

Université J.F. CHAMPOLLION Albi

Introduction à l'intelligence artificielle



Licence d'informatique S6

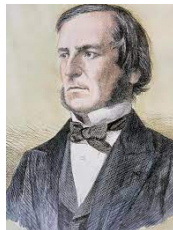
Thierry Montaut

Introduction à l'IA

- L'objectif de ce cours est découvrir en mode projet et par l'exemple, l'étendu des problématiques de l'intelligence artificielle et en particulier de l'apprentissage automatique.
- Pour chaque problématique, découvrir les notions théoriques clés, les possibilités des principales bibliothèques python dédiées à l'analyse des données et à l'apprentissage automatique et aller jusqu'à la réalisation pratique d'un modèle, son analyse et son utilisation pratique.
- Le but de ce premier chapitre est de présenter des éléments de contexte du machine learning en parcourant son histoire, de ses prémices jusqu'aux évolutions actuelles.

Un peu d'histoire : Quelques précurseurs

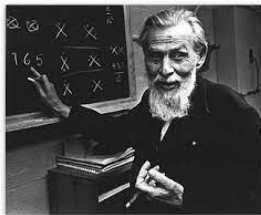
- George Boole (1815 - 1864)



1854 : An Investigation of The Laws of Thought on which are Founded The Mathematical Theories of Logic and Probabilities.

Quelques précurseurs

- Warren McCulloch (1898-1969) et Walter Pitts (1923-1969)



Neurologues, inventent en 1943 le premier modèle mathématique du neurone biologique : le neurone formel. Nous étudierons en détail ce premier modèle.

Quelques précurseurs

- Alan Turing (1912-1954)



Conceptualise les fonctions calculables avec la machine de Turing ;

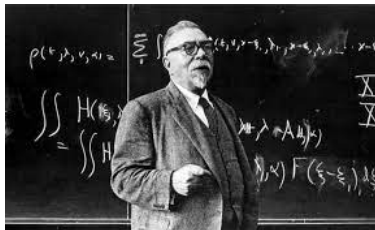
1948 : « Intelligent machinery »

1950 : « Computing machinery and intelligence » ;

1956 : « Can a machine think ? » : test de Turing

Quelques précurseurs

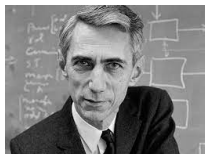
- Norbert Wiener (1894-1964)



Mathématicien, lance dans les années 1940 la cybernétique, comme science du fonctionnement de l'esprit humain.

Quelques précurseurs

- Claude Shannon (1916-2001)

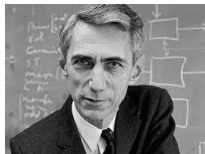
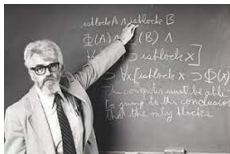


a défini la théorie de l'information. 1950 : « Programming a computer for playing chess. »

- Lotfi Zadeh (1921-2017) a défini la logique floue. 1950 : « Thinking machines - a new field in electrical engineering » - les machines à penser, un nouveau champ de l'ingénierie électrique.

Naissance "officielle"

- Été 1955 aux états unis : Les mathématiciens John McCarthy (1927-2011), Marvin Minsky (1927-2016), Nathaniel Rochester (1919-2001) et l'ingénieur électronicien Claude Shannon ont proposé un projet de recherche sur l'intelligence artificielle au Dartmouth college 1955 : "A proposal for the Dartmouth summer research project on artificial intelligence"
- Ils obtiennent de la fondation Rockefeller 7500 \$ afin d'organiser à l'université de Dartmouth, durant les 2 mois de l'été 1956, un atelier de travail sur les machines pensantes.



La conférence de Dartmouth

- La conférence de Dartmouth (Dartmouth Summer Research Project on Artificial Intelligence) : atelier scientifique organisé durant l'été 1956,
- considéré comme l'acte de naissance de l'intelligence artificielle en tant que domaine de recherche autonome.
- Organisée par Marvin Minsky et John McCarthy, sur le campus du Dartmouth College et a réuni vingt chercheurs dont Claude Shannon, Nathan Rochester, Ray Solomonoff, Trenchard More, Oliver Selfridge, Allen Newell et Herbert Simon.
- Ces deux derniers y présentent le programme Logic Theorist, capable de démontrer des théorèmes mathématiques.
- John Mc Carthy y présente le principe d'élagage alpha-bêta, un algorithme d'évaluation notamment utilisé par la grande majorité des programmes d'échecs.

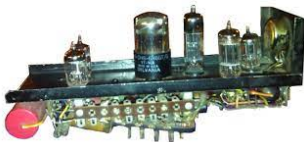
La conférence de Dartmouth

- Le nombre de scientifiques intéressés par le sujet est très limité
- Les recherches sont freinées par l'encombrement et les faibles performances des machines de l'époque



Quelques dates à retenir

- 1950 : Test de Turing
- 1951 : Minsky construit le SNARC, première machine à réseaux de neurones



- 1956 : 1er programme d'ordinateur capable de démontrer des théorèmes de logique
- 1957 : Frank Rosenblatt introduit la notion de perceptron : premier apprentissage supervisé d'un réseau de neurones

Quelques dates à retenir

- 1959 : Oliver Selfridge (1926-2008), reconnaissance des formes
Arthur Samuel (1901-1990), pionnier des programmes de jeux de dames et d'échec basé sur l'idée d'apprentissage, capable de gagner contre de bons joueurs
- 1970-1980 : Vives critiques et fin des investissements (premier hiver de l'IA)
- 1980-1987 : Le boom des systèmes experts

Quelques dates à retenir

- 1986 : Retour des réseaux de neurones grâce au Perceptron multicouche
- 1987-1993 : Le second hiver de l'IA
- 1995, Yann LeCun et Yoshua Bengio introduisent le concept de réseau de neurones convolutionnel ou Convolutional Neural Networks (CNN)
- 1997 : Victoire de Deep Blue sur Garry Kasparov

Quelques dates à retenir

- 2000 : Deep learning. Fort investissement des GAFAM
- Forte **mathématisation** de l'IA : Probabilité, graphes bayésiens, Calcul différentiel, algèbre linéaire...
- Explosion des **données disponibles** pour l'apprentissage
- Très forte progression des **capacité de calcul**
- 2015-2016 : Victoires d'AlphaGo sur les champions d'Europe et du monde

Problématiques de l'IA

L'ambition de l'IA : faire ce que fait le cerveau humain... c'est-à-dire, rendre la machine capable :

- de reconnaître, d'identifier ;
- de raisonner sur une situation statique ou dynamique ;
- de résoudre des problèmes combinatoires ;
- de faire un diagnostic ;
- de proposer une décision, un plan d'action ;
- d'expliquer, de communiquer ;
- de comprendre un texte, un dialogue en langage naturel ;
- d'acquérir de l'information, de la synthétiser ...

Les paradigmes

- représentation des connaissances ;
- formalisation des raisonnements et des décisions ;
- algorithmes généraux de résolution de problèmes ;
- apprentissage ;

Une définition ?

- Luc Julia (Siri) : L'IA n'existe pas, la machine ne "pense" pas mais exécute des algorithmes même si la nature, le type de ces algorithmes a changé.
- L'IA est l'ensemble des outils, des méthodes et des algorithmes permettant à une machine de réaliser des tâches complexes sans avoir été explicitement programmée pour cela.
- Dans cette introduction, on pourra donc pratiquement confondre l'IA et l'apprentissage (Machine learning).

Machine learning (ML)

- Sous-ensemble de l'IA qui concerne la création d'algorithmes qui permettent d'apprendre à partir de données précédemment collectées.
- 1959 : "Machine learning enables a machine to automatically learn from data, improve performance from experiences and predict things without being explicitly programmed".
- On peut toujours considérer qu'à partir d'un ensemble X , on cherche à apprendre pour tout x dans X la valeur $y = f(x)$ que prend une fonction f en x .
- si f est la fonction à apprendre, on notera \tilde{f} la fonction apprise et $e = \tilde{f} - f$, l'erreur d'apprentissage.

Différents problèmes d'apprentissage

Quand on cherche à apprendre la fonction

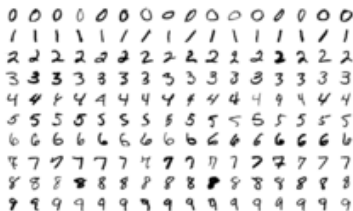
$$f : \begin{pmatrix} X & \rightarrow & Y \\ x & \mapsto & y = f(x) \end{pmatrix}$$

On distingue les problèmes d'apprentissage selon la nature de l'ensemble Y :

- Si Y est un ensemble fini, on parle de problème de **classification**. La fonction à prédire est un classifieur.
- Dans le cas particulier où $Y = \{\text{vrai}, \text{faux}\}$ on parle également de prédicteur.
- Si $Y \subseteq \mathbb{R}$, on parle de problème de **régression**.

Exemple de problèmes de classification

- Distinguer les éléphants des girafes : X est un ensemble de couples (poids,taille), Y est l'ensemble {éléphant,girafe}
- Reconnaître des chiffres manuscrits : X est un ensemble d'images, $Y = \{0, 1, \dots, 9\}$.

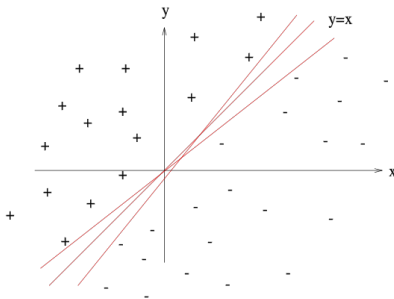


- Détection de spam, détection de visage, reconnaissance faciale...

Apprentissage supervisé

- On parle d'apprentissage supervisé lorsque la méthode d'apprentissage de la fonction f nécessite l'utilisation d'une base de données pour lesquelles on connaît exactement les résultats. Dit autrement, d'un ensemble de couples $\{(x_1, y_1), \dots, (x_n, y_n)\}$ pour lesquels $y_i = f(x_i)$.
- La séparation linéaire, les k plus proches voisins, les réseaux de neurones, les réseaux bayésiens sont des méthodes d'apprentissage supervisé que nous étudierons dans ce cours d'introduction à l'IA.

Un classifieur linéaire simple



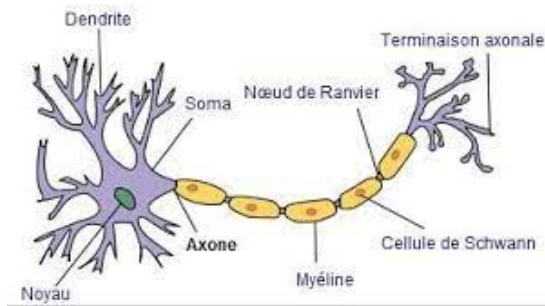
- Les girafes sont grandes avec un long cou, les éléphants sont massifs et lourds. Il doit donc être possible de les différencier à l'aide de deux caractéristiques : leur poids et de leur taille.
- On dispose de données nous fournissant ces caractéristiques pour un certain nombre de girafes et d'éléphants.
- Sur la figure ci-dessus, x est le poids de l'animal, y sa taille et on représente par + les girafes et par - les éléphants.

Un classifieur linéaire simple

- On cherche à déterminer la droite qui sépare au mieux notre jeu de données.
- Plus généralement si nous avons n caractéristiques et non seulement deux, nous chercherions dans \mathbb{R}^n l'équation d'un hyperplan (un sous-espace vectoriel de dimension $n - 1$ séparant au mieux nos données).

Réseaux de neurones formels

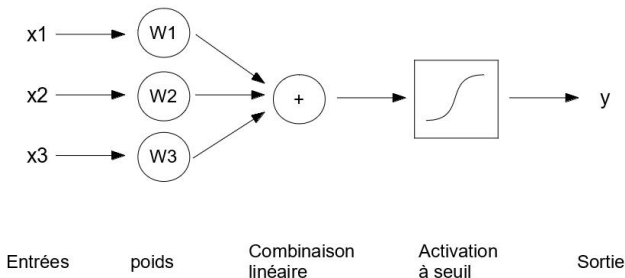
- Le neurone formel s'inspire directement du neurone biologique :



- Les études en neurophysiologie montrent que les dendrites correspondent aux entrées du neurone et l'axone à sa sortie. Le cerveau est ainsi constitué d'un réseau sophistiqué de neurones connectés les uns aux autres par le biais de dendrites et d'axones, d'entrées et de sorties.

Le neurone formel binaire

- Le modèle de neurone formel proposé par McCulloch et Pitts reproduit le comportement essentiel du neurone biologique : sa capacité à s'activer au delà d'un certain seuil :

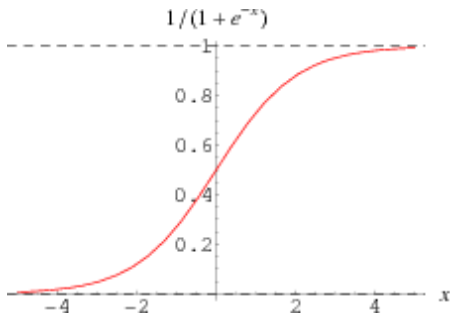


Le neurone formel binaire

- Le modèle initial de McCulloch et Pitts est binaire : Les entrées x_i valent 0 ou 1 et sont pondérées par des poids $w_i \in [0, 1]$. Si la combinaison linéaire obtenue est supérieure à un seuil donné, la sortie y vaut 1 (le neurone est activé), sinon elle vaut 0.
- Par exemple : Si l'entrée vaut (1,0,1), que les poids valent respectivement 0.5, 1 et 0.2, la combinaison vaudra 0.7. Si le seuil est fixé à 0.6, le neurone sera activé et sa sortie vaudra 1.

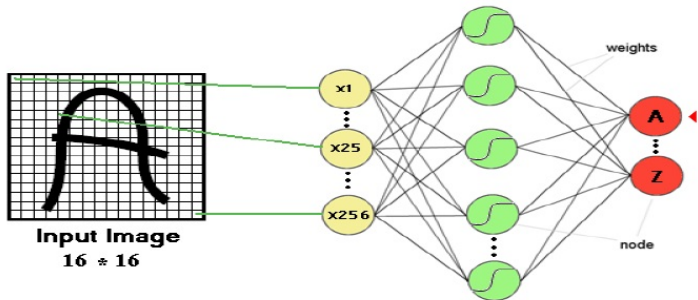
Le neurone formel numérique

- On peut améliorer le comportement du neurone formel en remplaçant la simple fonction à seuil par une fonction numérique qui donnera toutes les valeurs réelles entre 0 et 1. Les réseaux de neurones peuvent en utiliser plusieurs mais la plus répandue est la fonction sigmoïde $f(x) = \frac{1}{1+e^x}$:



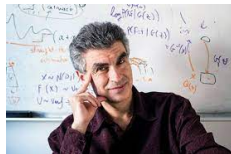
Réseau de neurones convolutifs (CNN)

- Obtenus en assemblant les sorties de neurones avec les entrées de nouveaux neurones. On obtient un réseau en couches.



Réseau de neurones convolutifs (CNN)

- Réseaux de neurones multicouches, introduits en 1995 par Yann LeCun et Yoshua Bengio,

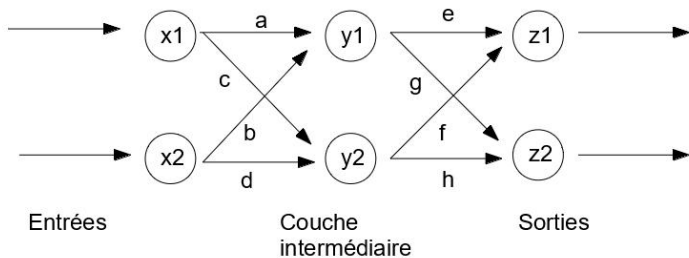


Réseau de neurones convolutifs (CNN)

- Inspirés de la neurobiologie des cellules nerveuses sélectives en fonction de l'orientation dans le cortex
- Proche du premier Perceptron de Rosenblatt en 57
- Non rebouclé sur lui même (feed-forward), l'information se propage des entrées vers les sorties.
- Capables de reconnaître des motifs avec beaucoup de variabilité. (caractères manuscrits par exemple).
- Le nombre de paramètres à apprendre devient rapidement très important

Un réseau de neurone élémentaire

Un réseau simple à trois couches de deux neurones



Plan du cours

- Introduction

- ▶ Histoire de l'IA
- ▶ Python pour les data Sciences (Numpy, Pandas, Matplotlib, Seaborn)

- L'apprentissage supervisé

- ▶ Critères de performance, surapprentissage
- ▶ Python pour l'IA (Scikit-learn, Tensor Flow)
- ▶ Le choix d'un modèle

Plan du cours

- Les problèmes de regression
 - ▶ A l'aide des statistiques
 - ▶ Par descente de gradient
 - ▶ à l'aide de scikit-learn
- Classifieurs KNN
 - ▶ Classification des iris
 - ▶ Auriez-vous survécu au naufrage du Titanic ?

Plan du cours

- Classifieurs à SVM

- ▶ Cas linéairement séparables (marges rigides)
- ▶ Cas linéaires, non séparables (marges souples)
- ▶ Cas non linéaires (l'astuce du noyau)

- Classification bayésienne

- ▶ Les tests Covid sont-ils fiables ?
- ▶ Construction d'un anti-spam

Plan du cours

- Les réseaux de neurones
 - ▶ Le perceptron. Quelles plantes sont toxiques ?
 - ▶ Réseau multicouche : Les chiens ne sont pas des chats !
- Clustering : Classification non supervisée
 - ▶ Performance d'une classification non supervisée
 - ▶ Algorithme des k moyennes