

Introduction à l'apprentissage automatique, la science de l'intelligence artificielle

Séance 1

Introduction, vocabulaire, apprentissage supervisé et non-supervisé

Frédéric Sur

https://members.loria.fr/FSur/enseignement/IMT_GE/

Plan

- 1 Introduction
 - Vocabulaire et principales notions
 - Exemples
- 2 Apprentissage non-supervisé et supervisé
- 3 Programme de la semaine

Qu'est-ce que l'apprentissage automatique ?

*L'apprentissage automatique (en anglais **machine learning**, littéralement « l'apprentissage machine ») ou **apprentissage statistique** est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches 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, cela concerne la conception, l'analyse, le développement et l'implémentation de telles méthodes.*

https://fr.wikipedia.org/wiki/Apprentissage_automatique

Intelligence Artificielle ?

« Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine »

Source : larousse.fr

Test de Turing
Imitation game

IA générale,
IA forte,
IA faible

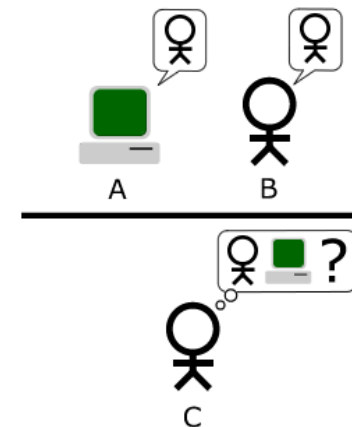


Illustration : By Bilby (Own work) [Public domain], via Wikimedia Commons

Quelques notions historiques

- 1763-1812 : théorème de Bayes (Bayes - Laplace)
- ~1945 : premiers ordinateurs
 - applications potentielles : traitement du langage naturel, « systèmes experts », OCR, computer vision, jeux. . .
- 1970s : "First AI winter"
- 1980s : réseaux de neurones artificiels et apprentissage
- mi-1980s / mi-1990s : "Second AI winter"
- 1990s : premières voitures « autonomes »
- fin 1990s : SVM, méthodes à noyau
- 1997 : Deep Blue bat Garry Kasparov
- début 2000 : robotique grand public
- 2010s : apprentissage profond (*deep learning*), réseaux de neurones convolutionnels
- 2016 : AlphaGo bat Lee Sedol

5/30

27 mars 2019 Le Monde

Le prix Turing récompense trois pionniers de l'intelligence artificielle (IA)

L'association américaine ACM a remis son prestigieux prix aux chercheurs français, canadien et britannique : Yann LeCun, Yoshua Bengio et Geoffrey Hinton.

Par David Larousserie - Publié le 27 mars 2019 à 11h01 - Mis à jour le 29 mars 2019 à 12h11

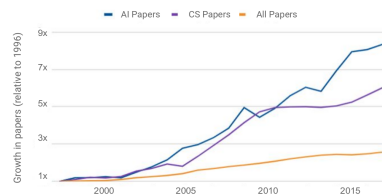
La plus importante société savante d'informatique, Association for Computing Machinery (ACM), a délivré, mercredi 27 mars, sa prestigieuse récompense, le prix Turing, à trois pionniers d'un domaine technologique très à la mode, l'intelligence artificielle. Le Canadien Yoshua Bengio, le Français Yann LeCun et le Britannique Geoffrey Hinton se partagent le million de dollars (889 000 euros) du prix, qui représente le sommet de cette discipline depuis la fin des années 1960. Yann LeCun est le second Français, après Joseph Sifakis en 2007, à être honoré par l'ACM.

Les trois lauréats ont chacun contribué à une partie de l'intelligence artificielle, en développant des méthodes d'apprentissage automatique, appelé *deep learning* ou réseaux de neurones profonds. Les paramètres de ces algorithmes ne sont pas introduits par les programmeurs mais « appris » à partir d'une base de données d'exemples, par essais-erreurs. Des millions, voire des milliards, de paramètres sont ainsi modifiés à partir de milliers, voire de millions, d'exemples.

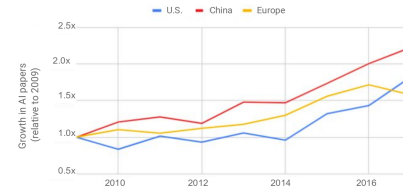
6/30

L'IA en quatre graphiques

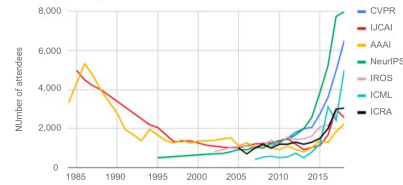
Growth of annually published papers by topic (1996–2017)
Source: Scopus



Growth in government-affiliated AI papers (2009–2017)
Source: Elsevier



Attendance at large conferences (1984–2018)
Source: Conference provided data



Annual VC funding of AI startups (U.S., 1995–2017)
Source: Sand Hill Econometrics



Source : AI Index 2018 Annual Report

7/30

Exemple : ressources humaines ?

L'Impact De L'IA Sur Les Ressources Humaines



David Mallon, vice-président et analyste chez Bersin, soutient que les services centralisés seront les premiers, en matière de ressources humaines, à être impactés par la robotisation et/ou l'automatisation. Rapidement, le travail humain devrait y être suppléé, voire même remplacé.

« Créés pour rationaliser, dimensionner et standardiser des processus ou des programmes administratifs RH répétitifs, les services centralisés se prêtent bien à la robotisation, qui systématise un travail de routine conséquent en plus d'être manuel. Imaginez tout ce qu'un humain pourrait faire en plus, en travaillant avec un système de données qui le soulage d'une charge d'une telle charge de travail, en imitant les actions humaines. »

L'influence de l'intelligence artificielle est déjà perceptible dans un domaine tel que le recrutement, selon Ksenia Zheltoukhova, responsable de la recherche au Chartered Institute of Personnel and Development. « Il s'agit de s'assurer que les candidats ont un parcours particulier, et une des étapes fondamentales du processus de recrutement, facilitée par l'IA, est l'interview vidéo ».

→ IA et biais ?

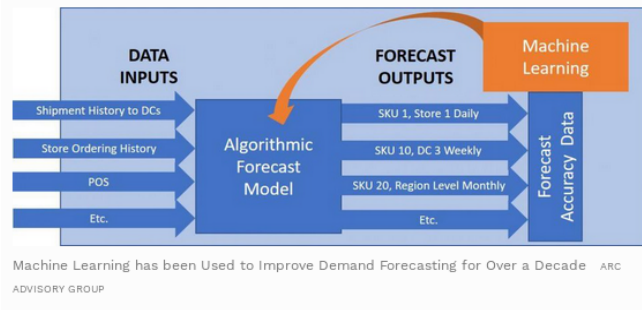
8/30

Exemple : chaîne logistique

20 Things To Know About Artificial Intelligence For Supply Chain Management



Steve Banker Contributor @
Transportation
I cover logistics and supply chain management.



Exemple : atomistic machine-learning



Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc



Amp: A modular approach to machine learning in atomistic simulations*



Alireza Khorshidi, Andrew A. Peterson*

School of Engineering, Brown University, Providence, RI, 02912, United States

ARTICLE INFO

Article history:
Received 4 March 2016
Received in revised form 4 May 2016
Accepted 8 May 2016
Available online 2 June 2016

Keywords:
Potential energy surface
Neural networks
Atomic Simulation Environment (ASE)
Density functional theory
Zernike polynomials

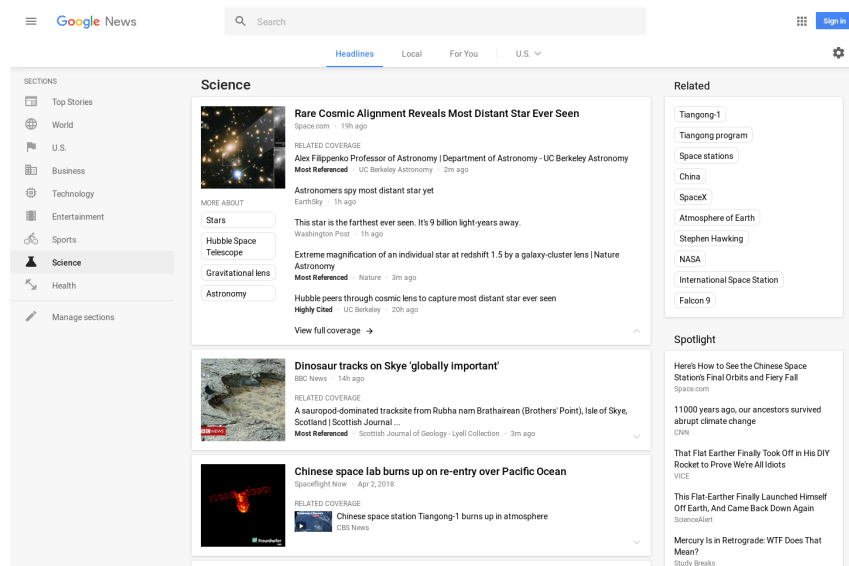
ABSTRACT

Electronic structure calculations, such as those employing Kohn-Sham density functional theory or *ab initio* wavefunction theories, have allowed for atomistic-level understandings of a wide variety of phenomena and properties of matter at small scales. However, the computational cost of electronic structure methods drastically increases with length and time scales, which makes these methods difficult for long time-scale molecular dynamics simulations or large-sized systems. Machine-learning techniques can provide accurate potentials that can match the quality of electronic structure calculations, provided sufficient training data. These potentials can then be used to rapidly simulate large and long time-scale phenomena at similar quality to the parent electronic structure approach. Machine-learning potentials usually take a bias-free mathematical form and can be readily developed for a wide variety of systems. Electronic structure calculations have favorable properties – namely that they are noiseless and targeted training data can be produced on-demand – that make them particularly well-suited for machine learning. This paper discusses our modular approach to atomistic machine learning through the development of

(crédit : Émilie Gaudry - dpt matériaux Mines Nancy)

10/30

Exemple : Google news



Exemple : Optical Character Recognition (OCR)

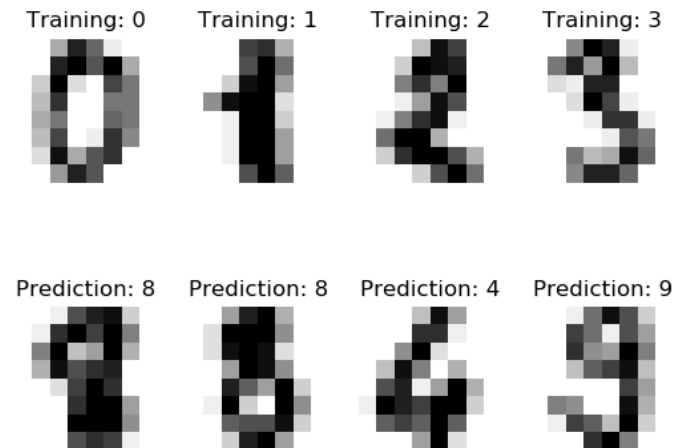


Illustration : scikit-learn

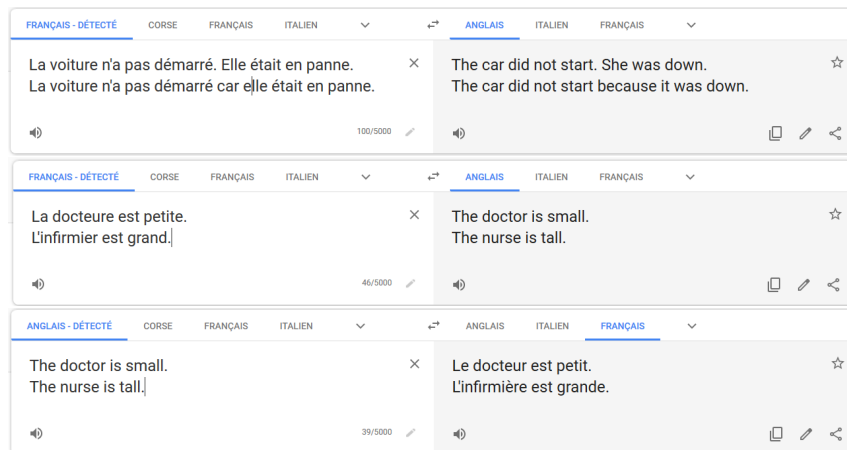
12/30

9/30

Source : Forbes 2019

11/30

Exemple : Google translate



→ IA et biais ?

→ essayez avec deepl.com

13/30

Exemple : reconnaissance de la parole



14/30

Exemple : génération automatique de texte

https://www.tensorflow.org/tutorials/text/text_generation

"We will work with a dataset of Shakespeare's writing from Andrej Karpathy's The Unreasonable Effectiveness of Recurrent Neural Networks. Given a sequence of characters from this data ("Shakespear"), train a model to predict the next character in the sequence ("e"). Longer sequences of text can be generated by calling the model repeatedly."

"The following is sample output when the model is [...] started with the string Q" :

QUEENE :

*I had thought thou hadst a Roman ; for the oracle,
Thus by All bids the man against the word,
Which are so weak of care, by old care done ;
Your children were in your holy love,
And the precipitation through the bleeding throne.*

Voir aussi :

Claude Shannon, *A mathematical theory of communication*, 1948.

15/30

Exemple : pixel recursive super-resolution

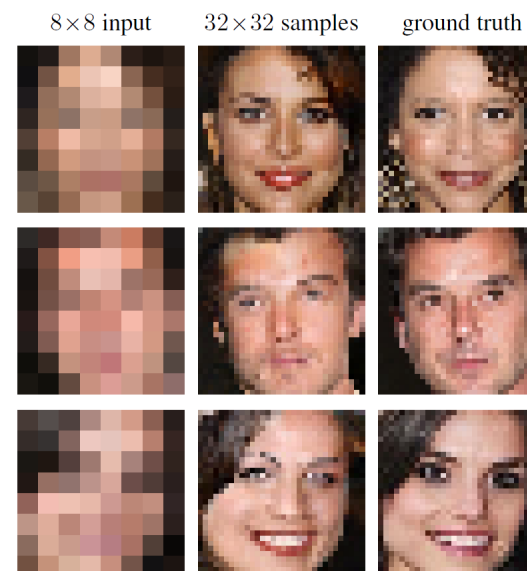


Illustration : Googlebrain, ICCV 2017

16/30

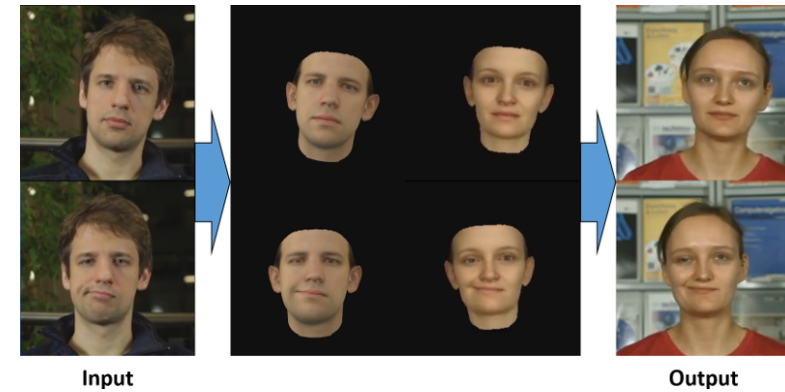
Exemple : “smart” city ?



Illustration : Saul Loeb/AFP/Getty Images

17/30

Exemple : deepfake



Deep Video Portraits (Siggraph 2018)

https://web.stanford.edu/~zollhofer/papers/SG2018_DeepVideo/page.html

18/30

Plan

- 1 Introduction
 - Vocabulaire et principales notions
 - Exemples
- 2 Apprentissage non-supervisé et supervisé
- 3 Programme de la semaine

19/30

Apprentissage non-supervisé et supervisé

Apprentissage non-supervisé :

on dispose d'un ensemble d'observations

→ on veut découvrir des caractéristiques communes à certaines observations (modèle probabiliste sous-jacent, classification...)

Exemples : *Google news*, *génération de texte*

Apprentissage supervisé :

classification : on dispose d'un ensemble d'observations, chaque observation appartenant à une classe

→ on veut prédire la classe d'une nouvelle observation

Exemples : *reconnaissance de caractères*, *reconnaissance faciale*

régression : on dispose d'un ensemble d'observations, chaque observation étant associée à une grandeur scalaire ou vectorielle

→ on veut prédire la grandeur associée à une nouvelle observation

Exemples : *super-resolution*, *deepfake*

20/30

Apprentissage non supervisé / partitionnement

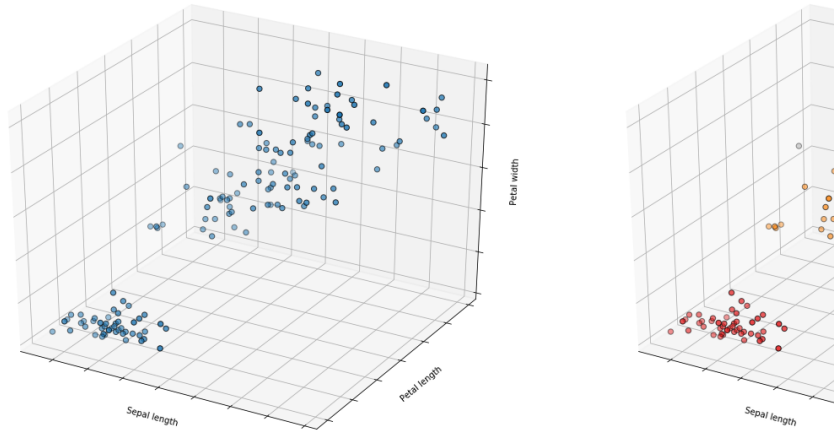


Illustration : scikit-learn

Iris de Fisher : identification automatique des espèces ?

21/30 **Difficultés** : dimension, métrique de similarité, ambiguïté

Qu'est-ce que le partitionnement ?

Les **données** / **exemples** / **observations** :

$x_i \in \mathbb{R}^d$, $i \in \{1, \dots, N\}$ (N : taille de l'échantillon)

Apprentissage non supervisé / partitionnement (*clustering*)

But : trouver des sous-ensembles des x_i partageant des caractéristiques communes

Exemples :

- x_i « vecteur » représentant un individu :
identification de profils similaires ?
- x_i « vecteur » représentant un article de journal :
identification de sujets ?

22/30

Apprentissage supervisé : classification

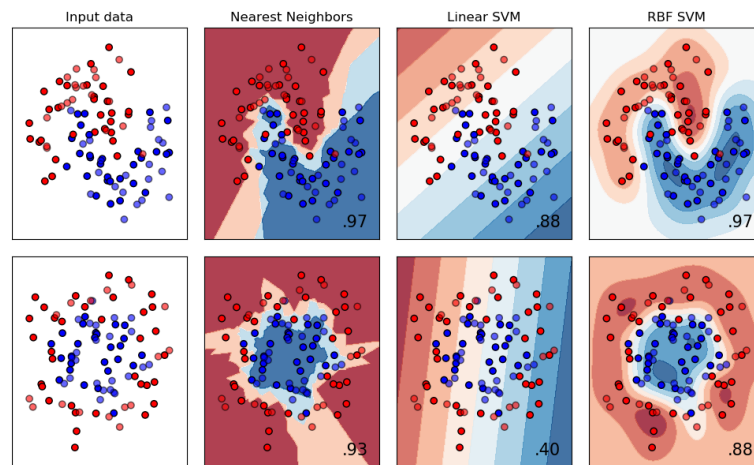


Illustration : scikit-learn

Difficultés : ambiguïté, dimension, fonction d'apprentissage, évaluation des performances...

23/30

Apprentissage supervisé : régression

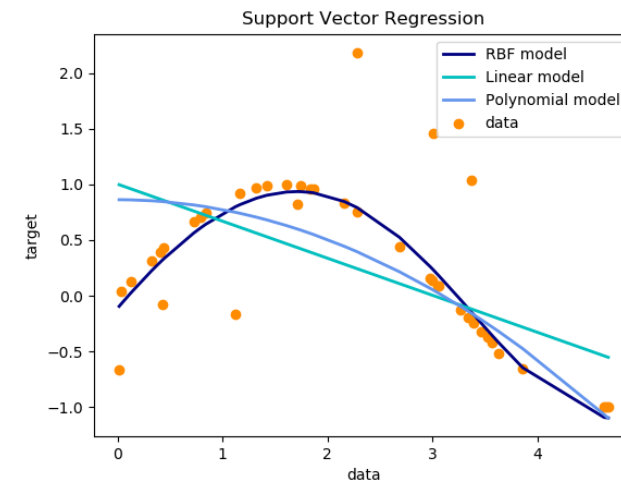


Illustration : scikit-learn

La régression linéaire est un algorithme de l'intelligence artificielle !

24/30

Qu'est-ce que l'apprentissage supervisé ?

Les **données** / **exemples** / **observations** :

$x_i \in \mathbb{R}^d$, $i \in \{1, \dots, N\}$ (N : taille de l'échantillon)

→ chaque donnée x_i est associée à une « sortie » attendue y_i

Apprentissage supervisé

pour une « fonction » f_w (ensemble de paramètres w)

Phase 1 : apprentissage, trouver \tilde{w} tel que $\forall i, f_{\tilde{w}}(x_i) \simeq y_i$

Phase 2 : prédiction avec x inconnu, $f_{\tilde{w}}(x) = ?$

$y_i \in \{1, 2, \dots, N\}$: *classification supervisée*

$y_i \in \mathbb{R}^{d'}$: *régression*

Exemples :

- $d = 1$ pour $(x_i, y_i) = (\text{vitesse, distance d'arrêt})$ (*régression*)
- $d = 8 \times 8$, $y_i \in \mathbb{R}^{32 \times 32}$ en slide 15 (*régression*)
- $d = 12 \cdot 10^6$ pour x_i image 12Mpx, $y_i \in \{\text{vélo, voiture}\}$ (*classification*)

Apprentissage par renforcement

→ troisième grande famille : thème non abordé dans ce cours

“apprendre les actions à prendre, à partir d'expériences, de façon à optimiser une récompense quantitative au cours du temps. L'agent est plongé au sein d'un environnement, et prend ses décisions en fonction de son état courant. En retour, l'environnement procure à l'agent une récompense, qui peut être positive ou négative.”

https://en.wikipedia.org/wiki/Reinforcement_learning

Exemples :

- AlphaGo
- *Intrinsically Motivated Goal Exploration, Automated Curriculum Learning and Emergence of Tool Use*
<https://www.youtube.com/watch?v=NOLAwD4ZTW0>

25/30

26/30

Plan

1 Introduction

- Vocabulaire et principales notions
- Exemples

2 Apprentissage non-supervisé et supervisé

3 Programme de la semaine

27/30

Objectifs de la semaine

- comprendre les problématiques scientifiques de l'IA
→ fondements, limites théoriques et pratiques
- comprendre ce que signifie “apprentissage”
→ principaux modèles, validation
- expérimenter dans un environnement standard utilisé en milieu professionnel / scientifique
→ carnets Jupyter et Python scikit-learn

28/30

Au programme. . .

Les grands principes de l'IA : malédiction de la dimension, sous-apprentissage vs. surapprentissage, introduction à la théorie statistique de l'apprentissage

Apprentissage non-supervisé : algorithmes de partitionnement

Apprentissages supervisé : plus proches voisins, régression logistique, perceptron, réseaux de neurones, machines à vecteurs supports

TP : comprendre les limites des algorithmes vus (exemples jouets) + applications réalistes

Modalités d'évaluation

À rendre chaque soir :

document Word avec les réponses rédigées aux TP

mail : sur@loria.fr