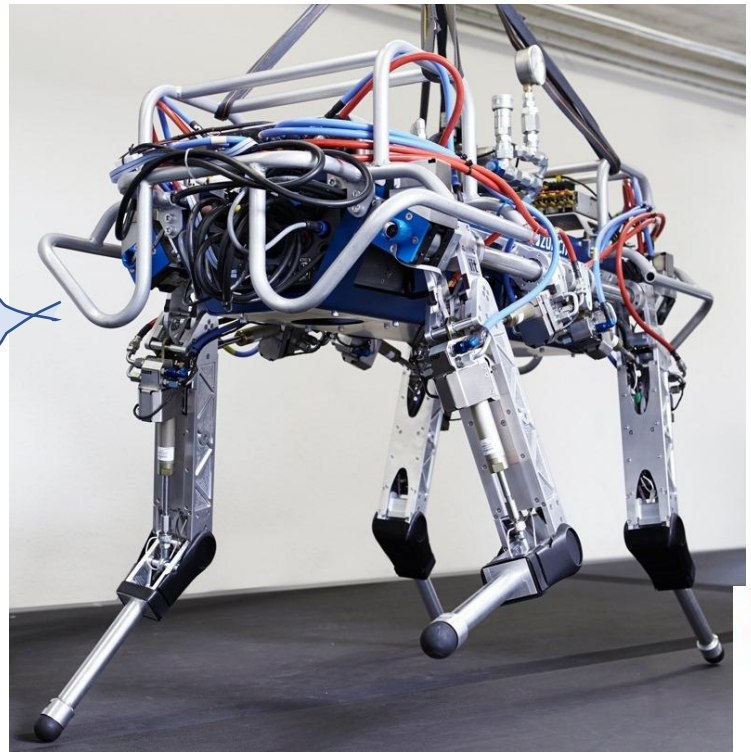
Appel aux adhérents de Old Up !

Je cherche un certain
Blaise PASCAL
Qui aurait dit « l'homme est un robot , le plus faible de la nature mais c'est un robot pensant !



Réponse de l’ami Blaise Pascal



Bonjour amies et amis

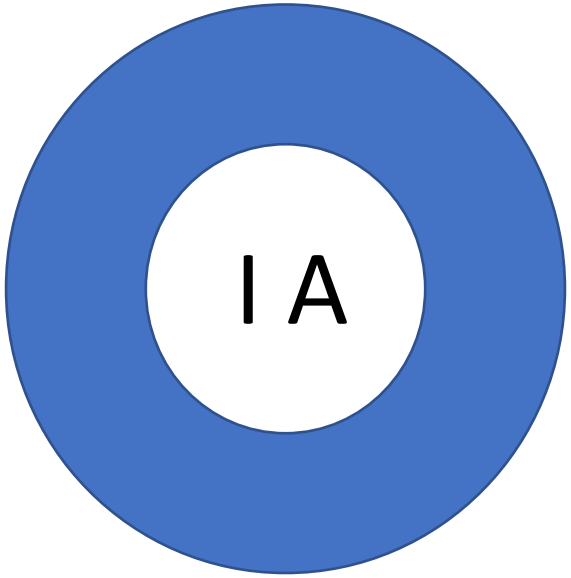
« roseaux » de Old up

Il y aura bientôt 4 siècles
Je m’émerveillait de cette dualité et complémentarité entre
« l’esprit de géométrie et l’esprit de finesse »

I A: nos principaux thèmes de réflexion et points d'entrée

Comprendre
Les systèmes
D' I A

Les usages:
Typologie
et métiers



Opportunités et
risques

La gouvernance ?
La régulation ? ,

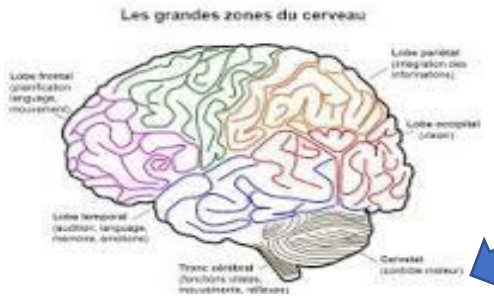
Les acteurs de
l'offre

comprendre les systèmes d'IA

avec *Yann Le Cun*;

focus sur « les machines apprenantes »

nos deux « apprenants » !



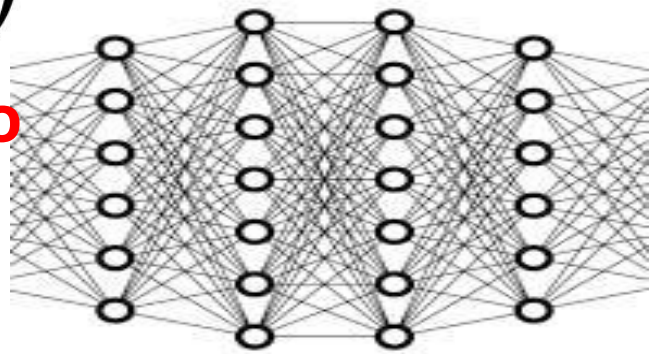


Deux logiciens et statisticiens !



$$\nabla f$$

In vivo vs **in silico**



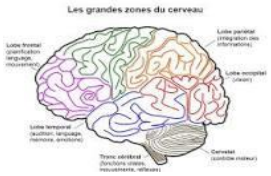
Stanislas **DEHAENE** et Jean **PIAGET**

Geoffrey HINTON et Yann LE CUN !

in vivo vs in silico !

humaine cognition

Sujets…
animés !



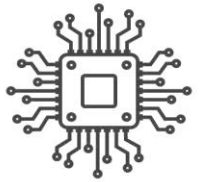




Inhumaine cognition

Objets…
inanimés !







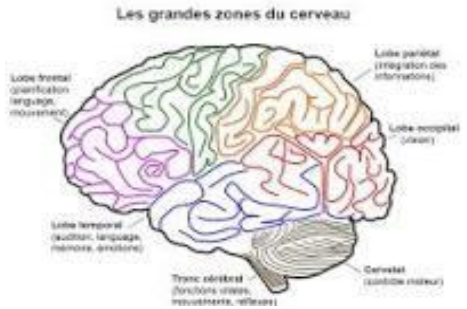
La « machinerie » biologique et sociale

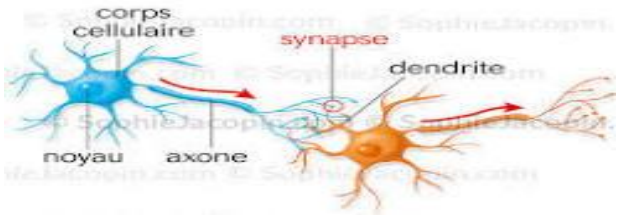
La « ferraille » et les algorithmes

in vivo vs in silico - suite !

humaine cognition

Sujets animés !





Courant ionique m/S

Parallélisme massif

100 Milliards de neurones

10^{11}

100 000 MILLIARDS DE SYNAPSES 10^{14}

Puissance requise: 25 WattS !



2 mots clés

Réseaux

Et

Systèmes

Inhumaine puissance de calcul et d' inférences

Objets inanimés !





25.000 $ pièce
Système multi
400.000 $

Propagation electromagnétique

2/3 lumière

Calculateur « Jean Zay »

125 900 000 000 000 000 ($> 10^{15}$ op/sec !

Puissance pour 1 seul GPU: 250 watts !

Les « formateurs »





Le parcours et l’apport de Yann Le Cun



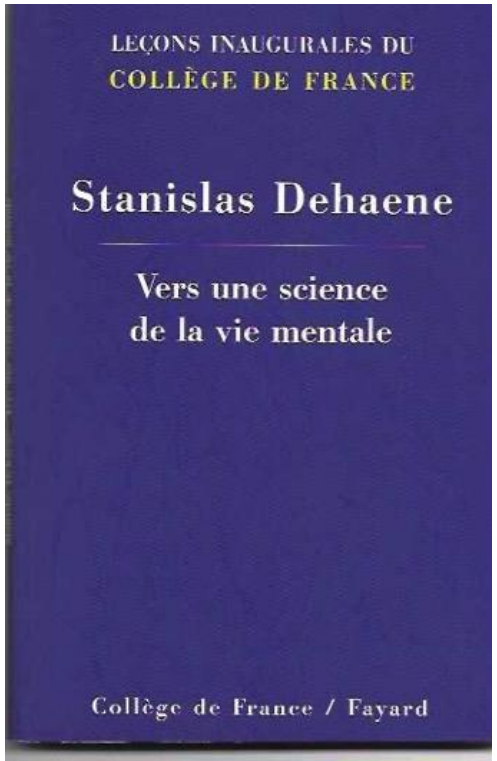
La démarche de **Yann Le Cun** par lui-même exposé à Sciences Po Paris - Septembre 2026. https://www.youtube.com/watch?v=Y4s8NadbZfU

- Eléments clés
  - - Son parcours…ESIEE (1983)…Toronto..Bell Labs… …META..Paris AMI Advanced Machine Intelligence Labs
  - - la reconnaissance de formes et les réseaux convolutifs
  - - les LLM
  - - Les World models
  - - l'Open source

retour sur **Stanislas Dehaene** et son livre **« APPRENDRE »**
« les talents du cerveau, le défi des machines »
RAPPEL réunion du 5 novembre 2019 !!!!! Claude JUHAN....JCB



Leçon inaugurale du 27 avril 2006
« vers une science de la vie mentale »



Stanislas DEHAENE : Qu'est-ce qu'apprendre ?

Comment un réseau de neurones apprend ?

« Apprendre, c'est: construire au sein de notre cerveau un nouveau modèle du monde » .

7 items clés au défi des machines:

- **ajuster les paramètres d'un modèle mental**
- **exploiter le potentiel de la combinatoire**
- **minimiser ses erreurs**
- **explorer l'espace des possibles**
- **optimiser une fonction de récompense**
- **restreindre l'espace de recherche**
- **projeter des hypothèses a priori**

1-ajuster les paramètres d'un modèle mental !

- Le défi du moustique !





- Cas des modèles de langues et de la hiérarchie dans une phrase

  En français: « mon oncle veut habiter à Paris »

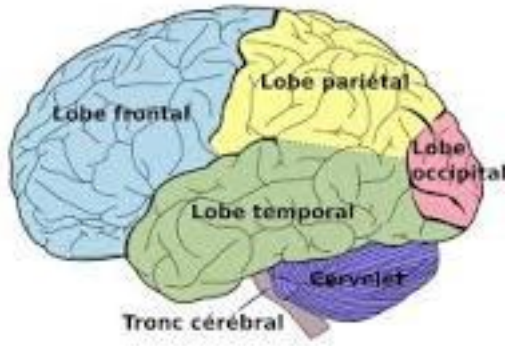
  En japonais : « Oncle mon Paris à habiter veut «

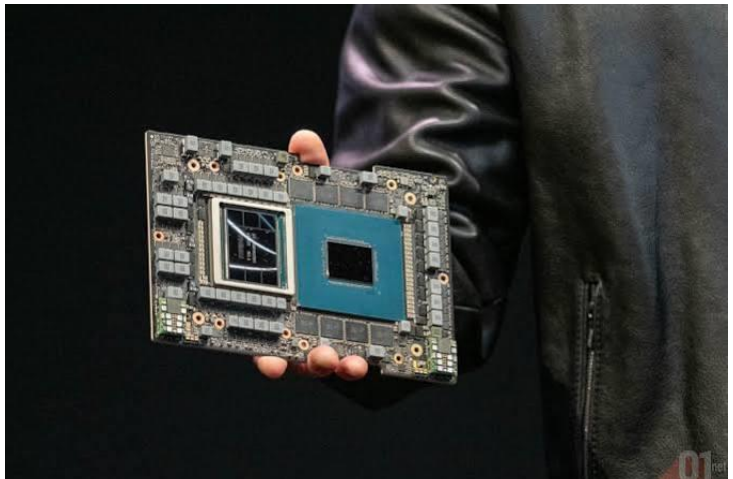
2- exploiter le potentiel de la combinatoire

- Exemple du lexique mental : 50 000 mots aux significations les plus diverses.
- Les vrais mots sont composés d'une **pyramide de syllabes** qui **s'assemblent selon des règles strictes**
- Les phrases sont des **assemblages réguliers** de mots qui sont des **assemblages réguliers** de syllabes qui sont des **assemblages réguliers** de phonèmes
- Apprendre une langue , c'est découvrir les paramètres qui régissent cette combinatoire à tous les niveaux……

Au travail !…..







3- minimiser ses erreurs

- image du système

  tireur – fusil – cible



- Le tireur, système homéostatique inconscient, s' adapte à l'environnement de l'action

- Le tireur conscient, par essais successifs et corrections ajuste sa position, sa respiration et la mire…il apprend de ses erreurs…, ses itérations, ce feed back lui permettent d'atteindre son objectif……le nombre de paramètres objectifs et conscients est ici limité……ce mécanisme itératifs de rétropropagation est à la base des **systèmes d'IA qui gèrent et ajustent des millions de paramètres par une opération mathématique dont Yann Le Cun a été le pionnier .**

4- Explorer l’espace des possibles

- L’arbre remarquable de Cucuron ou Grambois ?



5- optimiser une fonction de récompense

- On ne m’apprend pas à marcher en me disant quels muscles contracter ou détendre ! **J’agis**



- J’apprends sur la base du résultat

Je tombe ou je traverse la pièce !

Je m’évalue

On apprend aux machines à **agir** et à **s’autoévaluer**, passage de l’apprentissage supervisé à l’apprentissage par récompense qui est au cœur de **l’apprentissage par renforcement**

6- restreindre l'espace de recherche

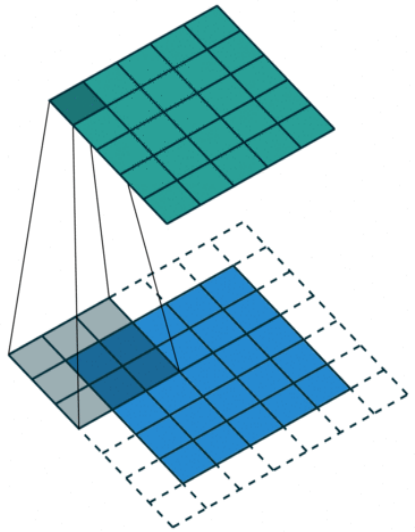
Un apport essentiel de Yann Le Cun…

…….**les réseaux de convolution** !

Le but: Eviter le sur-apprentissage en réduisant le nombre de paramètres à ajuster.

Le principe: appliquer en chaque zone d'une image , le même algorithme …exemple: rechercher un visage : une ovale, une paire d'yeux etc un même filtre qui parcourt l'image

Ex: cadrage des visages sur votre Smartphone !



Filtrage moyenneur glissant

7- projeter des hypothèses a priori

Réseaux à convolution….

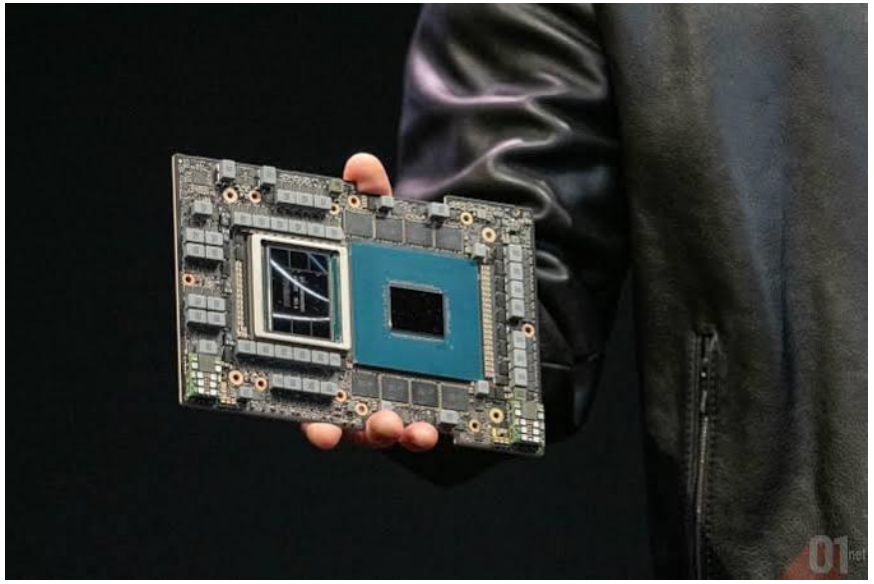
- « vous avez dit à connaissances innées ? »
- « Oui car ils incorporent dans leur architecture même une hypothèse forte: ce que j'apprends en un endroit peut se généraliser partout ailleurs: je dois reconnaitre un objet (ex un visage) quelles que soient sa position, sa taille, même s'il bouge ! c'est l'invariance
- « voulez-vous dire que le système apprend l'invariance ?
- « non, il ne l'apprend pas, il la suppose vraie a priori et l'utilise pour réduire l'espace d'apprentissage »

La morale de l'histoire:

L' apprentissage ne peut pas tout , il s'appuie presque toujours sur des a priori, sur des hypothèses par défaut …plus il y a d'hypothèses « innées » plus l'apprentissage sera rapide…

la parole à JP Changeux dans « l'Homme neuronal »: « Apprendre, c'est éliminer »

Sous le capot des calculateurs stand-alone et, par milliers, dans des data centers !



Puce NVIDIA Blackwell B200

200 milliards de transistors !

18 à 20 petaflpops ! 10^{15}

30.000 euros pièce !
Un hyperscaler :100.000 puces !
Soit 3 milliards d’ euros !

Sous le capot des data centers: vous avez dit « big data » !

Volume de données générées dans le monde en une année (Flux):

Estimation 2024 – source U.E- Statista

175 à 180 zetta octets 10^{21} !!

1 livre $\approx$ 10 millions d'octets 10^{7}

Volume estimé de l'archivage annuel (1 à 2 % du flux (Stock du Cloud)

Soit 1 à 2 zetta octets (10^{21})

Ordre de grandeur : 10^{14} équivalent livres

Soit pour une planète de 10^{10} habitants

un équivalent 10.OOO livres par an !!!!et par personne !!!

K. Fukushima + Y. Le Cun = Convolution !

- **K.Fukushima**, Inspiré par les découvertes sur le cortex visuel et notamment l'arrangement et la spécialisation des cellules du champ visuel , conçoit le néocognitron.

A retenir :

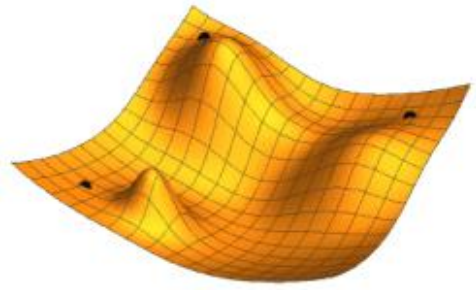
- le principe de pavage, de tuilage de l'image,

- à chaque tuile son neurone artificiel qui affecte un poids à chaque pixel,

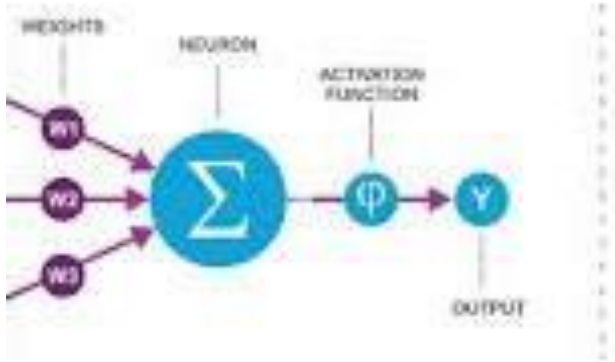
- l'analyse de la variation spatiale, **le gradient**, de chacun des paramètres de la tuile vers ses voisines.

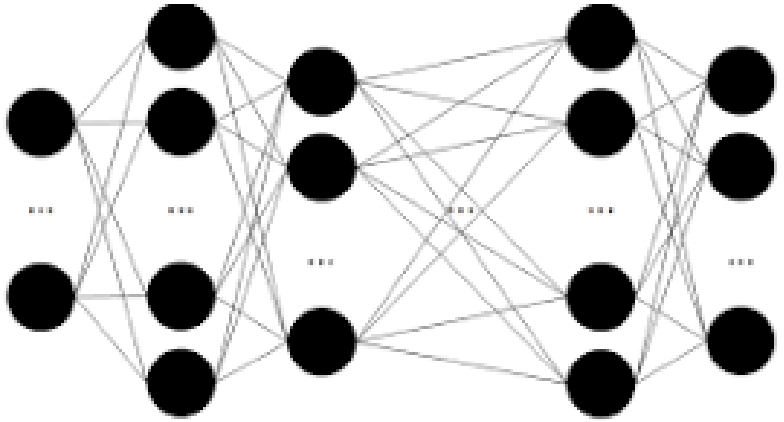
- **Y Le Cun et G Hinton** généraliseront et enrichiront (rétropropagation de gradient) cette approche aux

réseaux multicouches de l 'apprentissage profond



Je me présente à nouveau: neurone artificiel !

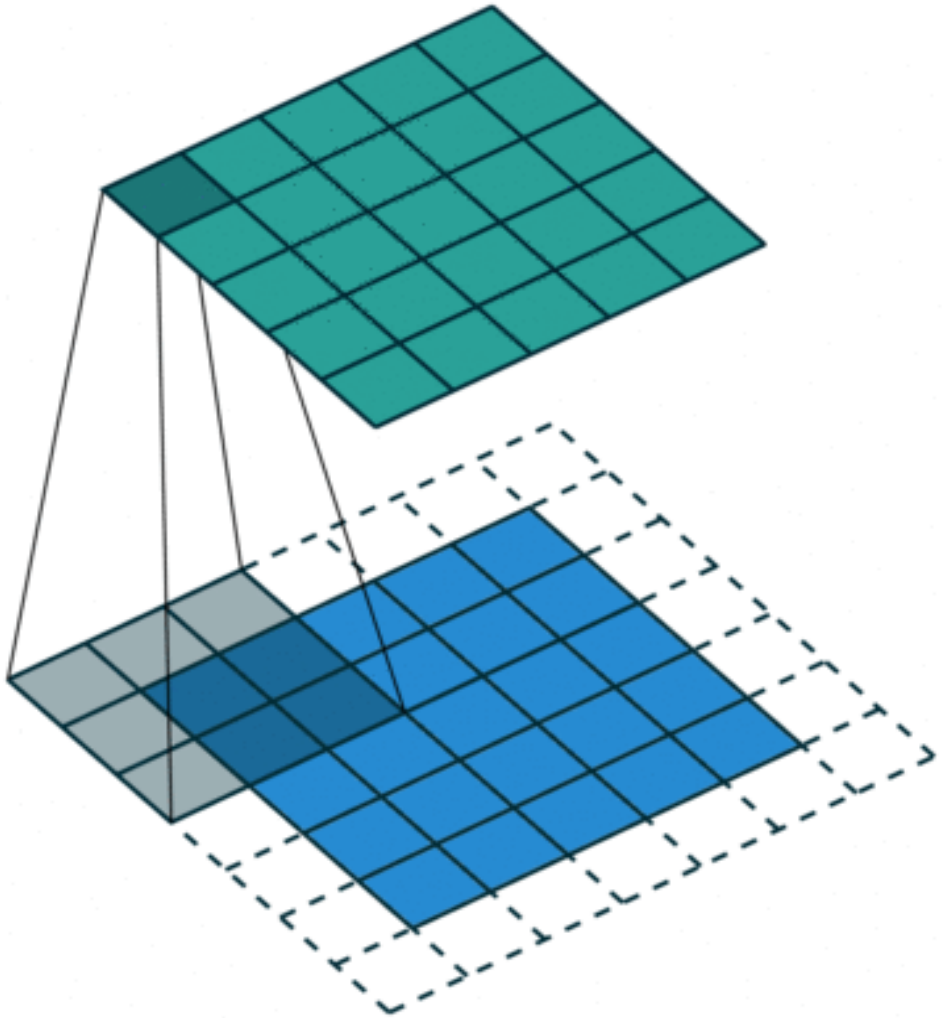






$$\nabla f$$

Reconnaissance des formes
La machine « observe », pixelise, calcule et reconnait !



Rappel 1 : 2 approches de l' I A

- Approche 1 dite **« symbolique »:** de **la logique avant toute chose** !

- Des données considérées comme vraies, des **mesures**, des **règles explicites** et les **algorithmes** qui leur correspondent.

Tous les humains sont mortels
+
Socrate est un humain
=
Socrate est mortel

Des faits vrais, des règles…c'est le temps des « systèmes experts »

– – – – –

- Approche 2 dite **« statistique »** s'appuie sur des **observations du monde réel** pour en extraire des **modèles généraux**.

- On appelle cela l'apprentissage automatique, **c' est le cœur de l'IA générative d'aujourd'hui**.

les chiens ont 4 pattes
Les chevaux ont 4 pattes
Les moutons ont 4 pattes
=
Les mammifères ont « très probablement » 4 pattes

- - - - -

Symbolique ou statistique
2 types de programmation des systèmes

La programmation impérative **– le COMMENT**

Le développeur décrit pas à pas la séquence d'étapes que l'ordinateur doit exécuter pour parvenir à un résultat.(langages C, C++, Java etc, Python version procédurale etc).

La programmation déclarative **-le QUOI**

Le développeur décrit le résultat souhaité sans détailler les instructions sous jacentes, c'est le moteur de recherches qui s'en charge….l'objectif final est d'établir des relations entre les données observées (langages et outils HTML, SQL, React etc…)

Approche statistique , vous avez dit : « observation du monde réel et génération de modèles »

2 innovations clés et 2 processus clés

La reconnaissance de formes et « la convolution »

Les interfaces homme-machine en langage naturel – les LLM- « l'apprentissage profond » et « l'attention »

+ une

L'accès mondial instantané à des bases de connaissances actualisées.

Explications et exemples

Types d’apprentissage pour chacune des innovations;
reconnaissance de formes et langage naturel

- Apprentissage supervisé
- Apprentissage non supervisé
- Apprentissage par renforcement
- Apprentissage profond , deep learning, LLM

supervisé ou non supervisé ?

supervisé

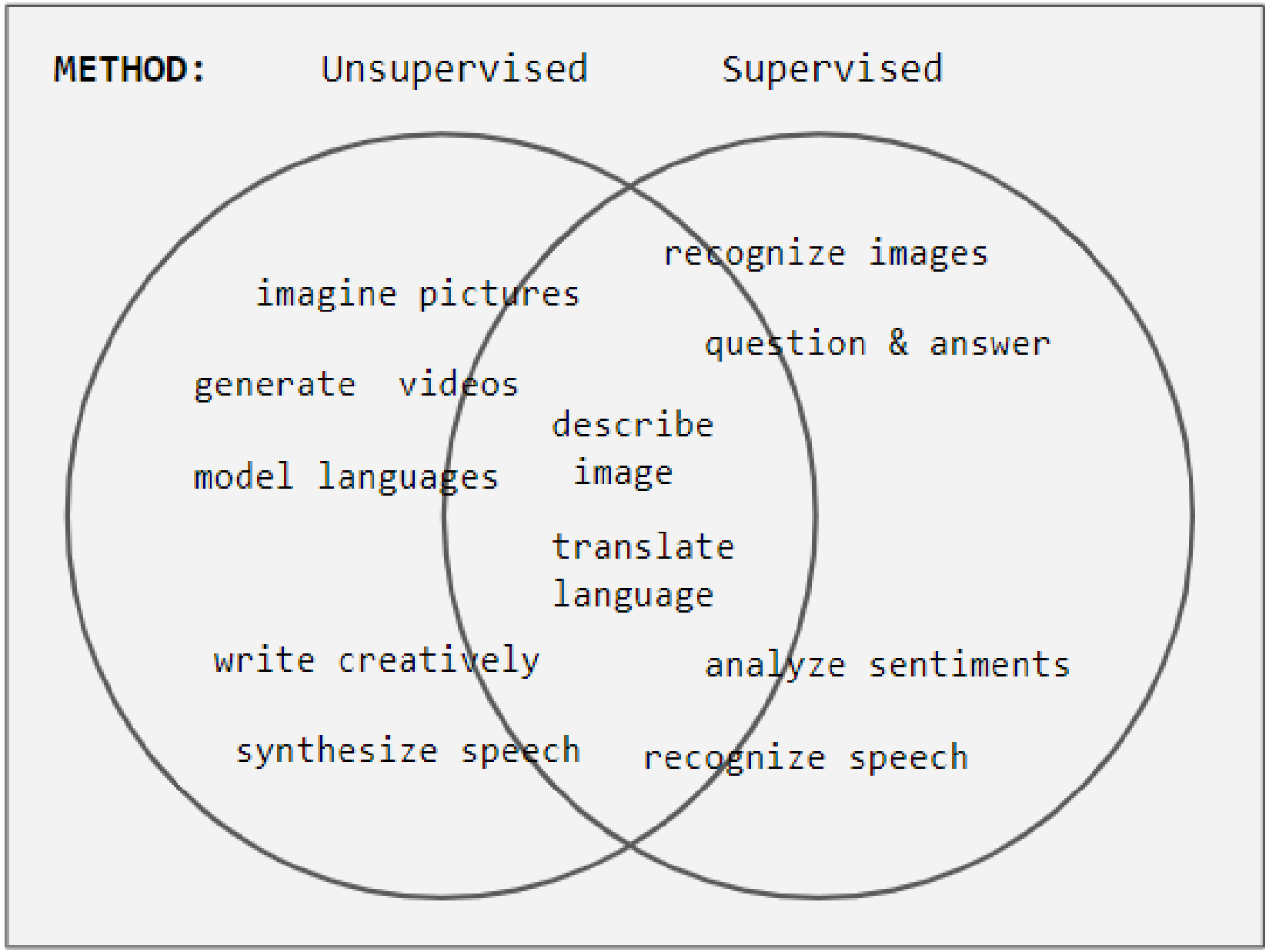
- « Le modèle est entrainé sur des données pour lesquelles le résultat est connu » - étiquetage.

« Entrainer un réseau, c'est, pas à pas, forger l'intelligence à partir des données, c'est optimiser les prédictions ».O Cappé

non supervisé

- « L'algorithme découvre des régularités et des structures par lui-même »

Supervisé , non supervisé



Apprentissage supervisé par étiquetage humain

- On m'a présenté 100 000 images d'animaux étiquettées « chien » par un humain…..

- j'ai « pixellisé, vectorisé, pondéré et modélisé » toutes ces images de « chien »….(voir slide suivant)

- On m'a « entrainé » à reconnaitre des chiens dans un paysage et au milieux d'autres animaux à pattes ou sans….

- À chaque erreur on a modifié les « paramètres de mon modèle ».

- je ne commets pratiquement plus d'erreurs, ***j'en ai fini avec « l'apprentissage automatique supervisé »***

Devenu « expert », Je vais pouvoir maintenant **« reconnaitre seul »** et
- **« générer »** moi-même le mot et des images de l'objet « chien »

Après attribution à tous mes « pixels » d'un « poids », je me présente : « chien » (source O. Cappé)



= 13959

CHIEN

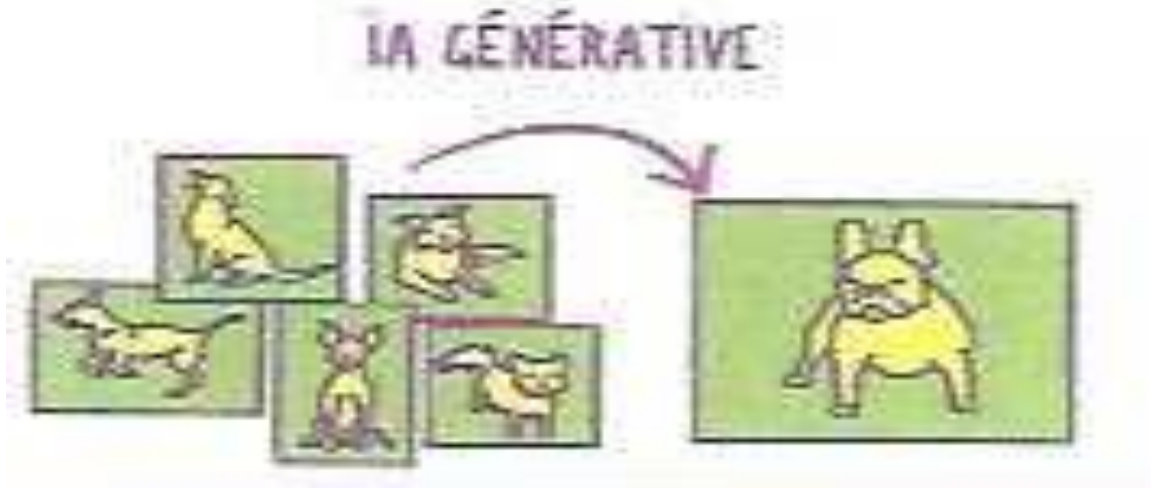
Si cette somme dépasse un certain seuil, le modèle décide que l'image représente un chien

CHIEN

Sinon, il considère l'image comme "pas un chien"

Moi « chien » en IA générative (source O. Cappé)

- Je peux « générer » une image de « chien »: <u>IA non conditionnelle</u>



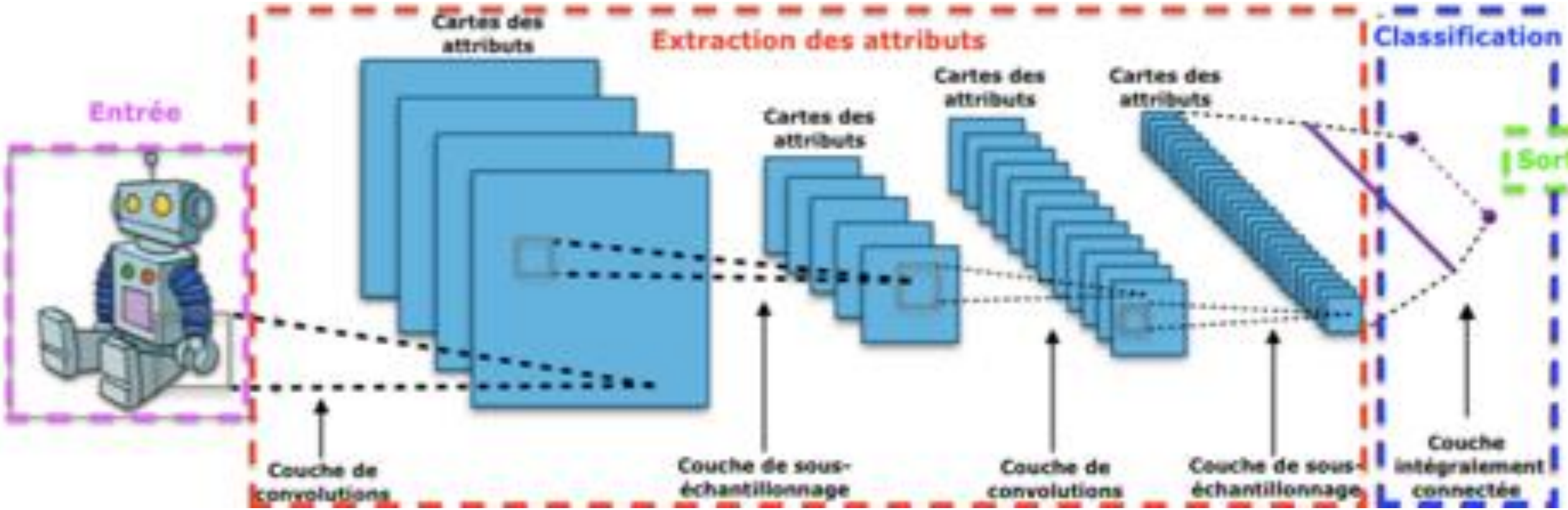
- Je peux « générer » un « berger allemand »: <u>IA conditionnelle</u> (signal de conditionnement par l'utilisateur))



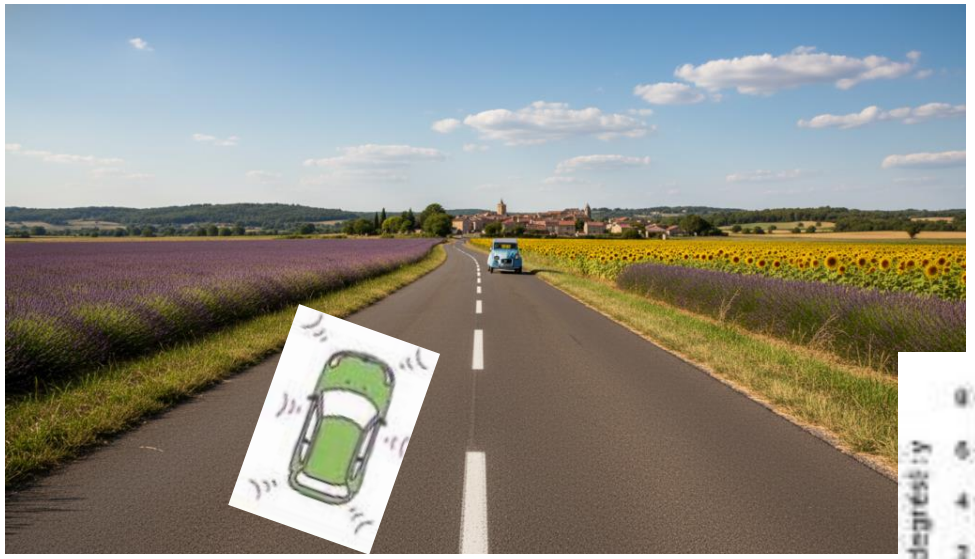
Reconnaissance de formes

- 5 étapes majeures (wikipédia)
- **1 acquisition** de l'image. Capture des données brutes via un capteur numérique (caméra, scanner, capteur biométrique etc)
- **2 Prétraitement**. Amélioration de la qualité des données (filtrage, amélioration des contrastes etc
- **3 Segmentation**. Isolement de la forme de l'objet d'intérêt du reste de l'image (le fond) Cette étape sépare la scène en régions homogènes.
- **4 Extraction de caractéristiques**. Sélection et mesure de propriétés discriminantes de la forme, (contours,, texture, histogramme de couleurs, points clés. Cette étape réduit la quantité de données à analyser en ne gardant que les descripteurs pertinents.
- **5 Classifcation** et décision. Un algorithme compare ces caractéristiques avec une base de données ou un modèle d'apprentissage pour attribuer une étiquette à la forme et valider son **IDENTIFICATION**.

Schéma type pour la reconnaissance des formes

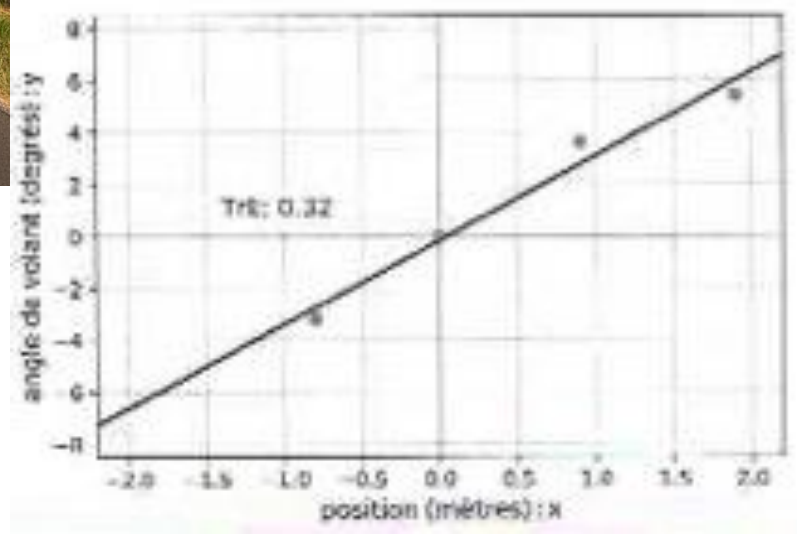


Autre exemple d'apprentissage supervisé
y a pas que les pixels……la minimisation de l'erreur



La supervision
Le conducteur
au volant



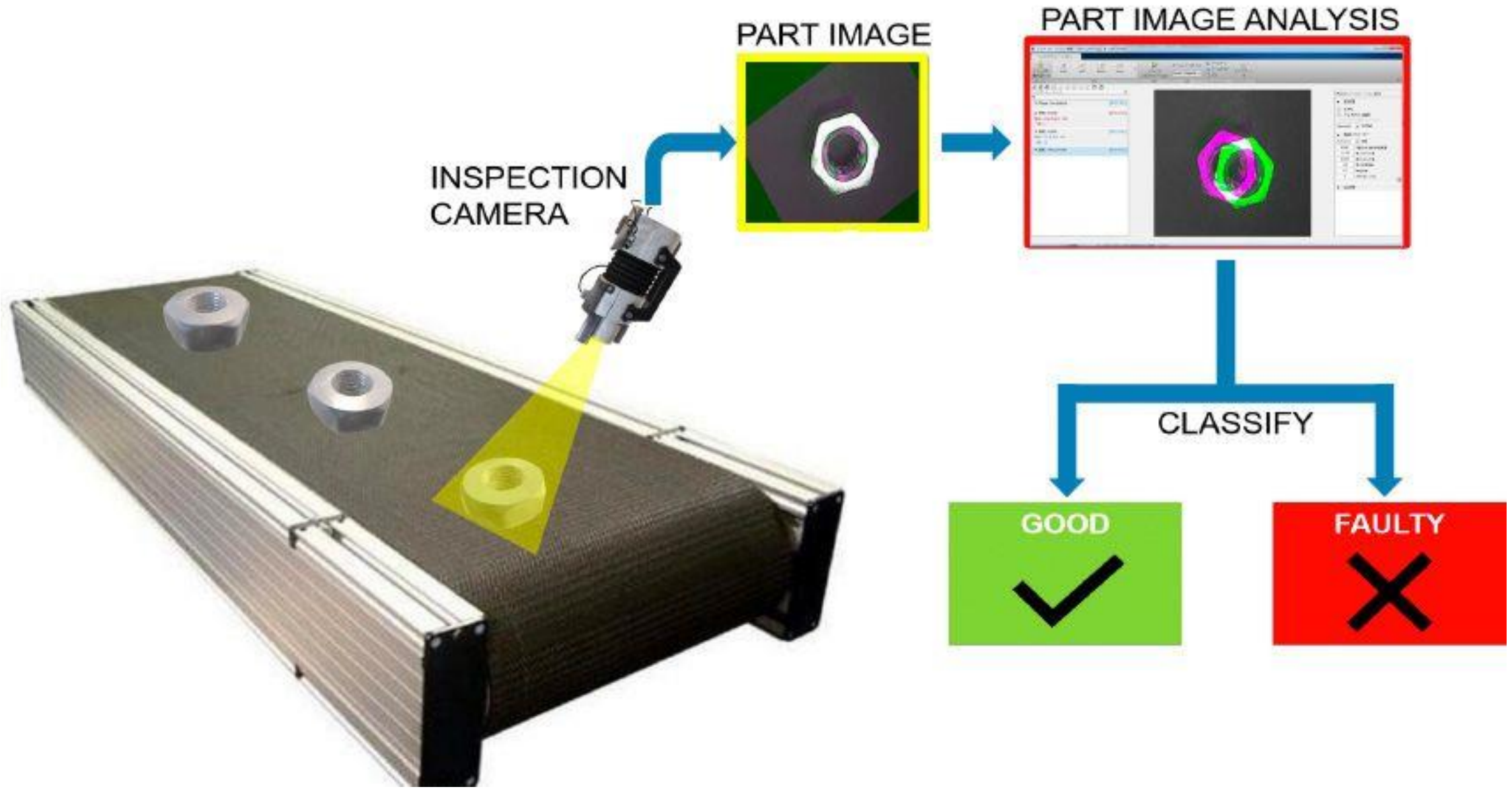


Une camera enregistre tous les 1/10 de sec
la position de la voiture les X
On enregistre l'angle du volant les Y
On trace la courbe des 36000 points (regression)
Les X prédisent les Y
La voiture est « autonome » (presque !)

après



Contrôle qualité de pièces

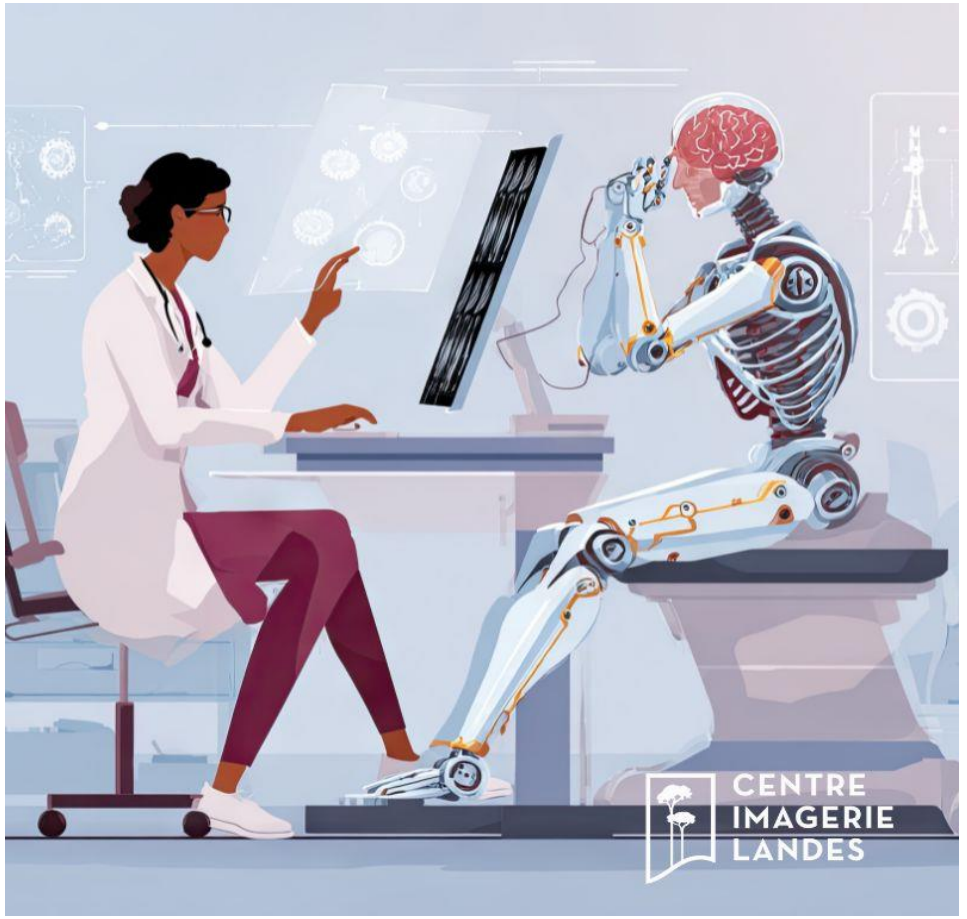


Retour aux pixels ! : L'aide au diagnostic médical

Les systèmes d'acquisition et d'analyse des images médicales:

Assistant, pas remplaçant

- Voir l'invisible
- Meilleure précision
- Accès à des bases de connaissances de référence
- Gain de temps
- Aide au diagnostic



Prix Nobel de chimie 2024
David Baker professeur de biologie
Demis Hassabis et John Jumper ingénieurs chez Deepmind

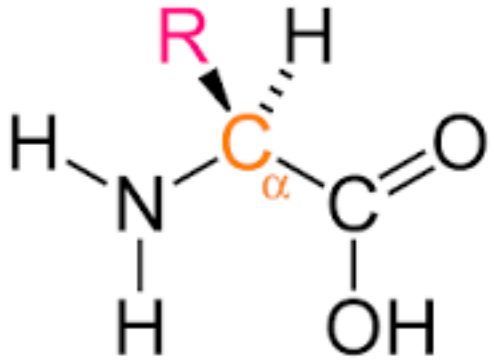
Objet: Design des protéines et prédiction de leur structure !

- **Une protéine est une chaîne d'acides aminés (des dizaines à plusieurs milliers selon les protéines).**
- **La chaîne d'acides aminés se replie en une structure tri-dimensionnelle qui dicte la fonctionnalité biologique de la protéine.**

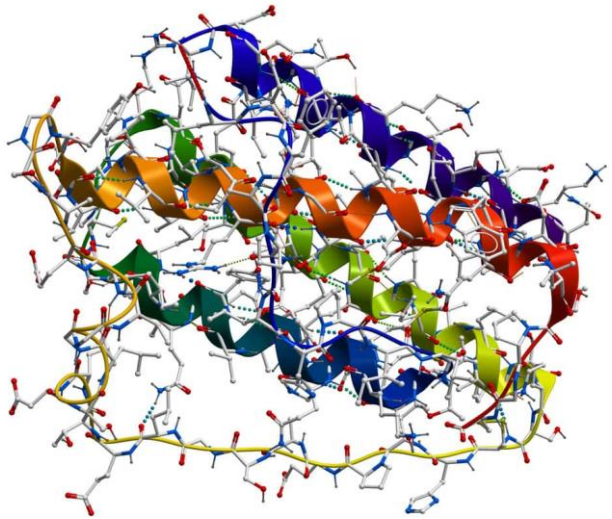
Ps de JCB: vous avez dit « chimie computationnelle » !

Le logiciel d'IA AlphaFold

- A partir de l'observation et de l'entrainement du logiciel sur plus de 175 000 structures protéiniques connues, l'IA modélise des formes 3D de millions de protéines !!







- Mise en ligne, par DeepMind , en accès ouvert d'une base de 365.000 protéines humaines devant évoluer vers des millions de diverses espèces !

Une question dans la salle ! À partir de cet exemple: Comment l'IA change-t-elle les méthodes scientifiques ?

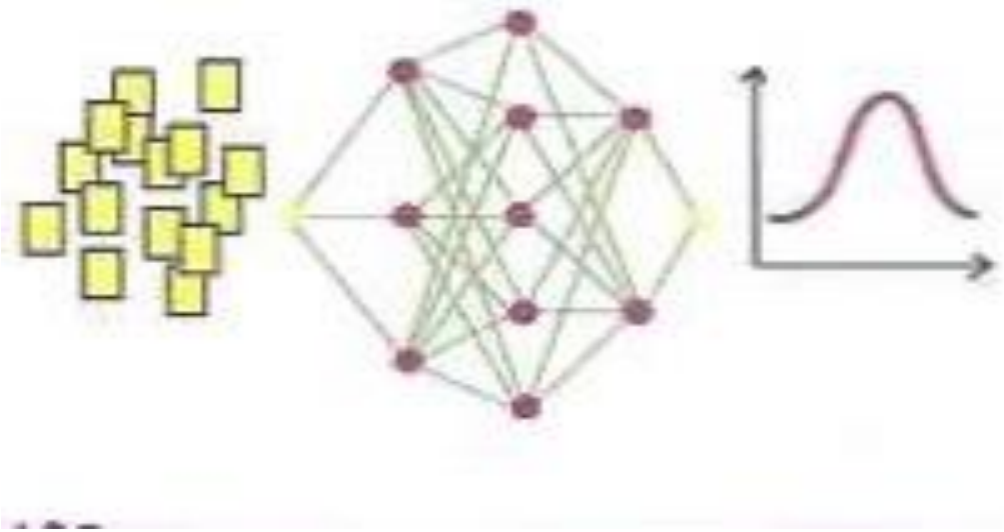
- L'approche hypothético- déductive.

- L'approche inductive……observer, mesurer, mettre en évidence des corrélations (≠ des relations de cause à effet) et en extraire des connaissances.

- Capter des motifs statistiques.

- Proposer des prédictions ou

des généralisations.



2 exemples qui ont agité les esprits…
et il y avait de quoi !

Deep Blue



Alpha Go



Deep blue et alpha Go



- **Deep blue** (IBM 1977).

- Système symbolique et recherche arborescente . Il calculait entre 100 et 200 millions de positions par seconde. Sa fonction d évaluation était entièrement codée par des programmeurs et de grands maitres d'échec. Il n'apprenait pas par lui-même mais exploitait d'immenses arbres de possibilités.

Alpha go (Deep Mind Google 2016) .



Réseaux de neurones + recherche arborescente.
Alpha go n'utilise pas des règles écrites par des humains, il apprend en analysant des millions de parties humaines puis en jouant contre lui-même des millions de fois pour inventer ses propres stratégies !

Puis…Zero est arrivé!



Alpha Go Zero



- Une version « **symbolique »** conçue en donnant au système les règles et sans aucune donnée humaine et en faisant jouer des machines entre elles - **apprentissage par renforcement**

- Alpha go Zéro a joué 4,9 millions de jeu contre lui-même dans les 3 premiers jours d'entrainement et a battu….alpha go
- 64 processeurs graphiques GPU et 19 processeurs CPU (coût 25 millions $)

- C'est une percée majeure pour l'IA (talents surhumains ?…et alors ?)

Retour à Chomsky

Grammaire générative, transformeur…l'IA d'aujourd'hui emprunte aux travaux, aux concepts et au vocabulaire du grand linguiste **Noam CHOMSKY**

- Sa proposition de « grammaires universellles » a inspiré les premiers modèles d'IA symbolique (cf Alain Colmerauer) qui cherchaient à reproduire des règles logiques et formelles pour comprendre et générer du langage. Il a inspiré les premiers **NLP** Natural Language Processing. Modélisation basée sur des règles.

- Ses critiques de l'IA statistique et du « deep learning » ont nourri les réflexions sur les limites et ambitions de l'IA.

Comment fonctionne un grand modèle de langue -1-

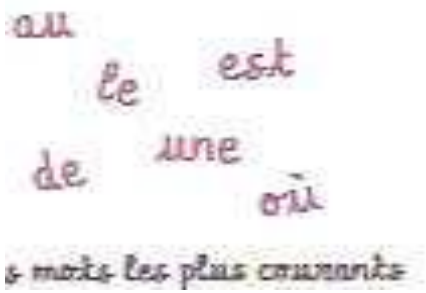
- source principale : Olivier Cappé

- Approche très différente des NLP initiaux

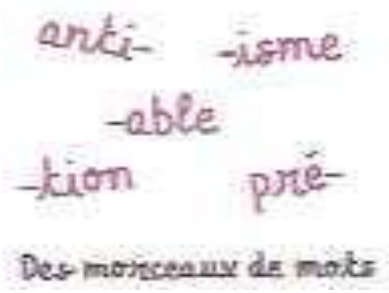
- Les LLM ne « comprennent pas » les textes, ils les « représentent » sous forme de séquences d'unités individuelles: les tokens

Exemple:

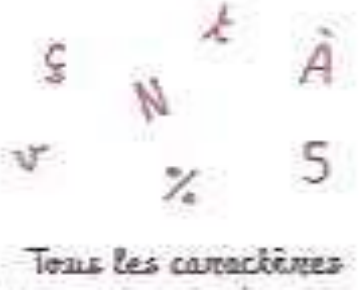
Une « soupe » d'unités individuelles (30 000 à 150 000 ! pour un modèle !)



5 mots les plus courants



Des morceaux de mots



Tous les caractères

« les marguerites s'effeuillent au ralenti »

devient

« les », « mar# », « guer # » « ites », « s' », « 'eff #»,

« eu# », « illent », « au », « ralenti », «. »

Comment fonctionne un grand modèle de langue -2-

- source principale : Olivier Cappé

Comment représenter ces unités individuelles et leur sens ?

La parole aux linguistes !

- Chaque unité individuelle est représentée par un ensemble de vecteurs correspondants aux « dimensions » de l'unité (de 500 à 1000) dimensions….l'unité devient….MESURABLE

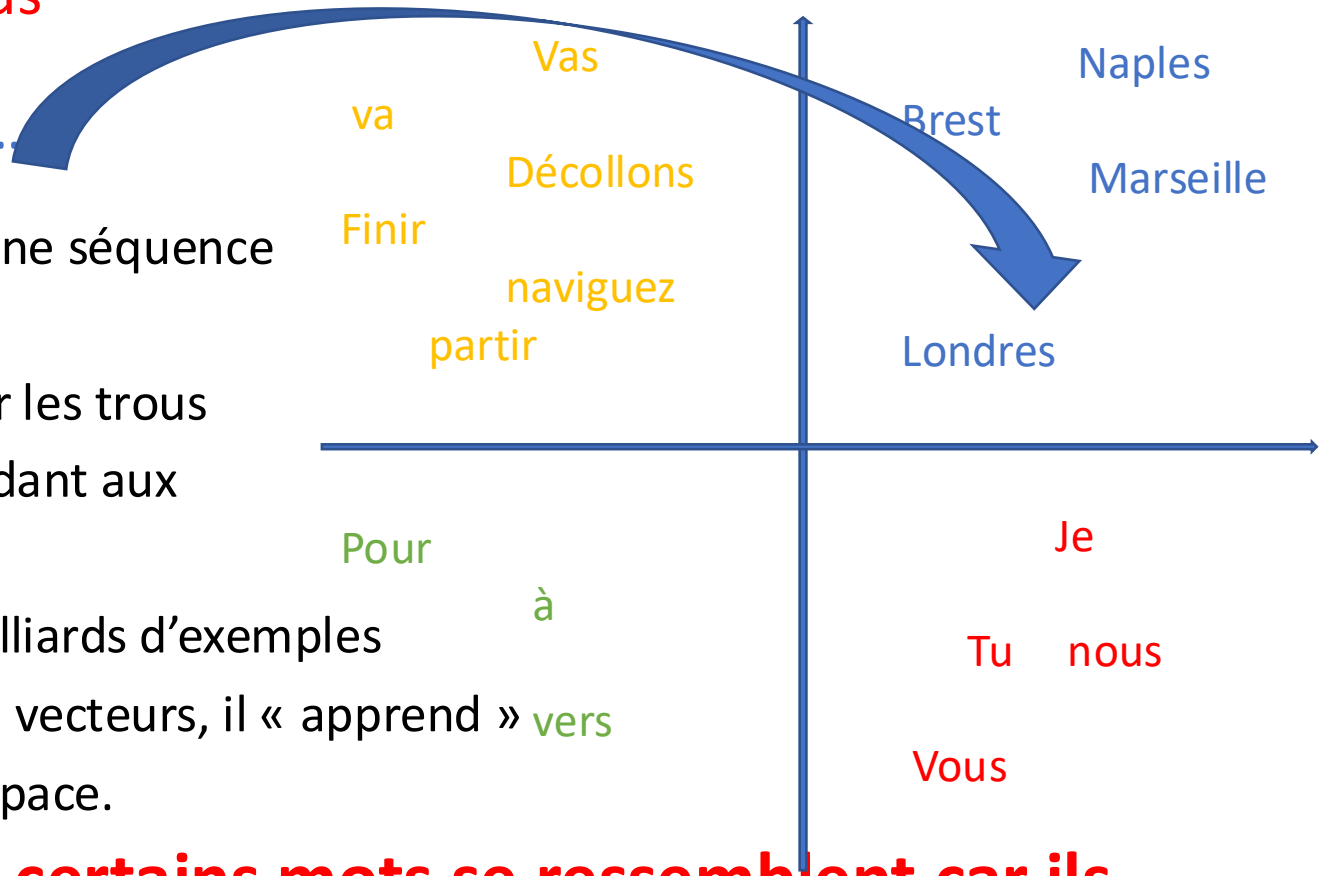
  - « ces dimensions et leur interprétation sont apprises automatiquement en s'appuyant sur un principe linguistique clé: « deux unités se ressemblent si elles apparaissent dans des contextes similaires »

- L'apprentissage se fait en faisant jouer un réseau de neurones artificiels au jeu des textes à trous à partir d'énormes quantités de textes….le modèle apprend que certains mots se ressemblent car ils partagent des contextes similaires….et la RESSEMBLANCE est MESURABLE

Le modèle « apprend »…les « ressemblances ! »

la phrase à trous

Vous naviguez vers ……



1- la phrase à trous est convertie en une séquence De vecteurs

2- le réseau de neurones doit combler les trous en proposant des vecteurs correspondant aux unités absentes

3- en répétant cet exercice sur des milliards d'exemples Le modèle ajuste progressivement les vecteurs, il « apprend » ainsi à positionner les unités dans l'espace.

4- Le modèle apprend que certains mots se ressemblent car ils partagent des contextes similaires

« apprentissage »..... suite

Le sens de chaque mot dépend du contexte et des mots voisins

La « convolution » entre en jeu (c'est une opération mathématique de calcul d'une somme pondérée des unités vectorisées)

Les réseaux de convolution capturent le contexte immédiat, ils analysent les mots voisins et créent des vecteurs enrichis.

Cette première contextualisation prépare l'étape finale qui est le mécanisme **de l'attention**

« L’attention » source Mistral

- Mettre en relation chaque mot ou token avec les autres mots du texte pour le comprendre
- Chaque token est transformé en vecteur qui sera comparé aux autres puis en retour, après pondération, le vecteur initial sera enrichi.
- Les tokens « s’observent entre eux » et ceci non pas séquentiellement mais en parallèle
- Les LLM génératifs ne regardent que les mots précédents, pas les suivants..Ils sont prédictifs
- C’est la capacité à relier dynamiquement n’importe quels tokens qui explique la puissance des LLM modernes

L'heure est venue de me présenter, je m'appelle : GPT

Generative Pretrained Transformer

Le cœur des systèmes génératifs:

- Le **« transformer »** (ou modèle auto-attentif) est une architecture d'apprentissage profond introduite en 2017.
- Il améliore la contextualisation en prenant en compte l'ensemble d'une séquence (une phrase, une image) grâce à un mécanisme clé appelé « l'attention »
- L'attention aide le modèle à déterminer quels mots d'une phrase sont les plus significatifs pour comprendre le contexte. Ce mécanisme génère un vecteur unique qui possède une pondération spécifique à chaque mot.
- L'approche est statistique, probabiliste mais, entrainements après entrainements et corrections après corrections sur des milliards de phrases ou d'images…..les résultats deviennent ceux attendus d'une compréhension humaine.La machine réussit au-delà des espérances…….le test de TURING !

Synthèse sur les **LLM**….c'est extraordinaire !

le lexique → la syntaxe → la sémantique

les mots → les règles → le sens ?

Le sens !!!

→ décomposition en « unités linguistiques »

→ détection des ressemblances par les contextes

→ vectorisation des ressemblances

→ détection des termes significatifs - **contexte**

→ **« transformation », création d'un vecteur unique de pondération**

Du bon usage « intelligent » d'un outil extraordinaire….qui ne l'est pas…intelligent !

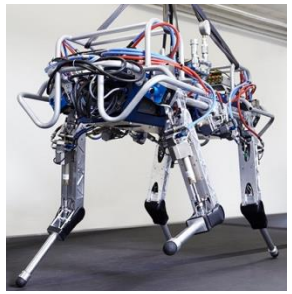
Qualité de la question

- « *tu vois le château avec la tour* »
- « *je veux laver ma voiture, la station de lavage est à 200 mètres, j'y vais à pied ou en voiture ?* »
- Etc..

Analyse critique, compréhension de la réponse, traquage des explications complémentaires et les sources ….bref, validation intelligente !

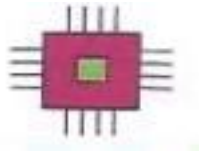
Passage par les robots

- Que suis-je ?









Moteurs capteurs ordinateurs

Je suis un automate programmable…..

- Que fais-je ? Que veux-je ?

- Aujourd'hui, des tâches prédéfinies mais demain, grâce à l'augmentation de mes capacités de perception et de calcul, je veux développer mes capacités d'apprentissage automatique et mes capacités d'adaptation en temps réel……je veux développer mon AUTONOMIE DE DECISION

Ça se complique ?



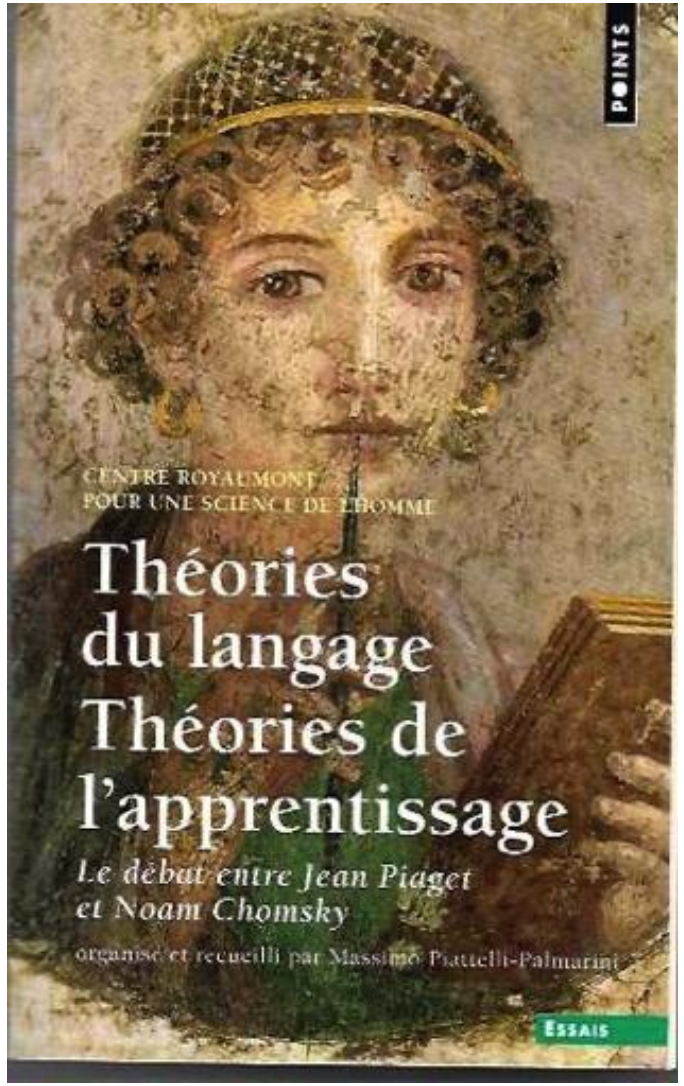


Théories du LANGAGE et Théories de l' APPRENTISSAGE
Rencontre Noam CHOMSKY Jean PIAGET

<u>1975</u> rencontre à l' abbaye de Royaumont autour de ces deux éminents chercheurs.

Ce livre retrace in extenso les exposés et débats exceptionnels de cette rencontre. (Editions du Seuil collection Points 530 pages)

Débats organisés par Le biophysicien Massimo Piatelli-Palmarini avec la participation de scientifiques de grand renom parmi lesquels:
Jacques Monod, François Jacob, Jean-Pierre Changeux, Seymour Papert, René Thom, Maurice Godelier etc.

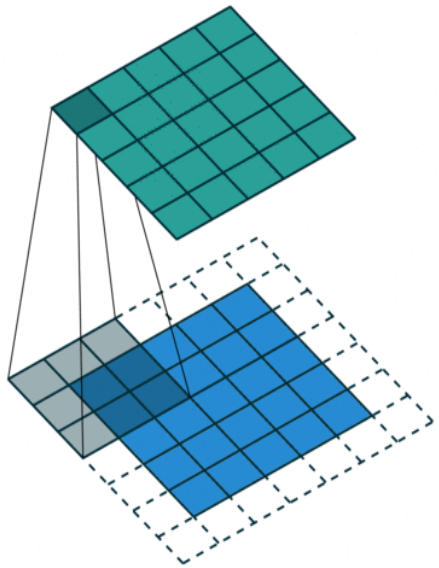


Pour le fun, petit exercice perso sur la « convolution »

- Photo avec votre appareil photo numérique ou votre smartphone

(images matricielles)

Observez le « cadrage » des visages…on a reconnu un ou des objets; un ou des visages au milieu d'un paysage varié……c'est un résultat de « convolution »



Filtrage moyenneur glissant