

Introduction aux réseaux neuronaux

Olivier Ricou

2023

Cas d'utilisations 1/2



Pub. ciblée



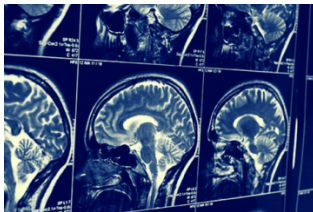
Recommendations



Description



Sécurité



Diagnostic



Jeux



Majordome

Cas d'utilisations 2/2



Conversation / Code



Génération d'images

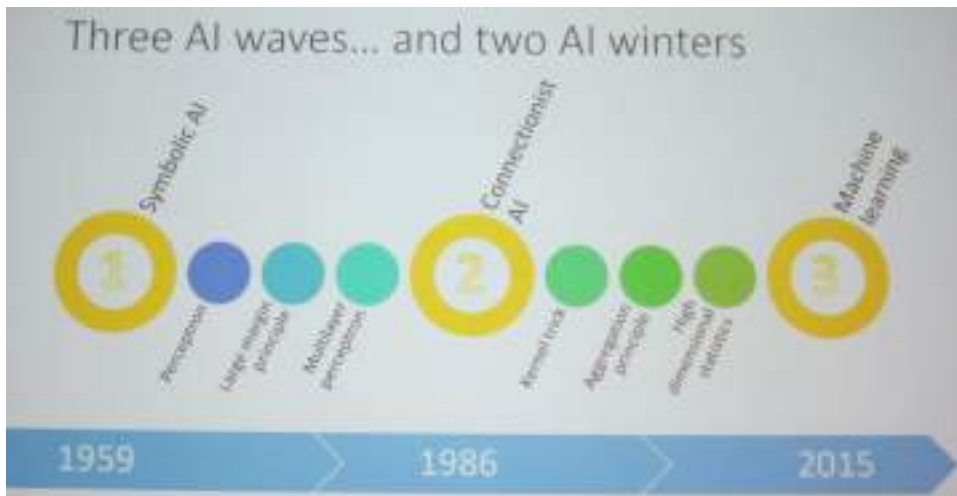


Conduite

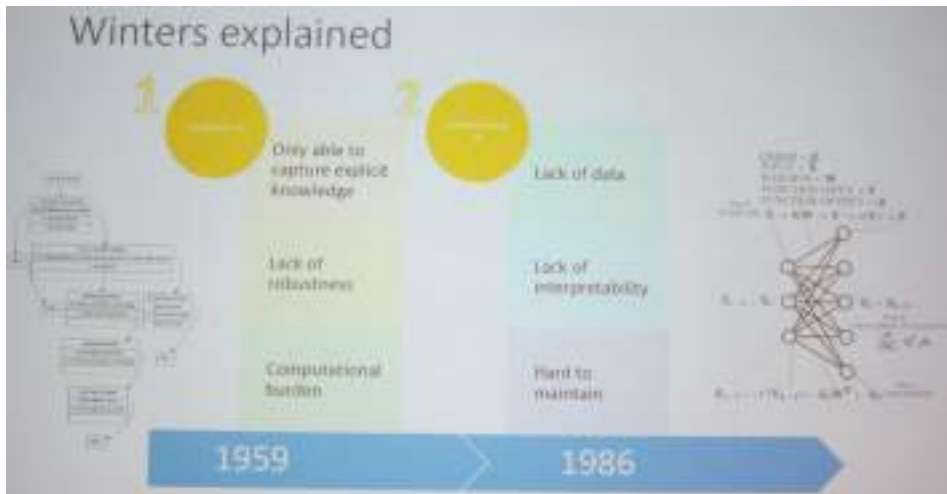


Guerre

Historique

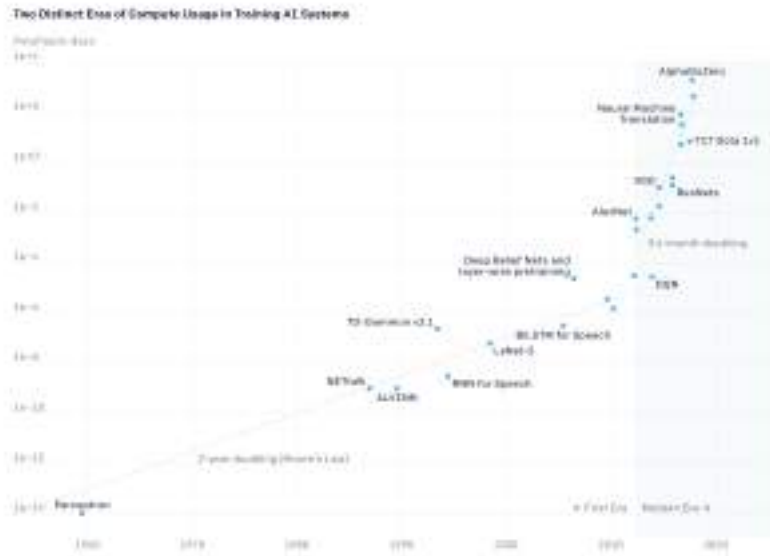


Les hivers passés



La renaissance est dues au triptique **données, hardware, théorie**.

Plus gros c'est mieux



Source : <https://openai.com/blog/ai-and-compute/>

Exemple d'une voiture autonome

ResNet-50 à besoin de 7,72 G opérations pour traiter une image 255x255.

→ 230 Gops pour 30 fps

→ 9,4 Tops pour du HD

→ 338 Tops pour 12 caméras et 3 couleurs par caméra

Nvidia

A100

1) Peak rates =
GPU boost clock

2) Effective |
using Sparsity

Peak FP64 ¹	19.3 TFLOPS
Peak FP64 Tensor Core ¹	19.3 TFLOPS
Peak FP32 ¹	19.5 TFLOPS
Peak FP16 ¹	78 TFLOPS
Peak BF16 ¹	39 TFLOPS
Peak TF32 Tensor Core ¹	156 TFLOPS 312 TFLOPS ²
Peak FP16 Tensor Core ¹	312 TFLOPS 624 TFLOPS ²
Peak BF16 Tensor Core ¹	312 TFLOPS 624 TFLOPS ²
Peak INT8 Tensor Core ¹	624 TOPS 1,248 TOPS ²
Peak INT4 Tensor Core ¹	1,248 TOPS 2,496 TOPS ²

Les leaders les plus visibles sont

- Google (Tensorflow, Keras, DeepMind)
- OpenAI (ChatGPT, DALL-E) – partenariat fort avec Microsoft
- Facebook (Torch, PyTorch)
- IBM (Watson)
- Baidu

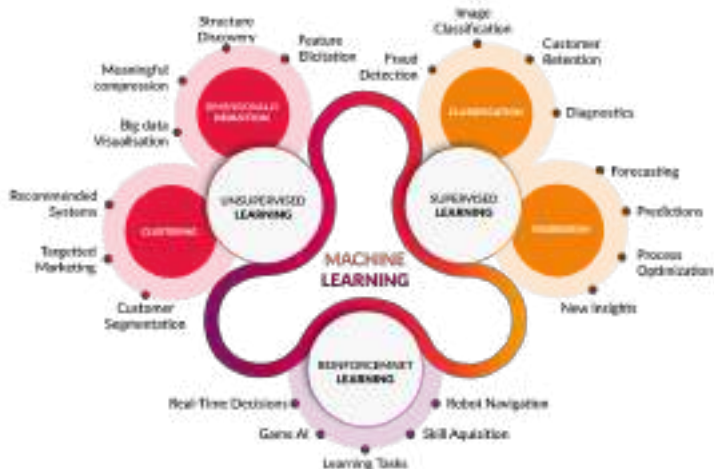
et bien sûr le principal fabricant : NVidia (Cuda, CuDNN)

Ceux qu'on voit moins

A côté de ceux qui participent activement à la recherche et au développement des outils, il y a ceux qui l'utilisent en interne.

- Amazon (Alexa, Amazon Go)
- Apple
- les constructeurs automobiles (Tesla, Uber, tous)
- tout ceux qui font du conseil (Netflix, Expedia...), de la pub (Critéo)
- plein de startups

Types d'apprentissage



L'apprentissage supervisé



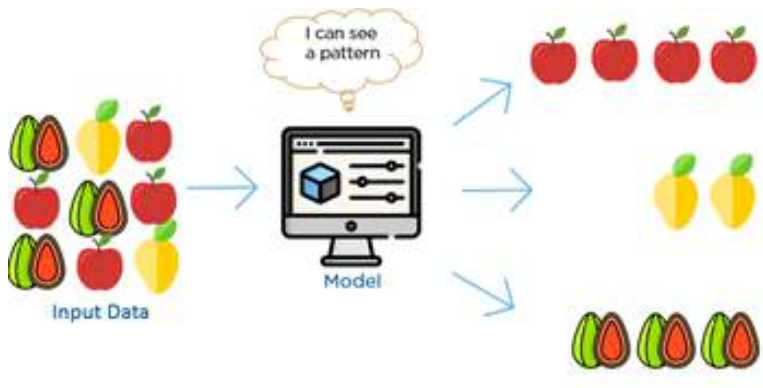
L'apprentissage supervisé

Régression	Classification
Moindres carrés	SVM
Régression polynomiale	Regression logistique Arbre de décision
Réseau neuronal	Réseau neuronal

La révolution vient des réseaux neuronaux :

- Mûr
- Demande des quantités énormes de données étiquetées
- Pas toujours simple à faire marcher
- De plus en plus complexe
- Produit des résultats remarquables en
 - ▶ traitement d'image
 - ▶ traitement de la parole

L'apprentissage non supervisé



- K-moyennes, ACP, des réseaux neuronaux
- Difficile d'en mesurer l'efficacité (besoin de juges humains)
- Usage limité mais en progrès

L'apprentissage auto supervisé

Il s'agit d'un entre deux. On utilise les données intrasèques :

- rotation d'images → trouver l'angle
- découpage d'images → trouver la position

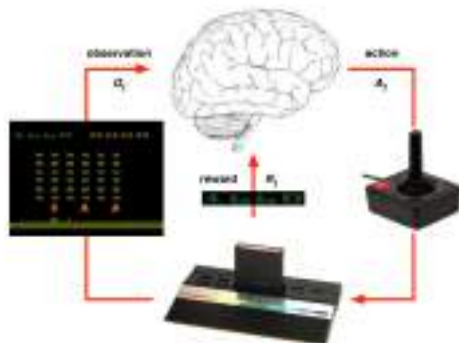
Cela permet d'avoir des premières couche d'un CNN.

Considéré comme la voie pour résoudre le blocage du supervisé.

« le problème principal est que l'I.A. manque de bon sens. Par exemple, elle est incapable de comprendre que j'ai des jambes si mes jambes lui apparaissent cachées derrière ce pupitre. De la même manière elle sera incapable de savoir ce qui se passera avec ce stylo si je le jette en l'air »

Yann LeCun 2020

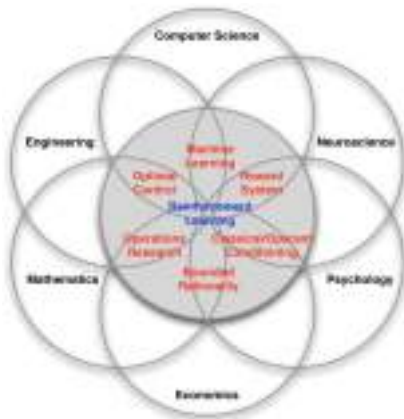
L'apprentissage par renforcement



- Rules of the game are unknown
- Learn directly from interactive game-play
- Pick actions on joystick, see pixels and scores

Points clefs du renforcement

- Pas de superviseur qui connait la solution, seulement une note
- Le retour d'information est décalé (pas immédiat)
- La notion de temps est importante → Système dynamique
- L'agent qui note a un impact sur la suite des données qu'on va recevoir



Test

Quel type d'apprentissage ?

- [Comparaison de CNN pour la vision sur route](#) – 2018
- DeepMind StarCraft II [combat](#) et [explications](#) – 2019
- [Appel au téléphone - Google](#) – 2018
- [Helicopter - Stanford Univ.](#) – 2008
- [Mélodie travaillée - Music VAE](#) – 2018
- Débat : L'État doit-il financer les écoles *pre*-maternelle ? (3 à 4 ans)
[Non – Harish Natarajan](#) [Oui – IBM Debater](#) – 2019
- [Un duo](#) et [l'artiste caché](#) (2019 pour la méthode)

VLC runs the video, F fullscreen, Crtl-Q quit

De AlphaGo à MuZero

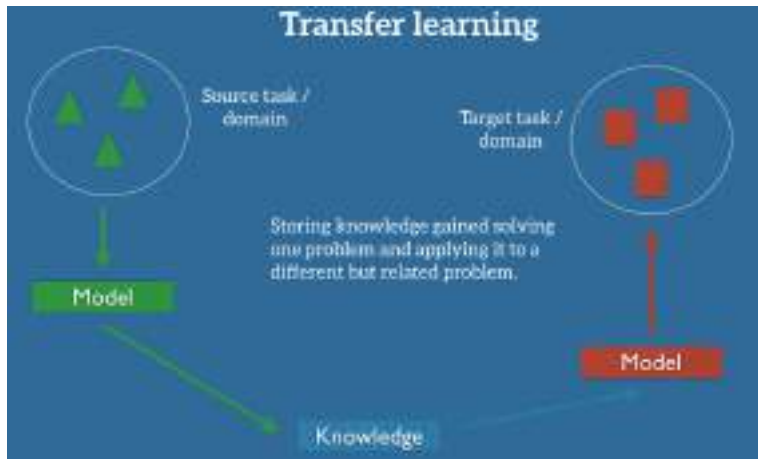


Usage futur des différents types d'apprentissage



Le monde académique/internet et industriel sont différents.

Transfert ML



Ainsi il est tout à fait possible d'utiliser un réseau neuronal entraîné pour une tâche A pour initier l'entraînement du réseau d'une tâche B proche.

IBM IA pour l'industrie

- IBM Watson Recrutement une aide à l'embauche pour les entreprises
- Watson solution pour la vente
- Watson Assistant pour le marketing
- Watson Decision Platform pour l'agriculture
- IBM Equipment Maintenance Assistant pour améliorer la qualité et réduire la maintenance
- IBM Watson Supply Chain Insights

<https://www.ibm.com/watson/ai-for-industries/>