Problème: énormément de variabilités dans les images
- ⇒ Il faut des techniques pour
- extraire des **caractéristiques** des exemples typiques
 - apprendre des **modèles** efficaces et compacts

Intelligence = Règles ?

- Une approche possible
 - ① Caractériser les choses par des règles
 - une girafe a 4 pattes, un long cou et une peau avec des tâches, ...
 - ② Vérifier si les règles s'appliquent à un objet inconnu
- Exemple : identification d'objets
 - les règles doivent être formulées par un spécialiste (**expert**)
 - il est difficile de couvrir toutes les exceptions
 - problème des objets incomplets ou bruités

Approches symboliques vs approches statistiques

Méthodes symboliques

- Représentations symboliques (concepts)
- Typiquement une **collection de règles**
- Expert humain spécifie les **règles**
- Des algorithmes pour
 - **combiner** les règles
 - **inférer** de nouvelles connaissances

Méthodes statistiques

- Représentations numériques (codes)
- Exemples typiques et la réponse souhaitée
- L'humain spécifie **quoi** mais pas **comment**
- Algorithmes d'apprentissage statistiques
 - Naive Bayes, RdN, SVM, HMM, ...

Dans ce cours : **méthodes statistiques**

Symbolique vs. Statistiques : un exemple intuitif

Exemples :

4 9 5 9

Méthodes symboliques

- Un **expert** analyse les exemples
- Un 4 contient 2 traits verticaux et 1 horizontal
- Un 9 contient un cercle et un trait vertical
- Un 5 contient 1 trait horizontal, 1 petit vertical et un cercle ouvert

Méthodes statistiques

- Exemples **annotés**:

4 9 5 9

→ Corpus d'apprentissage

- Entraînement d'un **modèle**

Classification : 4 ⇒ classe 4

Introduction : Apprentissage automatique



- Apprentissage automatique
 - Objectif: développer de nouveaux algorithmes d'apprentissage
 - Analyse théorique, capacité de généralisation
 - Évaluation sur une petite tâche, souvent avec données artificielles
 - Facile d'essayer de nouvelles idées
 - Approches remarquables : SVM, Boosting, Réseaux de Neurones, ...
 - Domaine qui évolue très vite

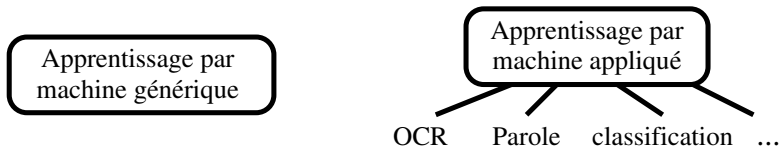
Introduction : Approches



- Apprentissage automatique **appliqué**

- Objectif: construire un système qui fonctionne bien pour une application spécifique
 - Pas nécessairement fondé sur une belle théorie
- ce sont les **résultats/performances** qui comptent
- Il est très difficile d'avoir un système avec des performances à l'état de l'art
 - Recherche plutôt lente et incrémentale

Introduction : Approches



- Interaction entre les communautés
 - Peu d'interaction entre les tâches
 - Il n'est pas aisé d'appliquer de nouvelles idées de l'apprentissage par machine générique à une tâche particulière

Introduction : Approches



- Interaction entre les communautés
 - Peu d'interaction entre les tâches
 - Il n'est pas aisé d'appliquer de nouvelles idées de l'apprentissage par machine générique à une tâche particulière
 - Objectifs de ce cours
 - Présenter les concepts et quelques algorithmes
- certains plus anciens et d'autres très récents
- Mise en oeuvre en TP
 - Programmation en Python

Courte histoire de l'Intelligence Artificielle

- Les origines:
 - Egypte (800 av. J.-C.) : statue dans la ville de Napata qui bouge son bras et qui parle aux spectateurs
 - Aristote (350 av. J.-C.) : crée les bases mathématiques des systèmes de déduction
 - Les éléments d'Euclide (300 av. J.-C.) : collection de 13 livres incluant entre autres des modèles de **raisonnement formel**, la géométrie et la théorie des nombres
 - al-Khwārizmī (783–850) : introduit le terme **algorithme**

Courte histoire de l'Intelligence Artificielle

- Les origines:
 - Egypte (800 av. J.-C.) : statue dans la ville de Napata qui bouge son bras et qui parle aux spectateurs
 - Aristote (350 av. J.-C.) : crée les bases mathématiques des systèmes de déduction
 - Les éléments d'Euclide (300 av. J.-C.) : collection de 13 livres incluant entre autres des modèles de **raisonnement formel**, la géométrie et la théorie des nombres
 - al-Khwārizmī (783–850) : introduit le terme **algorithme**
- 17^e-19^e siècle
 - Descartes : les animaux ne sont que des machines complexes
 - Thomas Hobbes : *reasoning is nothing but reckoning*
 - Gottfried Leibnitz : le raisonnement d'un être humain s'explique par des calculs algébriques
 - George Boole : *law of thoughts*, introduction de l'algèbre binaire

Histoire des Recherches en IA

- Début du 20^e siècle
 - Développement de la logique formelle
 - Machine de Turing
 - Google doodle Turing Machine
 - Vidéo machine de Turing : <https://www.youtube.com/watch?v=P66h8D5Lkwk>
 - λ -calcul de Church
 - Première mention du mot **robot**
- révolution d'esclaves mécaniques qui étaient mal traités, après être équipés avec de la "conscience"
- McCulloch & Pitts proposent le perceptron

Histoire des Recherches en IA

- La naissance

- Les années 1940 et 50 : quelques scientifiques envisagent la création d'un cerveau artificiel
- 1956 : l'IA est fondée en tant que domaine de recherche scientifique lors d'une conférence à Dartmouth
- Construction des premiers ordinateurs universels
 - déchiffrement de messages cryptés pendant la 2ème guerre mondiale (Enigma)

- Les années d'or : 1956 - 1974

- 4 centres de recherche obtiennent des financements importants :
 - MIT, CMU, Stanford et Edinburgh
- Raisonnement = recherche dans un espace de solutions possibles
 - heuristiques, back-tracking, ...
- Traitement du langage naturel (réseaux sémantiques, ELISA, ...)
- Vision par ordinateur : un bras articulé sait empiler des blocs

Histoire des Recherches en IA

- Un optimisme débordant

- H. Simon et A. Newell (CMU, 1958) :

- "within ten years a digital computer will be the world's chess champion"*

- "within ten years a digital computer will discover and prove an important new mathematical theorem."*

- H. Simon (1965) :

- "machines will be capable, within twenty years, of doing any work a man can do."*

- Marvin Minsky (1967) :

- "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved."*

- Marvin Minsky (1970 dans le magazine *Life*) :

- "In from three to eight years we will have a machine with the general intelligence of an average human being."*

Histoire des Recherches en IA

- La première récession (1974-1980)
 - On se rend compte que les prévisions étaient largement exagérées
 - Diminution importante des financements
 - Rapport ALPAC en 1966 sur la traduction automatique
- Problèmes
 - Les ordinateurs manquaient de puissance
 - traitement du langage prometteur, mais avec un vocabulaire de 20 mots
 - Passage à l'échelle semble être impossible à cause d'une complexité exponentielle

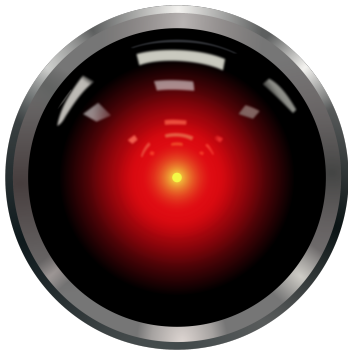
Histoire des Recherches en IA

- Le nouveau boom (1980-1987)
 - *Fifth generation computer project* au Japon (1981–1991) : réaliser des machines et programmes qui sont capables de faire des **conversations**, qui **traduisent des langues**, qui **interprètent des images** et qui **raisonnent** comme des êtres humains
 - Utilisation de systèmes d'experts
 - Le succès des réseaux de neurones (Hopfield, rétro-propagation, ...)
 - Langages de programmation spécialisés : Prolog et Lisp
- La deuxième récession (1987-1993)
 - Une fois de plus les promesses des chercheurs n'ont pas été réalisées
 - Des machines LISP dédiées sont devenues obsolètes avec l'arrivée des PCs

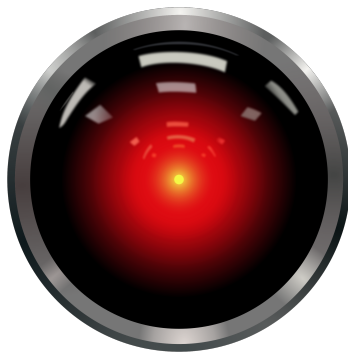
Histoire des Recherches en IA

- Le présent (depuis 1993)
 - L'augmentation continue de la puissance des ordinateurs permet d'utiliser des nouvelles méthodes
 - Big Blue de IBM est 10 million fois plus rapide qu'un ordinateur en 1951
 - On se focalise sur des **problèmes précis** au lieu de vouloir construire le « *general problem solver* »
 - Agents intelligents : un système qui perçoit son environnement et qui prend des décisions maximisant les chances de succès
 - Plus de maths ...

Histoire des Recherches en IA



Histoire des Recherches en IA



- HAL 900 n'existe toujours pas !
- « *space odyssey* », film de science fiction de 1968 :
 - HAL 9000, l'ordinateur au bord du vaisseau spatial « *Discovery one* » est doté d'intelligence
 - reconnaissance de la parole, des expressions faciales,
 - lecture des lèvres, raisonnement, émotions, ...

Définition

Apprentissage automatique (*machine learning*)

- **Wikipédia** : L'**apprentissage automatique** (en anglais : **machine learning**), **apprentissage artificiel** ou **apprentissage statistique** est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches **sans être explicitement programmés pour chacune**. Plus largement, il concerne la conception, l'analyse, l'optimisation, le développement et l'implémentation de telles méthodes.
- **Herbert Simon** (Prix Nobel d'économie en 1978) :
*L'apprentissage dénote des changements dans un système qui ...
lui permet de faire la même tâche plus efficacement la prochaine fois.*

⇒ L'objectif de l'apprentissage automatique est de concevoir **des programmes pouvant s'améliorer automatiquement avec l'expérience**

Applications de l'apprentissage Automatique

Tâche : distinguer plusieurs objets à partir d'exemples de données :

- Traitement du langage naturel (fouille de textes)
- Reconnaissances de formes
- Moteurs de recherche
- Diagnostic médical
- Bioinformatique
- Analyse des marchés boursiers
- Jeux

Types d'apprentissage

On distingue 3 grands types d'apprentissage :

- **Apprentissage supervisé**

- Chaque exemple est associé à une étiquette
- Objectif : prédire l'étiquette de chaque donnée
- Le système apprend à classer les données

- **Apprentissage non-supervisé**

- Les exemples ne sont pas étiquetés
- Objectif : trouver une structure aux données
- Le système apprend une classification des données

- **Apprentissage par renforcement**

- Les exemples sont associés à une **récompense** (reward)
- Objectif : maximiser les récompenses
- Le système apprend une politique de décision

→ Dans ce cours : apprentissage **supervisé** et **non-supervisé**

Apprentissage supervisé - classification supervisée

- Le nombre et le type des classes sont connus d'avance et fixes
 - **rejet** possible si aucune classe ne convient
- Les données d'apprentissage sont étiquetées préalablement par un **expert** ou **oracle**.
 - Objectif : classer les objets non vus à l'entraînement dans des classes
- Processus en deux phases :
 - ① **apprentissage** : déterminer un modèle à partir des données étiquetées
 - ② **test** : prédire l'étiquette d'une nouvelle donnée en utilisant le modèle préalablement appris.
- Exemples d'application
 - Reconnaissance de l'écriture, de la parole ...
 - Météo : classification d'image satellite

Classification non-supervisée (clustering)

- On parle d'**observations** plutôt que d'**exemples**
- Les données d'apprentissage **ne sont pas étiquetées**
- Aucun expert n'est requis

⇒ Objectif : faire **ressortir de l'information** à partir **des données**

- **Classer** les données en groupes **homogènes**
 - **Regrouper** les données selon leur **similarité**
 - **Associer** ou déduire du sens pour chaque groupe
-
- Exemples d'application
 - Analyse des données en général
 - Identifier des clients dans supermarché selon leurs profils
 - Identifier des groupes de risque pour une assurance

Apprentissage par renforcement

En intelligence artificielle / apprentissage automatique, l'apprentissage par renforcement consiste à **apprendre une stratégie** (appelée **politique**) à partir d'expériences, de façon à optimiser une **récompense** (**reward**) au cours du temps. Le système (appelé **agent**) est dans un environnement qui lui procure une **récompense** en fonction de ses décisions. L'agent cherche une **politique** optimale visant à maximiser la somme des **récompenses** au cours du temps.

- Concepts clés :

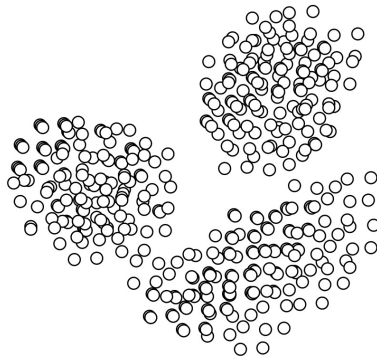
- L'action modifie l'environnement
- L'environnement retourne un feedback (positif ou négatif)
- L'agent modifie sa politique en fonction du retour obtenu pour améliorer ses actions futures

→ L'être humain fonctionne de cette manière !

- Exemples: apprentissage du vélo, de la marche, découverte du feu, ...

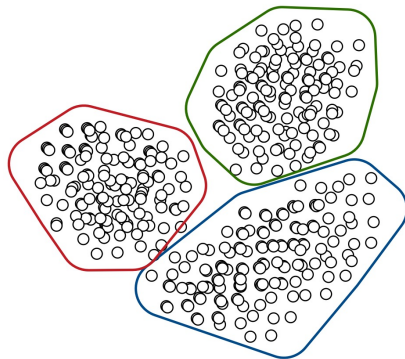
Dans ce cours

À partir de données non annotées



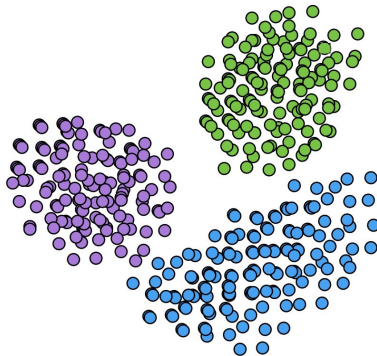
Dans ce cours

À partir de données non annotées → algorithmes de classification non-supervisée



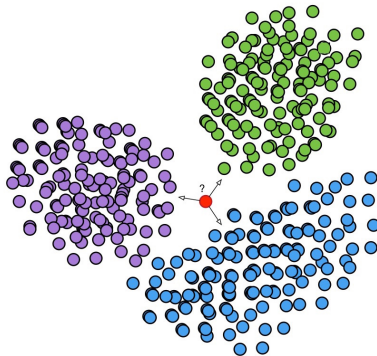
Dans ce cours

À partir de données annotées



Dans ce cours

À partir de données annotées → algorithmes de classification supervisée



Questions?

Fragen?

Domande?

Perguntas

¿Preguntas?

Questions?