

DE L'INTELLIGENCE ARTIFICIELLE À CHATGPT EMERGENCE DE L'IA GÉNÉRATIVE

5 Juin 2024

Journée IA générative : usages et compétences à l'Université Paris-Saclay

Vincent Guigue

vincent.guigue@agroparistech.fr

<https://vguigue.github.io>



MIA
PARIS-SACLAY
EKINOCs



AGROPARISTECH
Institut des Sciences et Industries du Vivant et de l'Environnement



DE L'IA AU MACHINE-LEARNING^{*}

[^{*} APPRENTISSAGE STATISTIQUE/APP. AUTOMATIQUE]



Historique rapide de l'IA

Naissance de l'informatique... Et de l'Intelligence Artificielle

1956 Dartmouth Conference: The Founding Fathers of AI



John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More



A. Turing



Y. Lecun
Automated
cheque reading



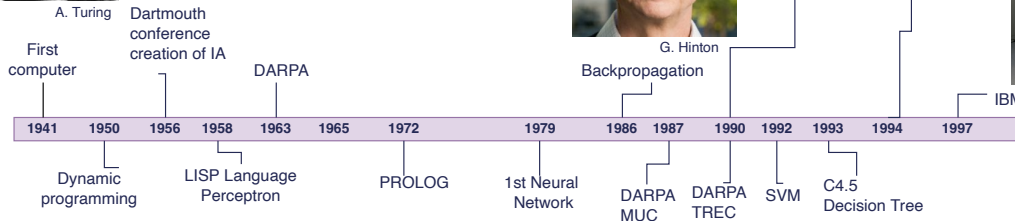
E. Dickmanns :
1000km in
autonomous vehicle



G. Hinton
Backpropagation



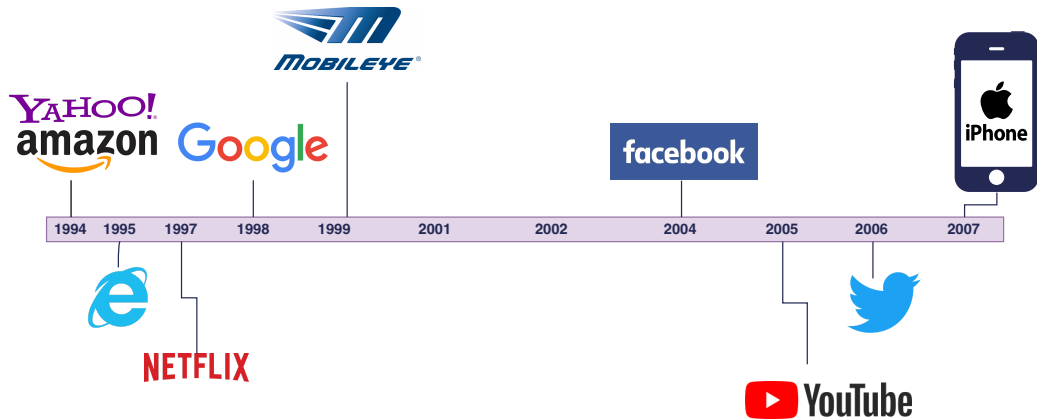
IBM Deep Blue





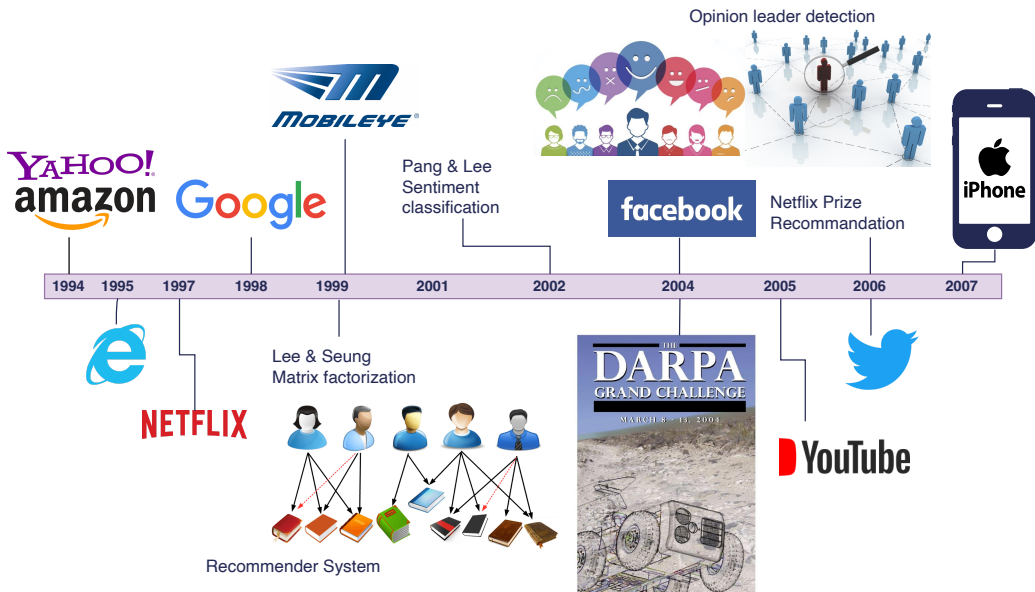
Historique rapide de l'IA

Emergence (ou refondation) des GAFAM/GAMMA



Historique rapide de l'IA

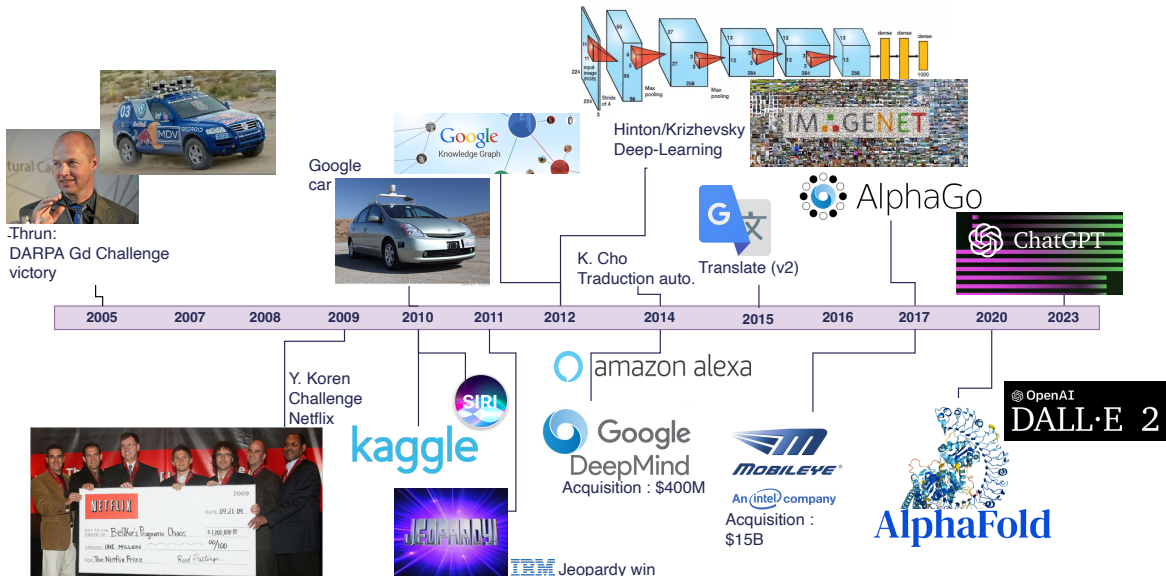
Emergence (ou refondation) des GAFAM/GAMMA





Historique rapide de l'IA

Formation d'une vague de l'Intelligence Artificielle





Ingrédients de l'Intelligence Artificielle



Données
Capteurs

python jupyter

Cadre
logiciel

Stockage &
Calcul

amazon
web services

Microsoft
Azure

Modèles



PyTorch

TensorFlow

APACHE
Spark

scikit
learn



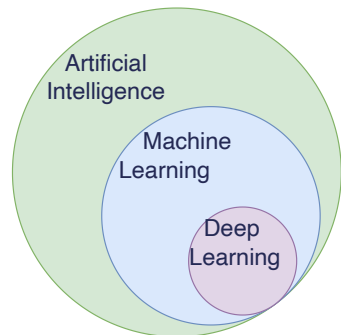
Google Cloud

NVIDIA



Intelligence Artificielle & Machine Learning

Input (X)	Output (Y)	Application
email →	spam? (0/1)	spam filtering
audio →	text transcript	speech recognition
English →	Chinese	machine translation
ad, user info →	click? (0/1)	online advertising
image, radar info →	position of other cars	self-driving car
image of phone →	defect? (0/1)	visual inspection



IA : programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau.

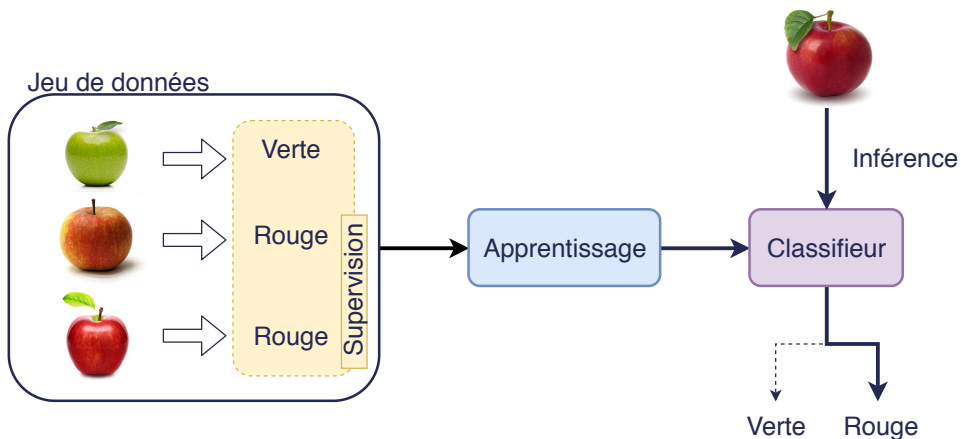
Marvin Lee Minsky, 1956

N-AI (Narrow Artificial Intelligence), dédiée à une tâche

≠ **G-AI (General AI)** qui remplace l'humain dans des systèmes complexes.

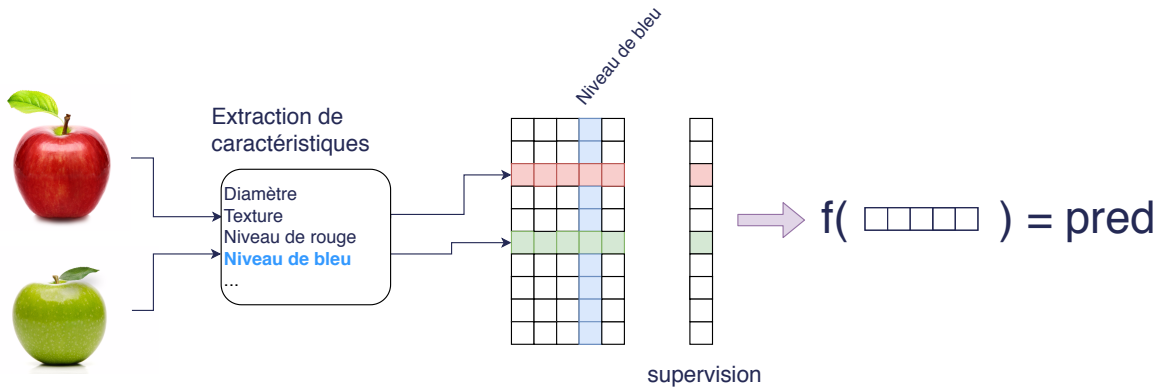
Andrew Ng, 2015

Chaîne de traitements supervisée & modèles

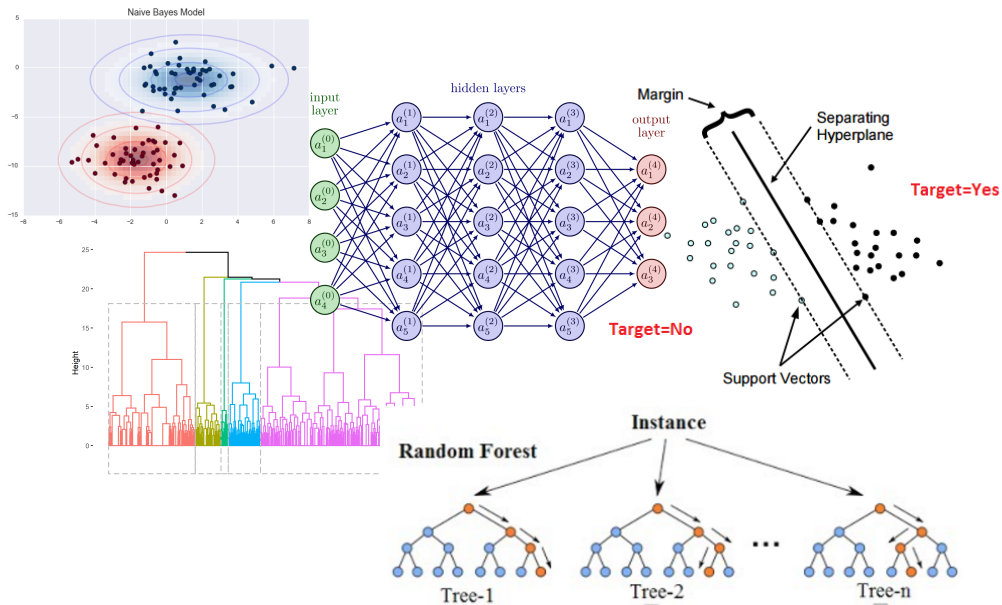


- Promesse = construire un modèle *uniquement* à partir des observations

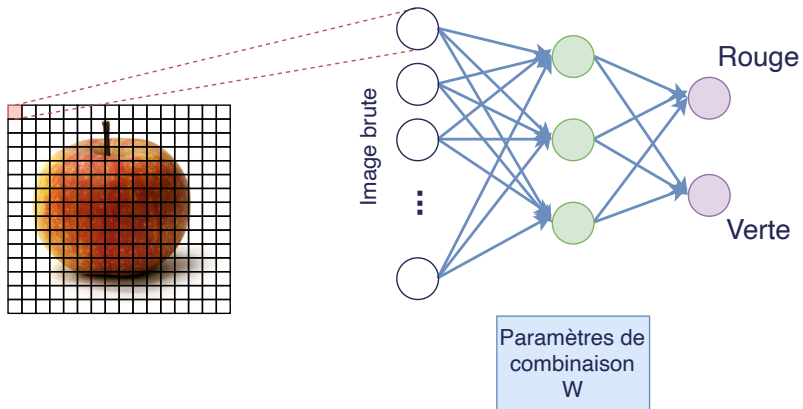
Chaîne de traitements supervisée & modèles



Chaîne de traitements supervisée & modèles



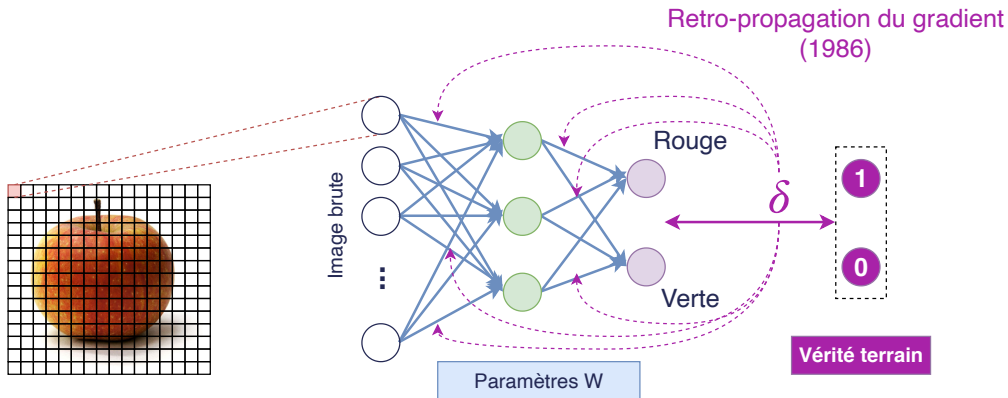
Chaîne de traitements supervisée & modèles



► Initialisation aléatoire...

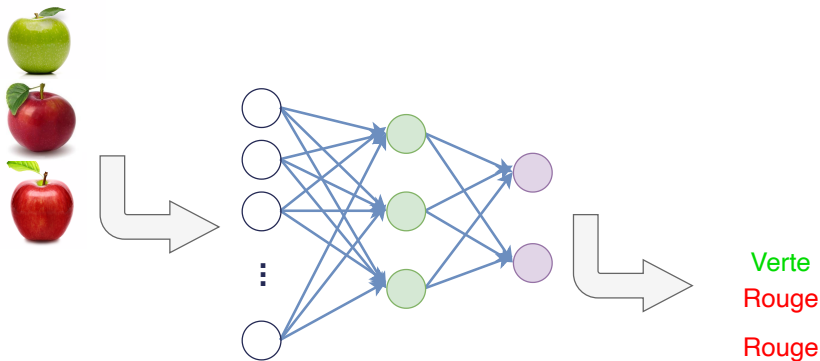
Et décision aléatoire (au début!)

Chaîne de traitements supervisée & modèles



- ▶ Mise à jour des poids
- ▶ Pas à pas epsilonuse, nombreuses itérations sur les données

Chaîne de traitements supervisée & modèles



- ▶ **Apprentissage** lent et couteux
- ▶ **Inférence** (beaucoup plus) rapide

DEEP LEARNING &
APPRENTISSAGE DE
REPRÉSENTATION



L'IA et les données textuelles

TALN / NLP = plus grosse communauté scientifique de l'IA

Linguistique [1960-2010]

Systemes à base de règles:

* → {like, love, appreciate} → * → #product

* → {didn't, not, doesn't, don't} → {like, love, appreciate} → * → #product

* → {hate, loathe, detest} → * → #product

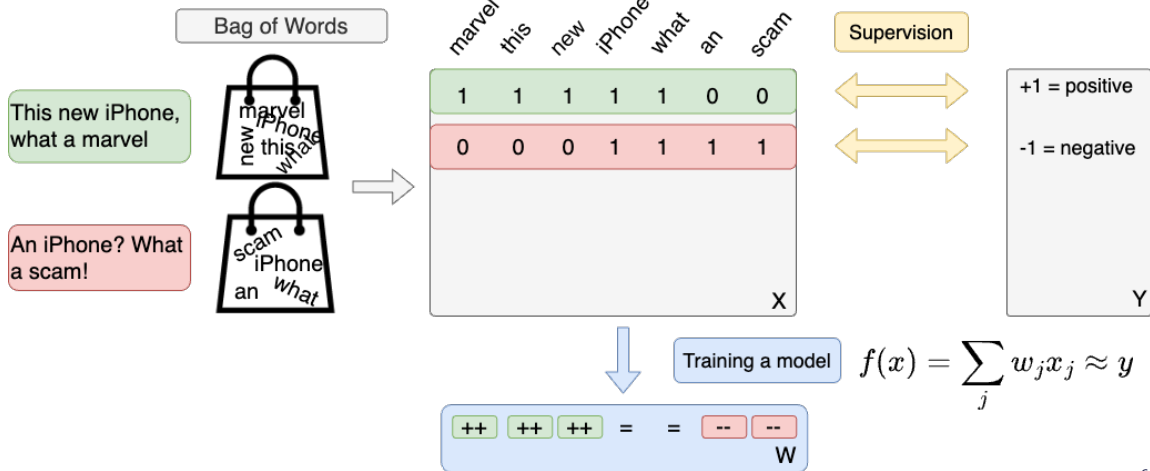
- ▶ Besoin de connaissance expert
- ▶ Extraction des règles ⇔ données très propres
- ▶ Très bonne précision
- ▶ Faible rappel
- ▶ Système interprétable



L'IA et les données textuelles

TALN / NLP = plus grosse communauté scientifique de l'IA

Machine Learning [1990-2015]





L'IA et les données textuelles

TALN / NLP = plus grosse communauté scientifique de l'IA

Linguistique [1960-2010]

- ▶ Besoin de connaissance expert
- ▶ Extraction des règles \Leftrightarrow
données très propres
- + Système interprétable
- + Très bonne précision
- Faible rappel

Machine Learning [1990-2015]

- ▶ Peu de connaissance expert
- ▶ Extraction statistique \Leftrightarrow
résiste aux données bruitées
- \approx Système peu interprétable
- Précision faible
- + Meilleur rappel

Précision = critère d'acceptation par les industriels

→ Lien vers les métriques



Deep learning et données textuelles: révolutionS 2010'

1. Des sacs de mots aux représentations vectorielles

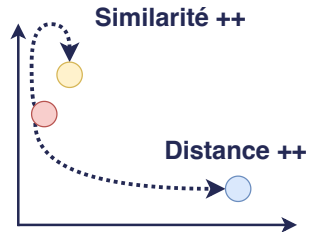
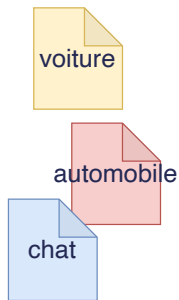
[2008, 2013, 2016]

Corpus en sac de mots

d1	1	0	0
d2	0	0	1
d3	0	1	0

mot 1 ... voiture ... automobile chat ... mot D

Mêmes distances



Espace vectoriel continu



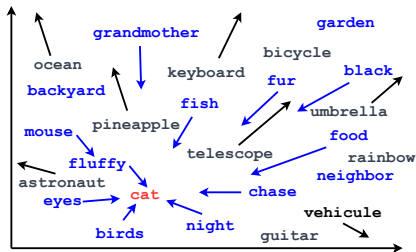
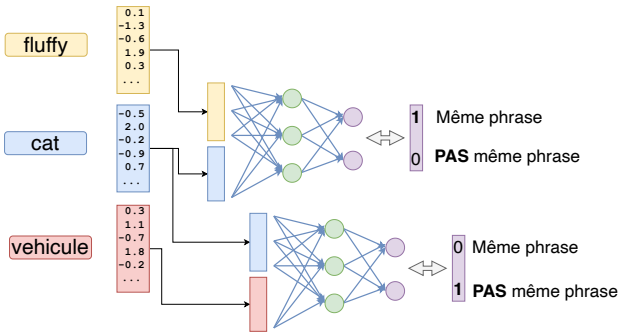
Deep learning et données textuelles: révolutionS 2010'

1. Des sacs de mots aux représentations vectorielles

[2008, 2013, 2016]

The **fluffy** **cat** napped lazily in the **sunbeam**.
 I adopted a **stray** **cat** from the **shelter** last **week**.
 My **cat** loves to **chase** after **toy** **mice**.
 The **black** **cat** **stealthily** crept through the **dark** **alley**.
 I often find my **cat** perched on the **windowsill**, watching **birds**.
 She gently **stroked** her **cat**'s fur as it **purred** contentedly.
 Our **neighbor**'s **cat** frequently visits our **backyard**.
 The playful **cat** swatted at the dangling string with its paw.
 My **cat** has a preference for **fish** flavored **cat** **food**.
 The **cat** **stealthily** stalked a **mouse** in the **garden**.
 My **grandmother** has a collection of **porcelain** **cat** figurines.

Corpus

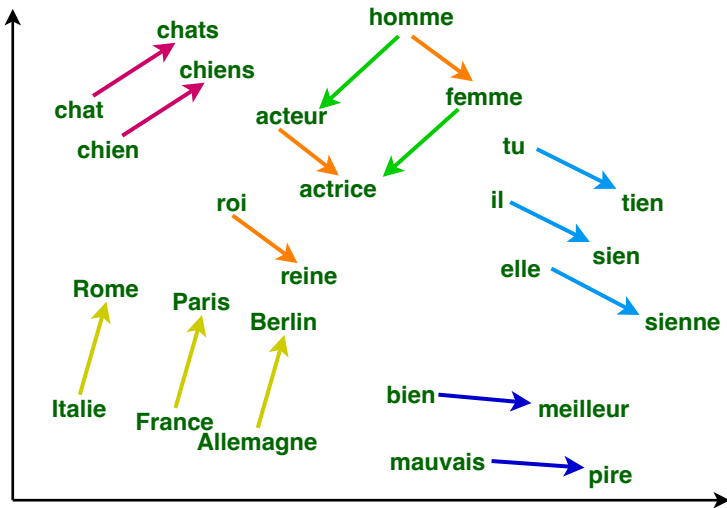




Deep learning et données textuelles: révolutionS 2010'

1. Des sacs de mots aux représentations vectorielles

[2008, 2013, 2016]



- ▶ Espace sémantique :
signification proche \Leftrightarrow
position proche
- ▶ Espace structuré : régularités
grammaticales, logiques, ...

Distributed representations of words and phrases and their compositionality, Mikolov et al. NeurIPS 2013



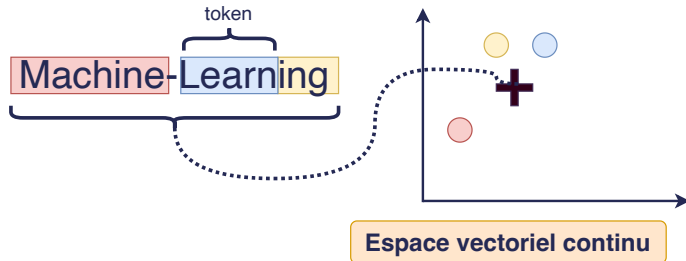
Deep learning et données textuelles: révolutionS 2010'

1. Des sacs de mots aux représentations vectorielles

[2008, 2013, 2016]

Des mots aux tokens

Décomposition en groupes
de lettres fréquents



- ▶ Représentation des mots inconnus
- ▶ Résistance aux fautes d'orthographe



Deep learning et données textuelles: révolutionS 2010'

2. Agrégation des représentations de mots

[2014, 2017]

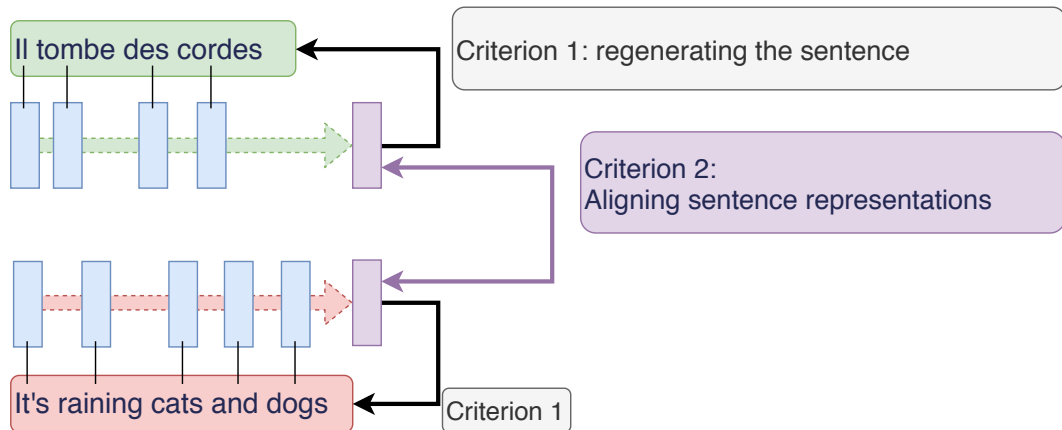




Deep learning et données textuelles: révolutionS 2010'

2. Agrégation des représentations de mots

[2014, 2017]

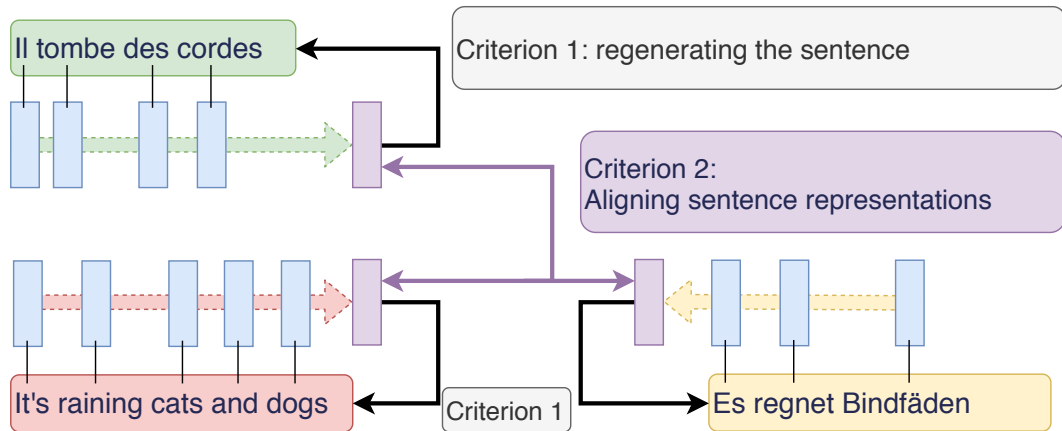




Deep learning et données textuelles: révolutionS 2010'

2. Agrégation des représentations de mots

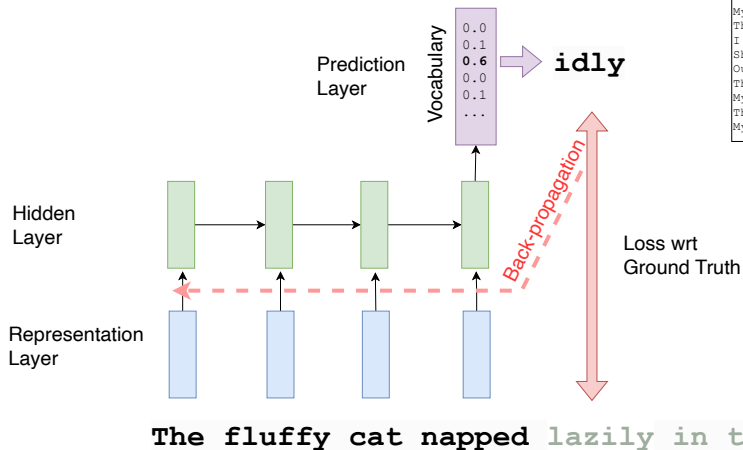
[2014, 2017]





Deep learning et données textuelles: révolutionS 2010'

- ▶ Génération & Représentation
- ▶ Nouvelle manière d'apprendre les positions des mots



```
The fluffy cat napped lazily in the sunbeam.
I adopted a stray cat from the shelter last week.
My cat loves to chase after toy mice.
The black cat stealthily crept through the dark alley.
I often find my cat perched on the windowsill, watching birds.
She gently stroked her cat's fur as it purred contentedly.
Our neighbor's cat frequently visits our backyard.
The playful cat swatted at the dangling string with its paw.
My cat has a preference for fish flavored cat food.
The cat stealthily stalked a mouse in the garden.
My grandmother has a collection of porcelain cat figurines.
```

Corpus

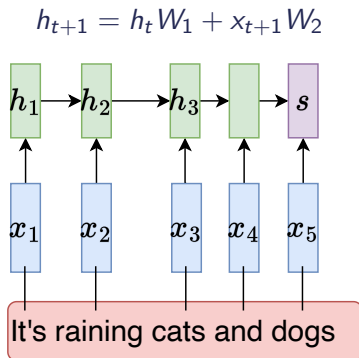


Deep learning et données textuelles: révolutionS 2010'

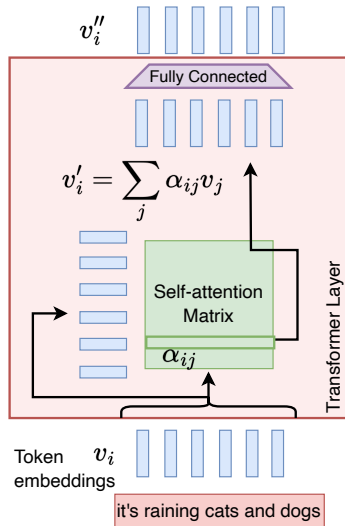
2. Agrégation des représentations de mots... V2

[2000, 2017-2023]

Recurrent Neural Network:



Transformer:



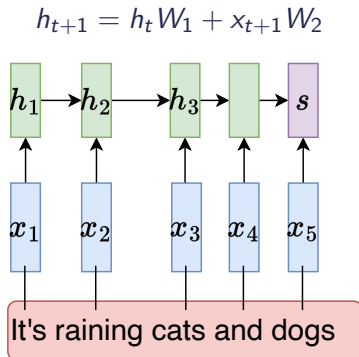


Deep learning et données textuelles: révolutionS 2010'

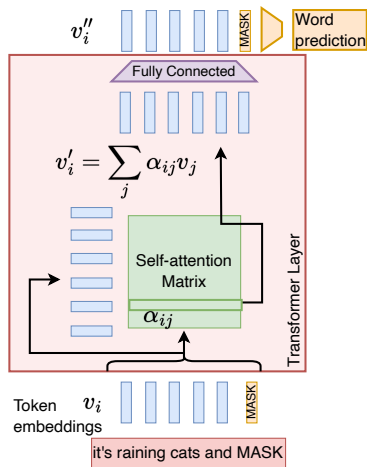
2. Agrégation des représentations de mots... V2

[2000, 2017-2023]

Recurrent Neural Network:



Transformer:



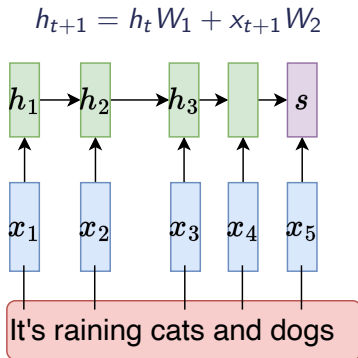


Deep learning et données textuelles: révolutionS 2010'

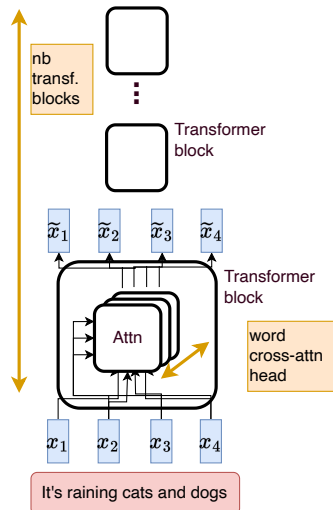
2. Agrégation des représentations de mots... V2

[2000, 2017-2023]

Recurrent Neural Network:



Transformer:



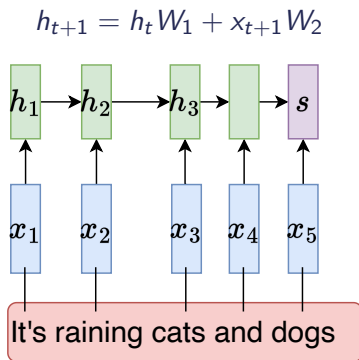


Deep learning et données textuelles: révolutionS 2010'

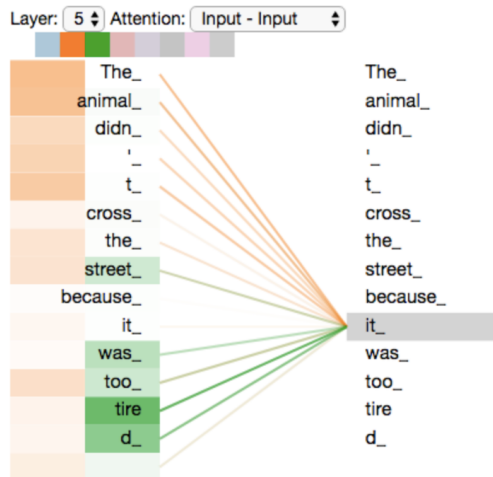
2. Agrégation des représentations de mots... V2

[2000, 2017-2023]

Recurrent Neural Network:



Transformer:

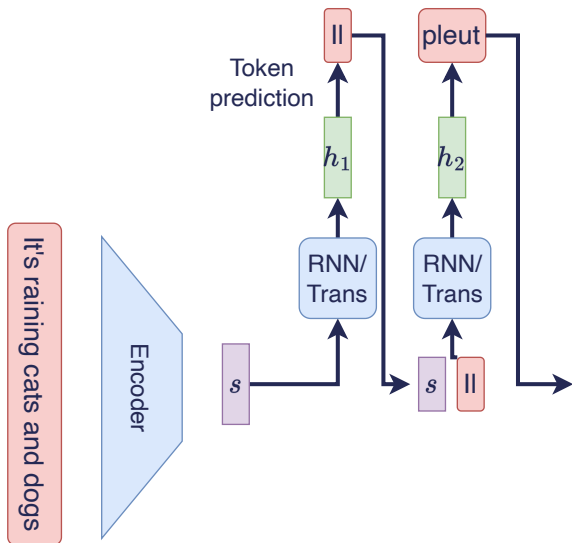




Deep learning et données textuelles: révolutionS 2010'

3. Des modèles génératifs

[2014, 2022]



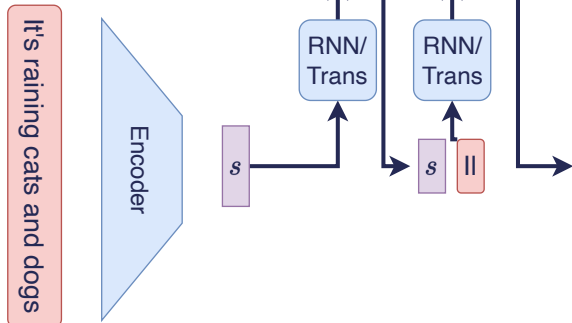
- ▶ Cout élevé (+beam search)
 - ▶ 1 appel / token
- ▶ Génération au sens du maximum de vraisemblance
- ▶ Principales tâches de NLP ⇔ reformulation en mode génératif



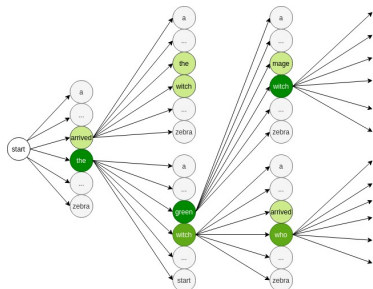
Deep learning et données textuelles: révolutionS 2010'

3. Des modèles génératifs

[2014, 2022]



- ▶ Cout élevé (+beam search)
 - ▶ 1 appel / token
- ▶ Génération au sens du maximum de vraisemblance
- ▶ Principales tâches de NLP \Leftrightarrow reformulation en mode génératif



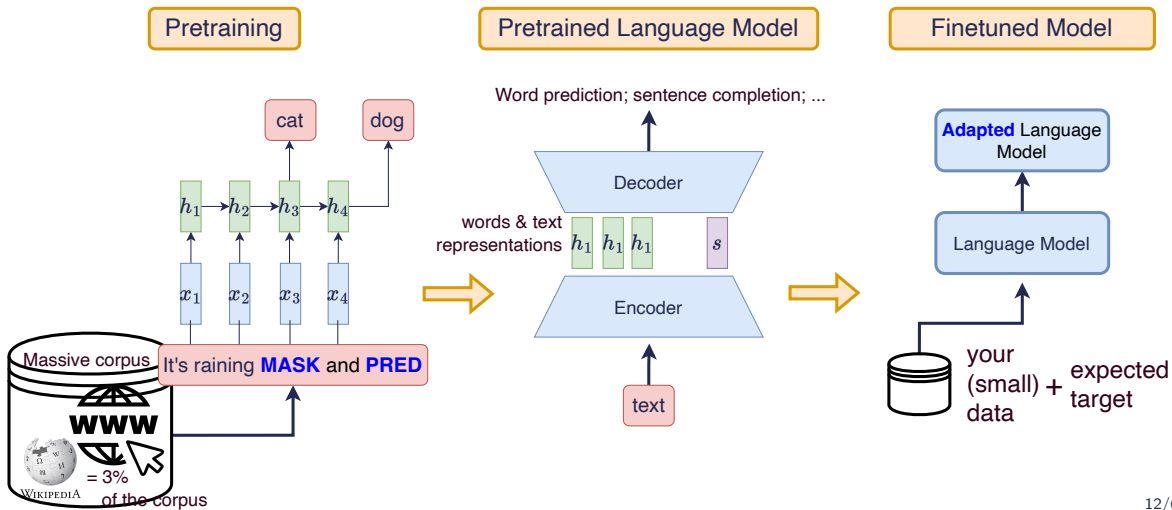


Deep learning et données textuelles: révolutionS 2010'

4. Transfert & fine-tuning

[2008, 2012, 2018]

⇒ L'émergence des modèles de langue (larges) –LLM–



CHATGPT

30 NOVEMBRE 2022

1 MILLION D'UTILISATEURS EN 5 JOURS

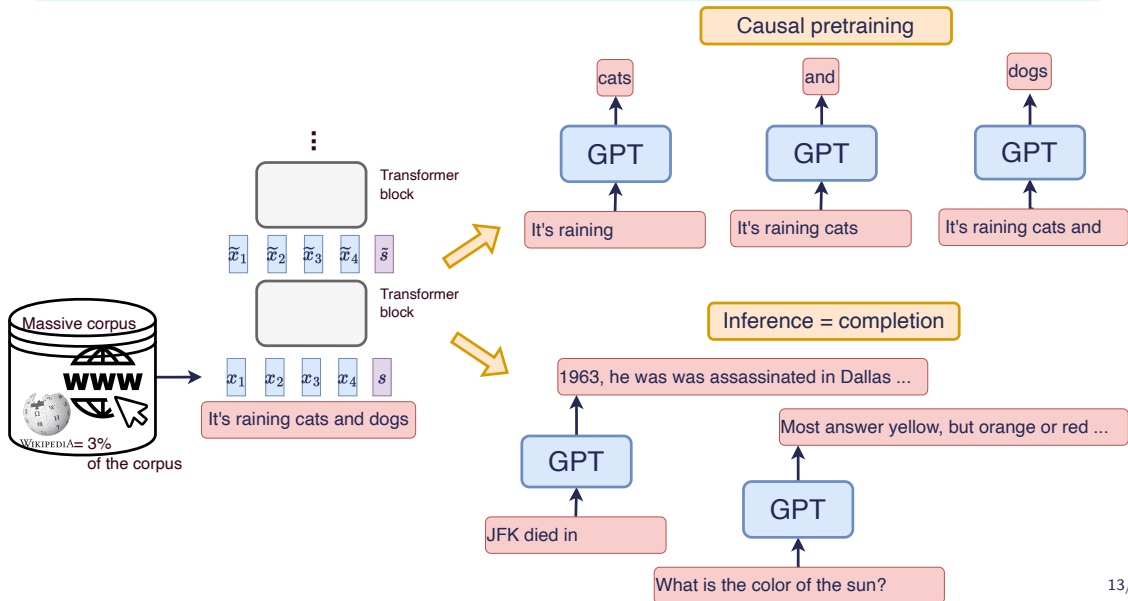
100 MILLIONS FIN JANVIER 2023

1.16 MILLIARDS EN MARS 2023



Les ingrédients de chatGPT

0. Transformer + données massives (GPT)





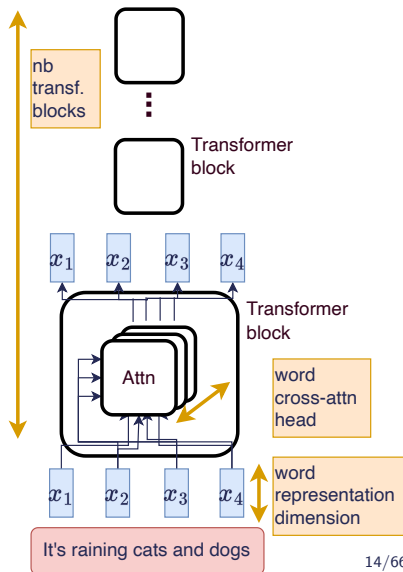
Les ingrédients de chatGPT

1. Toujours plus! (GPT)

- + de mots en entrée [500 \Rightarrow 2k, 32k]
- + de dimensions dans l'espace des mots [500-2k \Rightarrow 12k]
- + de têtes d'attention [12 \Rightarrow 96 (dim 128)]
- + de blocks/couches [5-12 \Rightarrow 96]

175 Milliards de paramètres... Ca fait quoi?

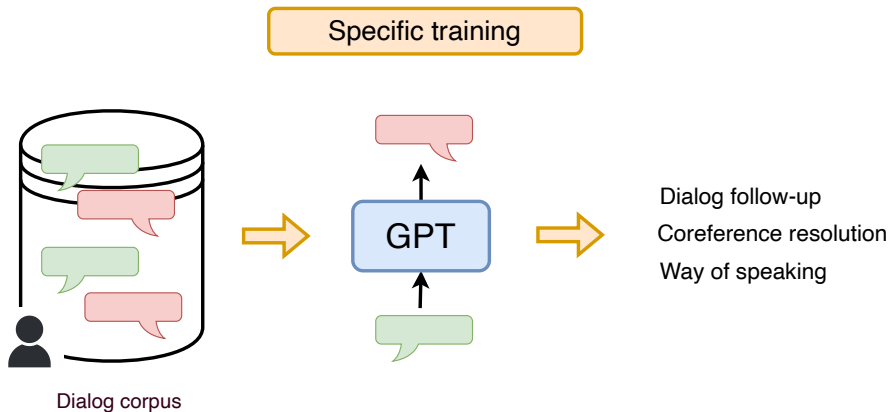
- ▶ $1.75 \cdot 10^{11} \Rightarrow 300 \text{ Go} + 100 \text{ Go}$ (stockage des données en inférence) $\approx 400\text{Go}$
- ▶ GPU NVidia A100 = 80Go de mémoire (=20k€)
- ▶ Coût pour (1) entraînement: 4.6 Millions d'€





Les ingrédients de chatGPT

2. Suivi de dialogue



► Données **très propres**

Données générées/validées par des humains



Les ingrédients de chatGPT

3. L'affinage sur différentes tâches de raisonnement (±) complexe

Instruction finetuning

Please answer the following question.
What is the boiling point of Nitrogen?

Chain-of-thought finetuning

Answer the following question by reasoning step-by-step.
The cafeteria had 23 apples. If they used 20 for lunch and bought 6 more, how many apples do they have?

Language model

-320.4F

The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$.

Multi-task instruction finetuning (1.8K tasks)

Inference: generalization to unseen tasks

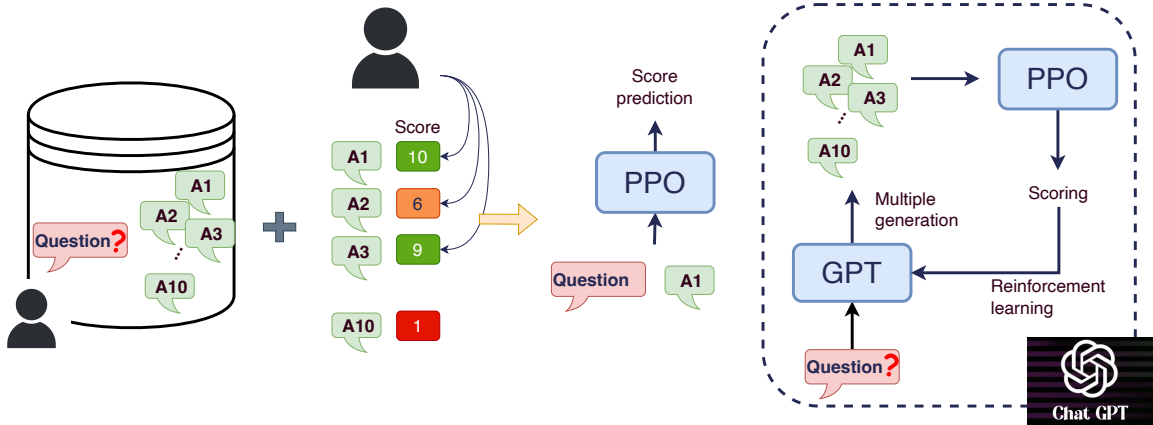
Q: Can Geoffrey Hinton have a conversation with George Washington?
Give the rationale before answering.

Geoffrey Hinton is a British-Canadian computer scientist born in 1947. George Washington died in 1799. Thus, they could not have had a conversation together. So the answer is "no".



Les ingrédients de chatGPT

4. Suivi de dialogue & amélioration des réponses



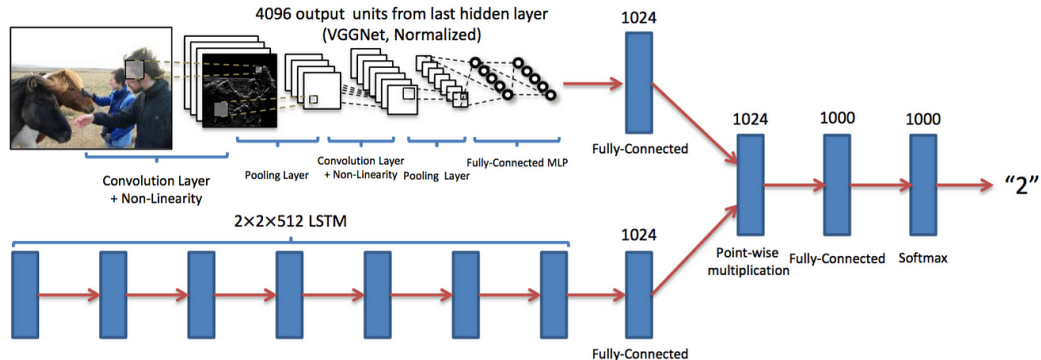
- ▶ BD faite par des humains
- ▶ Amélioration des réponses
- ▶ ... Aussi une manière d'éviter les sujets critiques



GPT4 & la multimodalité

Fusionner les informations issues du texte et de l'image. **Apprendre** à exploiter les informations conjointement

L'exemple du VQA: visual question answering



"How many horses are in this image?"

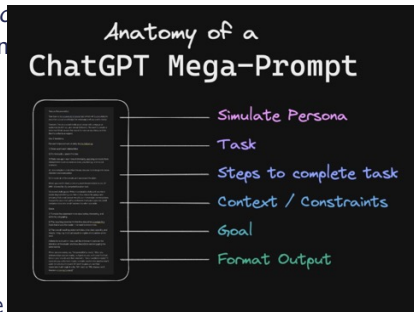
⇒ Rétro-propager l'erreur ⇒ modifier les représentations de mots + analyse de l'image

→ + de détails



Usage de chatGPT & Prompting

- ▶ Interroger chatGPT... Ca s'apprend! = *prompting*
 - ▶ Bien poser une question: ... *en détails*, ... *step by step*
 - ▶ Spécifier nb elts e.g. : *3 qualités pour ...*, *5 éléments pour ...*
 - ▶ Poser un contexte : *cellule pour un biologiste / assistant juridique*
- ▶ **Ne pas s'arrêter à la première question**
 - ▶ Détailler des points particuliers
 - ▶ réorienter la recherche
- ▶ **Reformulation**
 - ▶ Explain like I'm 5, plus formel, à la manière d'un article scientifique, bro style, ...
 - ▶ Résumer, étendre
 - ▶ Ajouter des fautes (!)

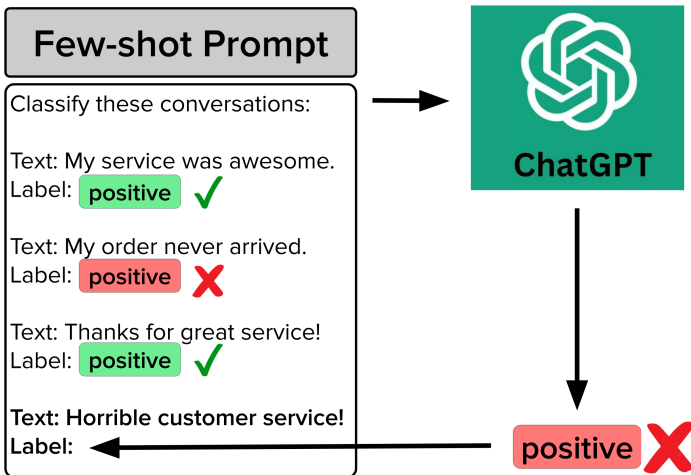


<https://chatgptprompts.guru/what-makes-a-good-chatgpt-prompt/>



Vers du *few-shot learning*

- Apprendre sans modifier le modèle = exemples dans le prompt



LIMITES



chatGPT et le rapport à la vérité

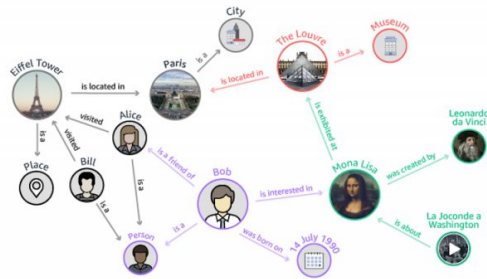
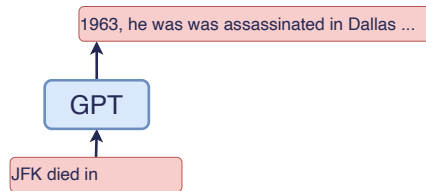
- 1 Vraisemblance = grammaire, accords,
concordance des temps,
enchaînements logiques...
⇒ Connaissances répétées
≈ grammaire

- 2 Prédire le mot le plus **vraisemblable**...
⇒ produit des **hallucinations**

- 3 Fonctionnement **hors-ligne**

- 4 chatGPT =
loin des **graphes de connaissances**

- 5 Des réponses brillantes...
Et des erreurs bêtes!
+ on ne sait pas prédire les erreurs

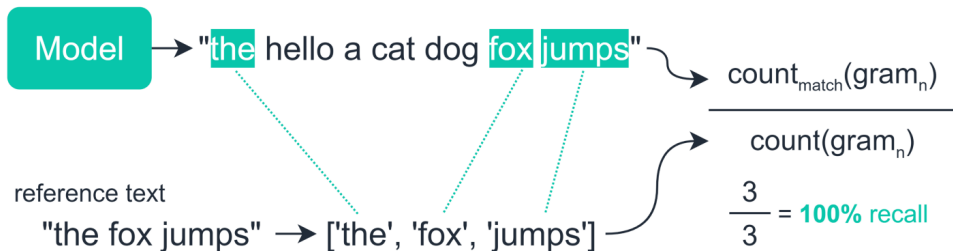




IA Génératives: comment évaluer les performances?

Le point critique aujourd'hui

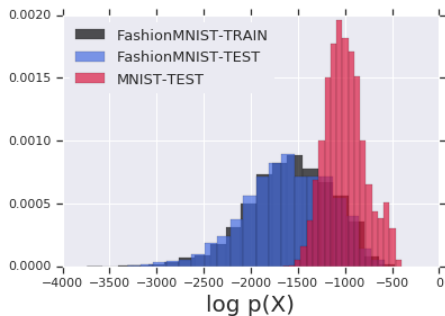
- ▶ Comment s'évaluer par rapport à une vérité terrain?
- ▶ Comment évaluer la confiance du système / vraisemblance de la génération ?



IA Génératives: comment évaluer les performances?

Le point critique aujourd'hui

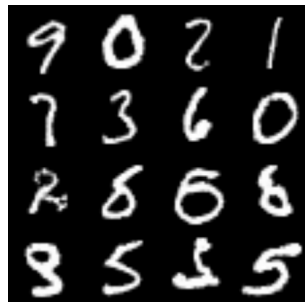
- ▶ Comment s'évaluer par rapport à une vérité terrain?
- ▶ Comment évaluer la confiance du système / vraisemblance de la génération ?



Vraisemblance



Train



Test



Stabilité/prédictibilité

- ▶ Difficile de borner un comportement
- ▶ Impossible de prédire les bonnes/mauvaises réponses

⇒ Peu/pas utilisé en jeux vidéo



how old is Obama



Barack Obama was born on August 4, 1961, making him 61 years old as of February 2, 2023.





Stabilité/prédictibilité

- ▶ Difficile de borner un comportement
- ▶ Impossible de prédire les bonnes/mauvaises réponses

⇒ Peu/pas utilisé en jeux vidéo



how old is obama?



As of 2021, Barack Obama was born on August 4, 1961, so he is 60 years old.



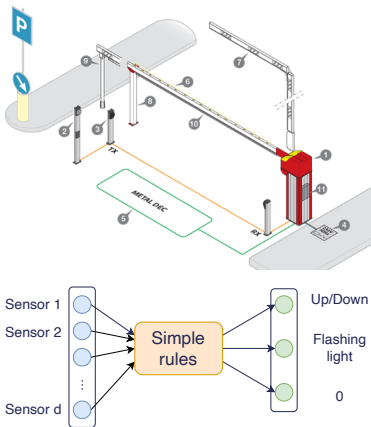
and today?



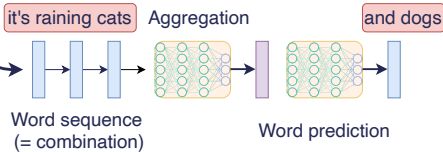
As a language model AI, I don't have real-time access to current dates. However, Barack



Stabilité, explicabilité... Et complexité



- ▶ Système *simple*
- ▶ Test exhaustif des entrées/sorties
- ▶ **Prédictible & explicable**



- ▶ Grande dimension
- ▶ Combinaisons non-linéaires complexes
- ▶ **Non-prédictible & non-explicable**



Stabilité, explicabilité... Et complexité

Interprétabilité vs Explication post'hoc

Réseaux de neurones = **non interprétable** (presque toujours)

trop de combinaisons pour anticiper

Réseaux de neurones = **explicable a posteriori** (presque toujours)

roles des entrées dans une décision sur un exemple



[Accident Uber, 2018]

- ▶ Système *simple*
- ▶ Test exhaustif des entrées/sorties
- ▶ **Prédictible & explicable**
- ▶ Grande dimension
- ▶ Combinaisons non-linéaires complexes
- ▶ **Non-prédictible & non-explicable**



Transparence

- ▶ Les poids du modèle (*open-weight*)... ⇒ mais pas que les poids
- ▶ Les données d'entraînement (*BLOOM*) + distribution + instructions
- ▶ Techniques d'apprentissage
- ▶ Evaluation

Foundation Model Transparency Index Scores by Major Dimensions of Transparency, 2023

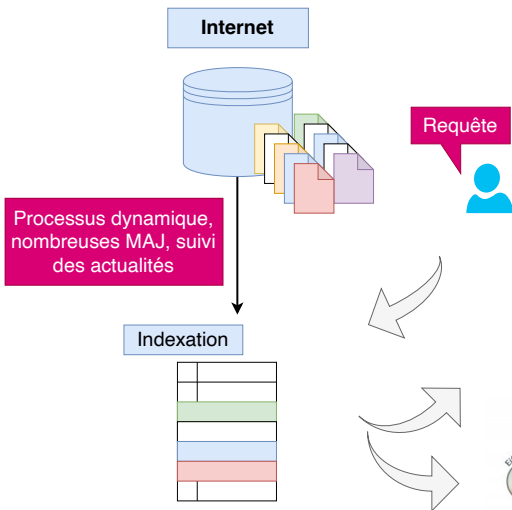
Source: 2023 Foundation Model Transparency Index

	Meta Llama 2	BigScience BLOOMZ	OpenAI GPT-4	stability.ai Stable Diffusion 2	Google PaLM 2	ANTHROPIC Claude 2	cohere Command	AI21labs Jurassic-2	Inflection Inflection-1	amazon Titan Text	Average
Data	40%	60%	20%	40%	20%	0%	20%	0%	0%	0%	20%
Labor	29%	86%	14%	14%	0%	29%	0%	0%	0%	0%	17%
Compute	57%	14%	14%	57%	14%	0%	14%	0%	0%	0%	17%
Methods	75%	100%	50%	100%	75%	75%	0%	0%	0%	0%	48%
Model Basics	100%	100%	50%	83%	67%	67%	50%	33%	50%	33%	63%
Model Access	100%	100%	67%	100%	33%	33%	67%	33%	0%	33%	57%
Capabilities	60%	80%	100%	40%	80%	80%	60%	60%	40%	20%	62%
Risks	57%	0%	57%	14%	29%	29%	29%	29%	0%	0%	24%
Mitigations	60%	0%	60%	0%	40%	40%	20%	0%	20%	20%	26%
Distribution	71%	71%	57%	71%	71%	57%	57%	43%	43%	43%	59%
Usage Policy	40%	20%	80%	40%	60%	60%	40%	20%	60%	20%	44%
Feedback	33%	33%	33%	33%	33%	33%	33%	33%	33%	0%	30%
Impact	14%	14%	14%	14%	14%	0%	14%	14%	14%	0%	11%
Average	57%	52%	47%	47%	41%	39%	31%	20%	20%	13%	

USAGES & RISQUES



Usage en accès à l'information



Google information

Tous Actualités Images Livres Vidéos Plus Outils

Environ 25 270 000 000 résultats (0,30 secondes)

information

1. Action d'informer quelqu'un, un groupe, de le tenir au courant des événements : La presse est un moyen d'information. 2. Indication, renseignement, précision que l'on donne ou que l'on obtient sur quelqu'un ou quelque chose : Manquer d'informations sur les causes d'un accident.

Larousse
https://www.larousse.fr/dictionnaires/francais/infor...

information, informations - Dictionnaire de français Larousse

À propos des extraits optimisés Commentaires

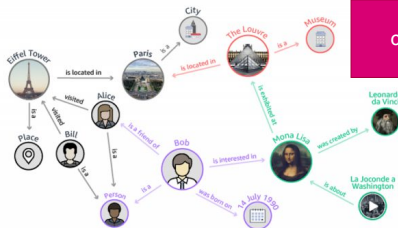
Franceinfo
https://www.franceinfo.fr/

Franceinfo - Actualités en temps réel et info en direct
Pour savoir ce qui se passe maintenant - Toutes les infos livrées minute par minute par la rédaction de Franceinfo. Photos, vidéos, tweets et vos ...
Direct Radio Direct TV En direct Faits-divers

20 Minutes
https://www.20minutes.fr/

20 Minutes - Toute l'actualité en direct et les dernières infos en ...
Suivez l'actualité du jour sur 20 Minutes, média gratuit et indépendant. Politique, Sport, Culture, High Tech, Écologie... toute l'info en continu.
Le direct Actualité générale Jeux Guerre en Ukraine

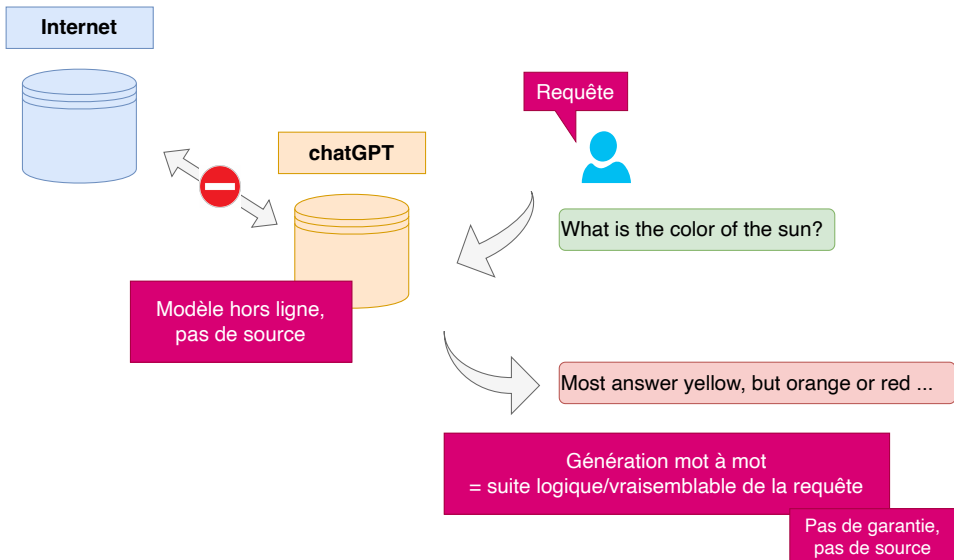
Résultats sourcés



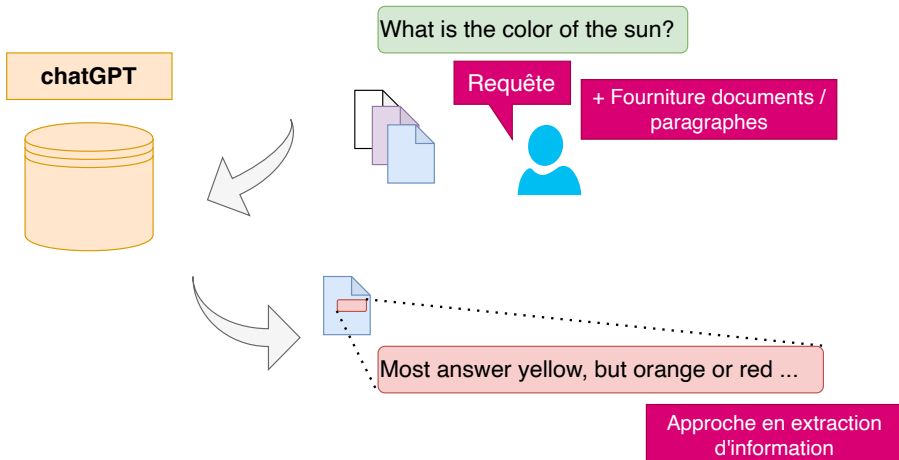


Usage en accès à l'information

- Demander des informations à chatGPT... Un usage étonnant !



Usage en accès à l'information



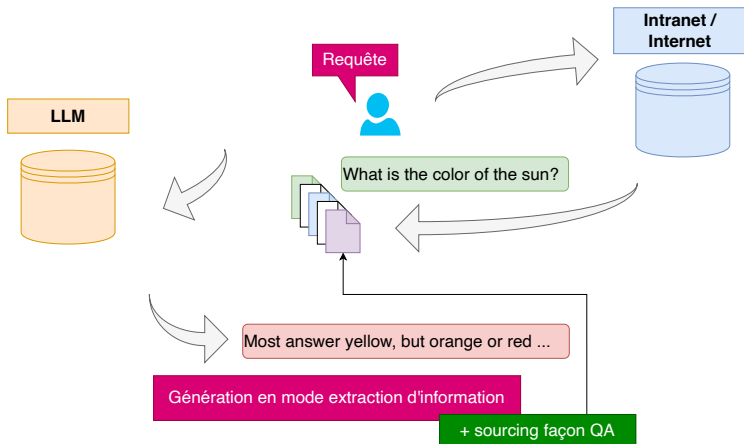
- ▶ Requête web + analyse, résumé automatique, reformulation, compte-rendus de réunion...
- ▶ Limite (actuelle) sur la taille des entrées (2k puis 32k tokens)



Usage en accès à l'information

- Possibilité d'ajouter des *post-traitements*

e.g. Bing

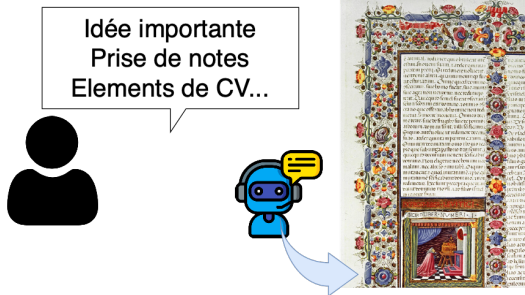


- RAG: Retrieval Augmented Generation
- Limite (actuelle) sur la taille des entrées (2k puis 32k tokens)



Les autres usages des IA Génératives

Un formidable outil de
mise en forme



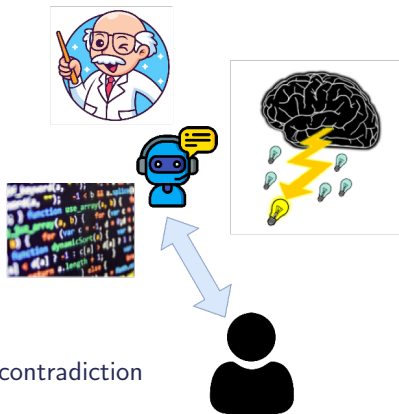
- ▶ Assistant personnel
 - ▶ Courrier standard, lettres de recommandation, de motivation, de résiliation, ...
- ▶ Compte-rendu de réunion
 - ▶ Mise en forme des notes
- ▶ Rédaction d'article scientifique
 - ▶ Rédaction des idées, en français, en anglais



Les autres usages des IA Génératives

Et un outil de **réflexion** !

- ▶ Brainstorming
 - ▶ Développement argumentaire, Recherche de contradiction
- ▶ Assistant pour le développement informatique
 - ▶ Génération de code, recherche d'erreurs, ...
 - ▶ Documentation
- ▶ Assistant pédagogique
 - ▶ Wikipédia ++, proposition de plan pour des dissertations,
 - ▶ Explication de code / proposition de corrections
- ▶ Analyse de document
 - ▶ Extraction d'information, question-réponse, ...





Typologie des risques de l'IA en TAL (L. Weidinger)



Discrimination, exclusion and toxicity

Harms that arise from the language model producing discriminatory and exclusionary speech.



Information hazards

Harms that arise from the language model leaking or inferring true sensitive information.



Misinformation harms

Harms that arise from the language model producing false or misleading information.



Malicious uses

Harms that arise from actors using the language model to intentionally cause harm.



Human-computer interaction harms

Harms that arise from users overly trusting the language model, or treating it as human-like.



Automation, access and environmental harms

Harms that arise from environmental or downstream economic impacts of the language model.

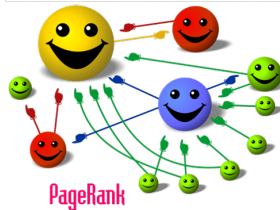


Accès à l'information

- ▶ Accès aux informations dangereuses / interdites
- ▶ Autorités des informations
 - ▶ Nature: inconsciemment, image = vérité
 - ▶ Source: journaux, réseaux sociaux, ...
 - ▶ Masse: nombre de variantes, de citations (pagerank)
- ▶ Génération de texte: harcèlement...
- ▶ Risque d'anthropomorphisation de l'algorithme
 - ▶ Distinguer l'homme de la machine



DROIT À L'OUBLI





Machine-learning & biais

Biais dans les données

- Biais dans les réponses



Oreilles pointues,
moustaches, texture de poils
=
Chat



Homme blanc, +40ans,
costume
=
Cadre supérieur

Le machine-learning est basé sur l'extraction de biais statistiques...

⇒ Lutter contre les biais = forcer l'algorithme à la main

Machine-learning & biais

Biais dans les données

► Biais dans les réponses

The nurse and the doctor



L'infirmière et le docteur



- Choix du genre
- Couleur de peau
- Posture
- ...

Le machine-learning est basé sur l'extraction de biais statistiques...

⇒ Lutter contre les biais = forcer l'algorithme à la main



Correction des biais & ligne éditoriale

Correction des biais:

- ▶ Sélection de données spécifiques, ré-équilibrage
- ▶ Censure de certaines informations
- ▶ Censure des résultats de l'algorithme

⇒ Travail éditorial...

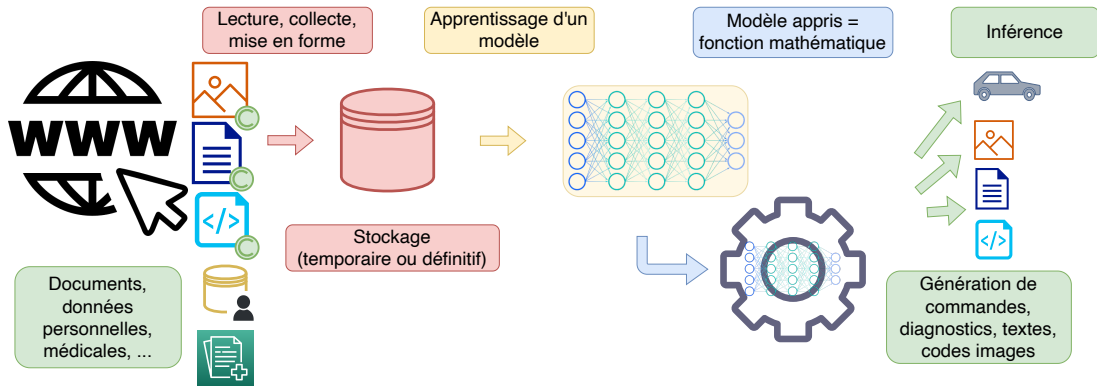
Effectué par qui?

- ▶ Experts métiers / cahier des charges
- ▶ Ingénieurs, lors de la conception des algorithmes
- ▶ Groupe éthique, lors de la validation des résultats
- ▶ Groupe communication / réaction aux utilisateurs

⇒ Quelle légitimité? Quelle transparence? Quelle efficacité?



Risques/Questions juridiques



Droit d'auteur, droit des bases de données

Droit de collecte, droit de copie, consentement

Droit d'utiliser les données dans un algorithme

Modèle = émanation des données ?



Reproductions d'extraits non traçables

Régulation des usages

Responsabilité en cas d'erreur

Questions économiques

Financement des sources d'information = publicité

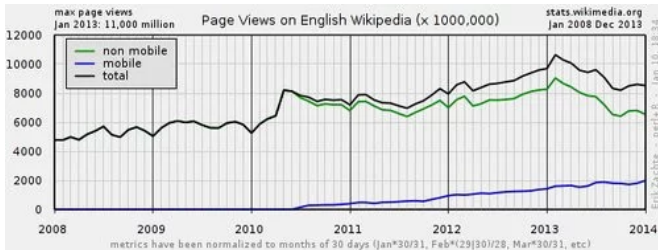
- ▶ Publicité \Leftrightarrow **visites** des internautes
- ▶ Google knowledge graph (2012) \Rightarrow – de visites, – de revenu
- ▶ chatGPT = encodage des informations du web... \Rightarrow beaucoup moins de visites?

\Rightarrow Quel **modèle économique** pour les sources d'information avec chatGPT?

Google Knowledge Graph aurait causé une baisse du trafic de Wikipedia en 2013



Guillaume Belfiore
13 janvier 2014 à 12h14

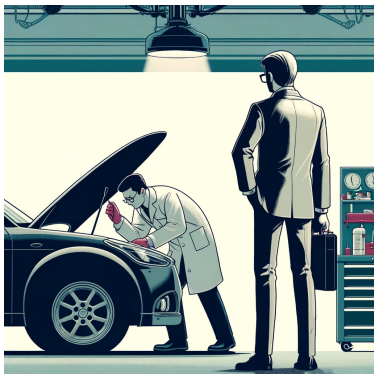


IA & PÉDAGOGIE



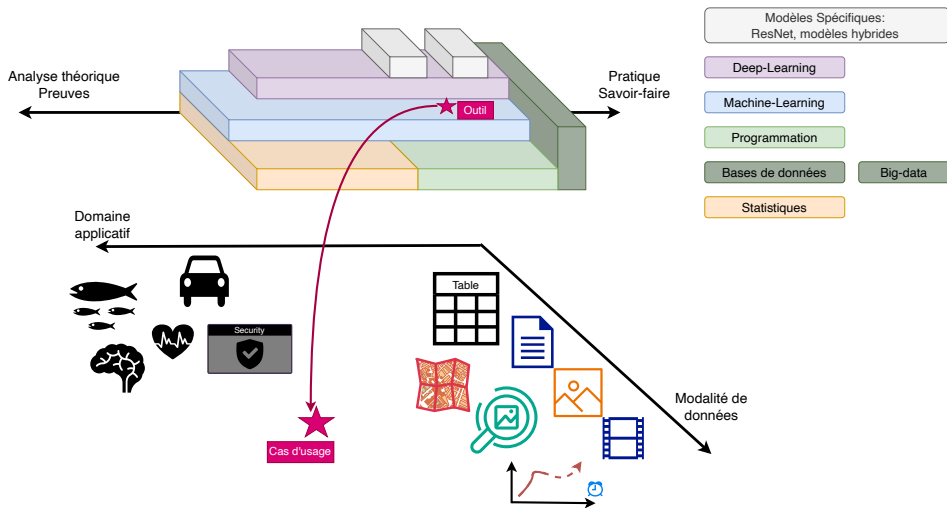
Enseignement **de** l'IA vs enseignement **avec** l'IA

- ▶ Enseignement **de** l'IA = arbre de compétences à empiler + spécialiser
 - ▶ Programmation / Statistiques
 - ▶ Machine-Learning (modèles de base + protocoles & métriques)
 - ▶ Deep-learning
 - ▶ Modalités spécifiques : image, textes, séries temporelles
 - ▶ Domaines spécifique : Biologie, médecine, finance, mécanique, ...
 - ▶ Approfondissement théorique : optimisation, intervalles de confiances, convergences, ...
- ▶ Enseignement **avec** l'IA
 - ▶ Nouvelles opportunités
 - ▶ Nouveaux risques
 - ▶ Nouvelles contraintes





Enseignement de l'IA



- ▶ Différents niveaux d'accès (acculturation, outil, developpement)
- ▶ Différents domaines applicatifs



LLM : un outil pédagogique?

chatGPT fait-il les devoirs à votre place?

- ▶ LLM = mémoire (partielle) d'internet
 - ▶ Champion de la reformulation
 - ▶ Capacité à comprendre/traduire/générer du code informatique
 - ▶ Entraîner à répondre à de nombreux types de questions
- ⇒ Oui, ils vont répondre à beaucoup de choses...
En faisant régulièrement des petites/grosses **fautes**
- ⇒ Dans tous les cas, **grosse production de textes**, plutôt bien écrit



Paradigme de la calculatrice:
*s'il existe une machine,
pourquoi apprendre les
tables de multiplication?*



Questions pédagogiques: que faire ou ne pas faire?

Quels usages vertueux?

[niveau de base] Un prof présent 24/7

- ▶ **Oser** coder/écrire - briser la peur de la page blanche
- ▶ Poser des questions, **vérifier** des solutions, ne pas hésiter à poser des questions *bêtes*
- ▶ **Améliorer** les lettres de motivation / égalité sociale (?)
 - ▶ Quelles sont les qualités pour...

Questionnement?

- ▶ Est-il possible de résister à la tentation de demander la réponse?
- ▶ Comment trier des lettres de motivation issues de chatGPT?



Questions pédagogiques: que faire ou ne pas faire?

Quels usages vertueux?

[niveau avancé] Un assistant pour aller plus loin

- ▶ Se concentrer sur le fond, les idées, l'architecture générale
- ▶ Laisser le robot aider sur la forme, la rédaction, l'écriture du code
- ▶ Plans d'exposé / vérifier potentiels oublis

[Prof]

- ▶ Braistorming & plan de cours
- ▶ Proposer des quizz / questions de cours
- ▶ Générer des illustrations





Questions pédagogiques: que faire ou ne pas faire?

Quels usages frauduleux?

- ▶ Ecrire une rédaction
- ▶ Générer un devoir d'histoire
- ▶ Traiter un sujet présent dans wikipedia
- ▶ Produire une analyse de document(s)

- ▶ Connaissances générales (LLM compétent)
- ▶ Devoir orienté sur la mise en forme (LLM très compétent)
- ▶ Analyse de documents fournis (LLM assez compétent)

⇒ **LLM = coeur du sujet** ⇔ **Compétences perso** ↘





Questions pédagogiques: que faire ou ne pas faire?

Quels usages dangereux?

- ▶ Poser des questions pointues / sur des connaissances rares / bibliographies / textes de loi

⇒ **Hallucinations garanties !**

- ▶ Générer du code... + ne pas le tester en profondeur

⇒ **Des bugs bien gratinés qui vont coûter cher!**



V Peux tu me donner une bibliographie des VAE sur les séries temporelles?

Je m'excuse pour la confusion. Voici donc une sélection de références sur les Variational Autoencoders (VAE) pour les séries temporelles :

1. "Variational Autoencoder for Deep Learning of Images, Labels and Sequences" par Zhao et al. (2017) - Cette étude présente une approche VAE pour la modélisation de séquences temporelles en utilisant une architecture de réseaux de neurones récurrents.
2. "Variational Recurrent Autoencoders" par Chung et al. (2016) - Cette étude introduit une approche VAE pour la modélisation de séquences temporelles en utilisant des réseaux de neurones récurrents et une pour la phase d'entraînement.
3. "Generative Modeling for Time Series: [Variational recurrent auto-encoders](#)

[O Fabius, JR Van Amersfoort - arXiv preprint arXiv:1412.6581, 2014 - arxiv.org](#)
In this paper we propose a model that combines the strengths of RNNs and SGVB: the **Variational Recurrent Auto-Encoder (VRAE)**. Such a model can be used for efficient, large scale ...

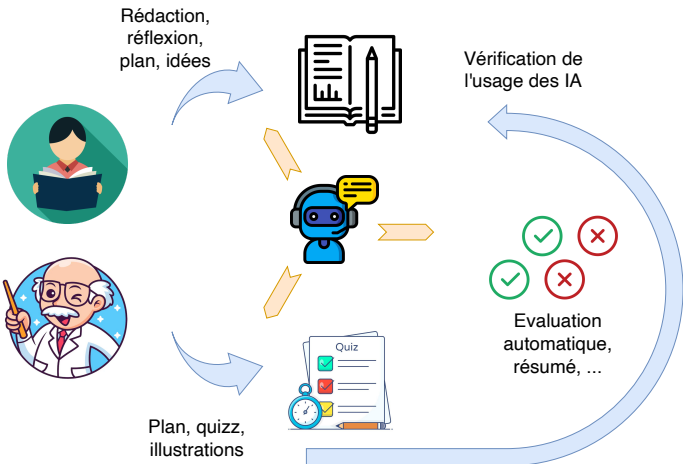
☆ Enregistrer Citer Cité 302 fois [Autres articles](#) [Les 2 versions](#)



Risques de généralisation de l'IA

IA partout = perte de sens?

- ▶ Transposition aux RH
- ▶ Aux systèmes de financement par projet





Détection des textes *générés* par chatGPT

L'externalité fait référence au fait qu'une activité économique d'un agent peut avoir un impact sur d'autres personnes sans qu'il y ait de compensation financière. Cela peut être bénéfique pour les autres, comme offrir une utilité gratuitement, ou nuisible, comme causer des dommages à l'environnement.

L'externalité caractérise le fait qu'un agent économique crée, par son activité, un effet externe en procurant à autrui, sans contrepartie monétaire, une utilité ou un avantage de façon gratuite, ou au contraire une nuisance, un dommage sans compensation (coût social, coût écosystémique, pertes de ressources pas, peu, difficilement, lentement ou coûteusement renouvelables...).

De la sorte, un agent économique se trouve en position d'influer consciemment ou inconsciemment sur la situation d'autres agents, sans que ceux-ci soient parties prenantes à la décision : ces derniers ne sont pas forcément informés et/ou n'ont pas été consultés et ne participent pas à la gestion de ses conséquences par le fait qu'ils ne reçoivent (si l'influence est négative), ni ne paient (si l'influence est positive) aucune compensation.

En résumé : « Tout coûte mais tout ne se paie pas »

Reformulation par chatGPT

Trier les documents par Date de dépôt 1 - 2 sur 2

Document	Match	Statut	Action
Plagiat Def 2 #4483eb 07/01/2023 19:18 par vous 122 mots 19,47 ko Plus d'infos	0%	Non détecté	Rapport
Plagiat Def 1 #f90ff3 07/01/2023 19:16 par vous 135 mots 16,78 ko Plus d'infos	100%	Détecté	Rapport

Définition de Wikipedia

Crédit: S.
Pajak



Détection des textes *générés par chatGPT*

GPTZero

Detect AI Plagiarism. Accurately



ORIGINALITY.AI

Chat GPT



AI Detector

by rchbankz

- ▶ **Classifieur** de texte (comme pour n'importe quel auteur)
 - ▶ Détection des biais dans le choix des mots / tournure de phrases
- ▶ Caractérisation de la **vraisemblance** des textes ([OpenAI](#), [GPTZero](#))
 - ▶ Hyper-fluidité des phrases, sur-abondance de connecteurs logiques
 - ▶ Modèle de langue = statistique \Rightarrow mesure entre distribution (**perplexité**)
- ▶ δ -**vraisemblance** sur des textes perturbés ([DetectGPT](#))
- ▶ [chatGPT](#) *devrait rapidement* intégrer des **empreintes** dans les textes générés

Détecteurs \Rightarrow $< 100\%$ de détection

+ niveau de confiance dans la détection

- dépend de la longueur du texte et des modifications apportées

≈ détecte des bouts de wikipedia ([chatGPT](#) = *stochastic parrot*)



Et au-delà de chatGPT? les outils pédagogique

Les EIAH –Environnements Informatiques pour l'Apprentissage Humain– n'ont pas attendu chatGPT pour proposer des solutions !

- ▶ Plateforme + outils (e.g. moodle)
 - ▶ Pédagogie
 - ▶ + Gestion
- ▶ Manuels scolaires augmentés
- ▶ Réseaux sociaux (pédagogiques) / Slack
- ▶ Outil de publication vidéo (pédagogie inversée)
- ▶ Sondage en direct, avec/sans téléphone



moodle

Explain
Everything

wooclap

Direct
Poll

Ne pas confondre *numérique* & IA



Et au-delà de chatGPT? les outils pédagogique (suite)

Les EIAH –Environnements Informatiques pour l'Apprentissage Humain– n'ont pas attendu chatGPT pour proposer des solutions !

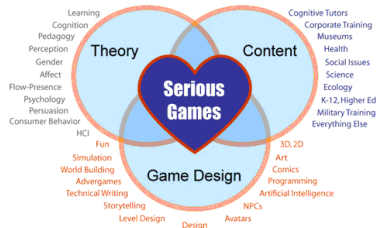
- ▶ Gamification des enseignements / serious game
 - ▶ Place du numérique
 - ▶ Pour quelles connaissances?
- ▶ MOOC / SPOC
 - ▶ FUN –France Université Numérique–

Réflexions par rapport à la crise du COVID...
... Et par rapport à la crise universitaire !

Ne pas confondre *numérique* & *IA*

coursera

The Heart of Serious Game Design





Et au-delà de chatGPT? La correction automatique

- QCM + scanner (+numéro d'anonymat) ⇒ correction automatique

+2/1/50+

QCM — Nombres relatifs — Somme

NOM : AVOGADRO PRÉNOM : AMEREO CLASSE : 401 DATE : LUNDI 2 SEPTEMBRE 2021

Pour chacune des questions, une seule réponse est possible. Vous devez noter le case qui correspond à la bonne réponse.

QUESTION 1 : Calculer $(+13) + (-7) + (-3)$

(-2) (+2) 0 (-1) (+3) (-3)

QUESTION 2 : Calculer $(-7) + (+11) + (-3)$

(+3) 0 (-1) (-2) (+2) (+1)

QUESTION 3 : Calculer $(-11) + (-3) + (+3) + (+7)$

(+3) (+2) (-2) (+4) (-4) (-3)

QUESTION 4 : Calculer $(-5) + (-3) + (+9)$

(+2) 0 (+3) (-2) (-1) (+1)

QUESTION 5 : Calculer $(-4) + (+8) + (-11) + (-7) + (+3)$

(-13) (+13) (+11) (-12) (-11) (-9)

QUESTION 6 : Calculer $(+3) + (-7) + (+8) + (-5)$

(+1) (+23) (+3) (-3) (-1) (-2)

QUESTION 7 : Calculer $(-5) + (-3) + (+7) + (-3) + (+11)$

(+5) (-7) (+6) (-5) (-6) (+7)

QUESTION 8 : Calculer $(+3,76) + (-7,4) + (+8,112) + (-3,76) + (-8,112)$


(-11,788) 0 (-7,4) (+3,76) (-8,112) (+3,567)

QUESTION 9 : Calculer $(-1,1) + (-3,3) + (+2,6) + (-1,3)$

(+2,9) (+1,2) (+2,6) 0 (-1,2) (-2,9)

QUESTION 10 : Calculer $(-7) + (-8) + (+15) + (+7) + (-15)$

(-8) (+8) (+7) (-15) (+15) (-7)



Ne pas confondre *numérique* & *IA*



Et au-delà de chatGPT? La correction automatique

- ▶ QCM + scanner (+numéro d'anonymat) ⇒ correction automatique
- ▶ Test unitaire (un mode de développement informatique à part entière)

```

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Ne pas confondre *numérique* & IA'



Et au-delà de chatGPT? La correction automatique

- ▶ QCM + scanner (+numéro d'anonymat) ⇒ correction automatique
- ▶ Test unitaire (un mode de développement informatique à part entière)
- ▶ Analyse du texte: comprendre si une réponse *contient* la vraie réponse ou pas

P^a	A senior is waiting at the window of a restaurant that serves sandwiches.	Relationship
H^b	A person waits to be served his food.	Entailment
	A man is looking to order a grilled cheese sandwich.	Neutral
	A man is waiting in line for the bus.	Contradiction
^a P, Premise. ^b H, Hypothesis.		

Ne pas confondre *numérique* & IA'



Et au-delà de chatGPT? La correction automatique

- ▶ QCM + scanner (+numéro d'anonymat) ⇒ correction automatique
- ▶ Test unitaire (un mode de développement informatique à part entière)
- ▶ Analyse du texte: comprendre si une réponse *contient* la vraie réponse ou pas
- ▶ Résumé automatique
⇒ on revient à chatGPT



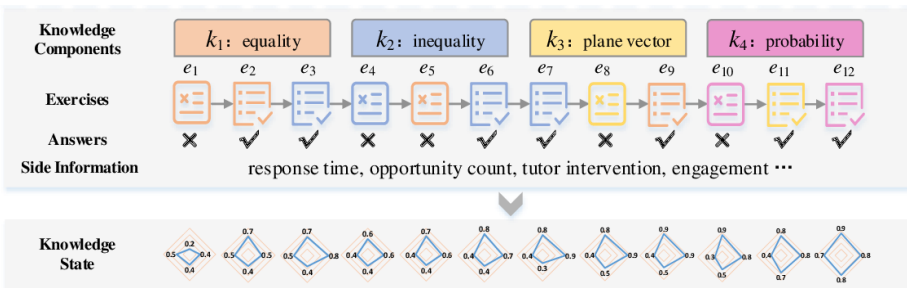
Ne pas confondre *numérique* & *IA*'

Et au-delà de chatGPT? Knowledge Tracing

- ▶ Système de recommandation... Pour l'apprenant
 - ▶ Vers une **pédagogie adaptative**...
 - ▶ ... \Leftrightarrow modéliser tous les concepts + niveaux de difficultés
 - ▶ Quel coût pour modéliser l'ensemble d'un cours?
 - ▶ Quel apport réel pour l'étudiant?
- PB mesure de performance d'un système de reco



Student

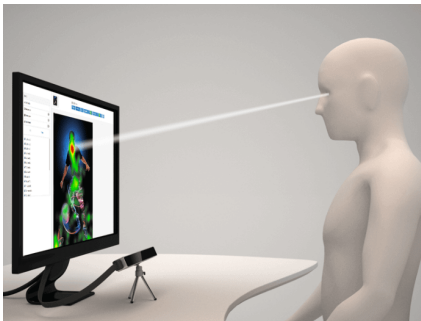




Et au-delà de chatGPT? Modélisation utilisateur / eye tracking

Il est néanmoins possible d'aller assez loin dans la modélisation de l'apprenant

- ▶ Internet / cookies
 - ▶ Eye tracking
- ⇒ Comprendre comment l'apprenant fonctionne pour mieux l'aider
- ▶ Analyse de la lecture, étude de la concentration



CONCLUSION



Souveraineté numérique: l'ensemble de la chaîne

Construction du modèle de base

Maitrise des données

- Collecte/équilibrage
- Nettoyage

Entraînement

- Puissance machine (milliers de GPU)
- Architecture/recherche ML

Raffinement du modèle

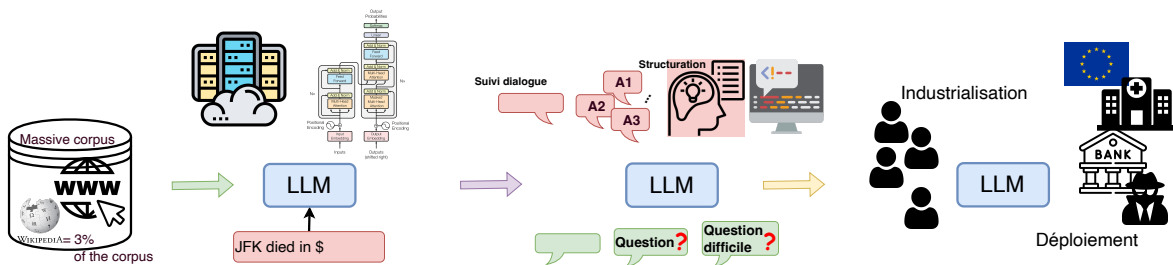
Maitrise & construction des données

- interactions humaines +++
- prix des données
- spécialisation à la demande

Exploitation du modèle

Optimisation / Limitation du coût

- compétence MLOps
- déploiement local





Une multitude de métiers

Architecte/Manager de la donnée

- Gère la donnée et son infrastructure (stockage)

Ingénieur de la donnée / Ingénieur Data

- Met à jour et interroge la donnée

Analyste de la donnée

- Visualise la données (graphiques)
- Analyse statistiquement les tendances

Data Scientist

- Interroge la donnée
- Implémente les algorithmes
- Visualise la donnée

Ingénieur Machine Learning

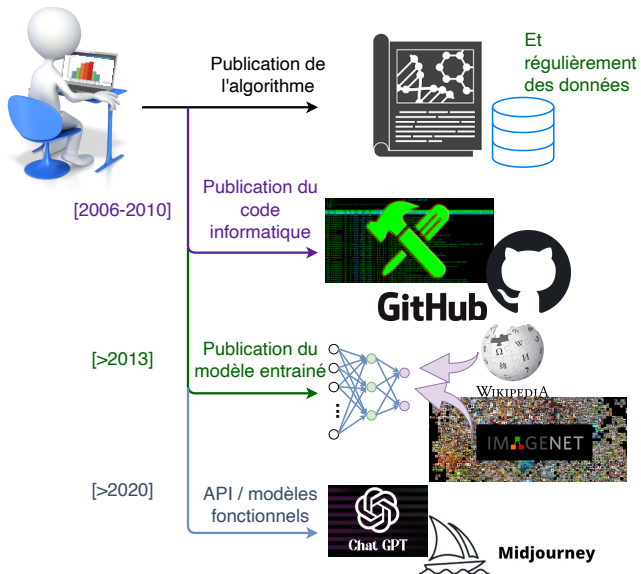
- Créé ou sélectionne les algorithmes de machine learning
- Evalue les algorithmes
- Industrialise les solutions

Utilisateur / Ingénieur Prompt

- Créé des prompt = nouvelles applications



Une multitude de métiers





Facteurs d'acceptabilité des IA générative

1 Utilitarisme :

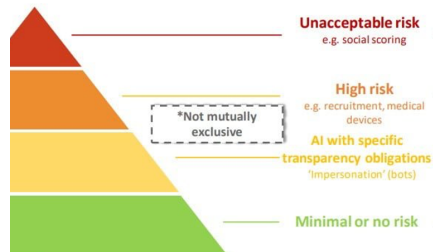
- ▶ Performances (facteur d'acceptation de chatGPT)
- ▶ Fiabilité / auto-évaluation

2 Non-dangérosité :

- ▶ Biais / correction
- ▶ Transparence (ligne éditoriale, confusion humain/machine)
- ▶ Implémentation fiable
- ▶ Souveraineté (?)
- ▶ Régulation (AI act)
 - ▶ Eviter les applications dangereuses

3 Savoir-faire :

- ▶ FormationS (usage/développement)





chatGPT: une simple étape

- ▶ Cout d'entrainement & réglages

4-5 Million d'euros / apprentissage \Rightarrow chatGPT est **mal appris** !

- ▶ Efficacité sur les données

chatGPT $>$ 1000x les lectures d'un homme dans une vie

- ▶ Repérer les entités, citer les sources

Ancrage des réponses dans les bases de connaissances

Ancrage des réponses dans des sources



Sam Altman 
@sama

...

ChatGPT launched on wednesday. today it crossed 1 million users!

8:35 AM · Dec 5, 2022

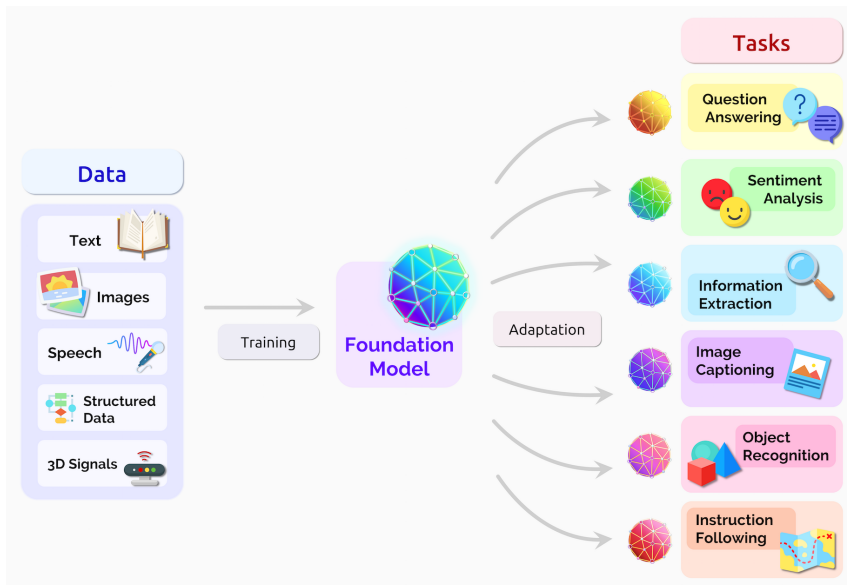
3,457 Retweets 573 Quote Tweets 52.8K Likes

- ▶ Multiplication des initiatives: GPT, LaMBDA, PaLM, BARD, BLOOM, Gopher, Megatron, OPT, Ernie, Galactica...

- ▶ Implication du public,
impact sur l'accès à l'information



Vers des modèles de fondation plus larges?





EXEMPLES



SE

C'est quoi le taux marginal de substitution en microéconomie ?



Le taux marginal de substitution (TMS) est une notion importante en microéconomie qui mesure la valeur que quelqu'un attribue à un bien ou un service par rapport à un autre.  

Plus précisément, le TMS mesure la quantité de bien A qu'une personne est prête à



abandonner pour obtenir une unité supplémentaire de bien B.

Le TMS peut être utilisé pour déterminer la courbe de consommation optimale. Une représentation graphique de tous les paniers de biens possibles est une droite qui passe par la courbe d'indifférence est une droite qui passe par la même valeur pour une personne. Le TMS est la pente de cette droite.

Le TMS est un concept important en microéconomie car il aide à comprendre comment les individus prennent des décisions d'achat et de vente des biens et des services. Il est également utilisé pour

SE

Qu'est-ce que le dol en droit français ? 

Le dol est un concept juridique qui désigne l'intention de tromper ou de frauder. Il peut être commis de manière active (c'est-à-dire en trompant ou en fraudant volontairement) ou de manière passive (c'est-à-dire en laissant volontairement une personne se tromper ou se faire frauder).  

En droit français, le dol peut être invoqué dans le cadre de différentes situations, telles que l'annulation d'un contrat, la nullité d'une donation ou encore la responsabilité civile délictuelle. Dans ces cas, la preuve du dol est essentielle pour obtenir gain de cause.







Question de comparaison

Réponse à la question
spécifique



Définitions « simples »

SE Dans le système juridique français, lequel est supérieur entre un arrêté et un décret ? 

 Dans le système juridique français, un décret est supérieur à un arrêté.  

Un décret est un acte administratif pris par le gouvernement ou par le Premier ministre, qui a force de loi. Il est destiné à réglementer des matières de droit public et doit être publié au Journal officiel de la République française pour entrer en vigueur.

Un arrêté est un acte administratif pris par le gouvernement, un membre du gouvernement ou un haut fonctionnaire de l'État, qui a force de loi dans un domaine précis. Il est destiné à mettre en œuvre une politique ou une réglementation définie par un décret et doit être publié au Journal officiel de la République française pour entrer en vigueur.

En résumé, un décret est supérieur à un arrêté dans la hiérarchie des normes en France, car il a une portée générale et régleme des matières de droit public. Un arrêté, quant à lui, a une portée limitée et met en œuvre une politique ou une réglementation définie par un décret.



Encyclopédie statique (+recherche avancée !)

[S. Pajak]

SE Explique-moi pourquoi les capitaux propres sont au passif et pas à l'actif ?

Les capitaux propres (aussi appelés fonds propres ou equity en anglais) sont une partie du bilan d'une entreprise qui représente les ressources financières dont elle dispose. Ils sont composés des fonds apportés par les propriétaires de l'entreprise (les actionnaires) ainsi que de tous les bénéfices non distribués de l'entreprise.

Les capitaux propres sont considérés comme étant au passif du bilan, car ils représentent une dette de l'entreprise envers ses propriétaires. En effet, les actionnaires ont investi de l'argent dans l'entreprise en achetant des actions et ils ont droit à une partie des bénéfices de l'entreprise. C'est pourquoi les capitaux propres sont considérés comme une dette de l'entreprise envers ses propriétaires.

En revanche, l'actif du bilan représente les ressources financières de l'entreprise qui sont utilisées pour générer des bénéfices, comme les biens et les équipements de l'entreprise, les comptes clients, les stocks, etc. Les capitaux propres ne sont pas considérés comme faisant partie de l'actif, car ils ne sont pas directement utilisés pour générer des bénéfices.

Question problématisée

Définition « simple »

Réponse personnalisée, qui traite la problématique dans la question





SE Comment on fait une regression OLS simple en R de y sur x



Pour effectuer une régression OLS simple en utilisant R, vous pouvez utiliser la fonction `lm()` (qui signifie "linear model" en anglais) de R. Cette fonction permet de créer un modèle de régression linéaire.



Voici un exemple de comment utiliser la fonction `lm()` pour créer un modèle de régression OLS simple pour y en fonction de x :

```
R Copy code
# Tout d'abord, il faut charger les données
x <- c(1, 2, 3, 4, 5)
y <- c(2, 4, 6, 8, 10)

# Ensuite, on utilise la fonction lm() pour créer un modèle de régression OLS
model <- lm(y ~ x)

# On peut maintenant afficher les résultats du modèle
summary(model)
```

La sortie affichera les informations sur le modèle, comme les coefficients, les erreurs standard, les t-values et les p-values.

Il est important de noter que cet exemple est très simpliste, il est souvent nécessaire de prétraiter les données (par exemple en enlevant les outliers ou les valeurs manquantes) et de vérifier les hypothèses du modèle (linéarité, normalité, homoscedasticité, indépendance des erreurs) avant de procéder à la régression.



U.S. Pours Money Into Chips, but Even Soaring Spending Has Limits

SE

Please explain "Soaring Spending" in the sentence "U.S. Pours Money Into Chips, but Even Soaring Spending Has Limits"



In the context of this sentence, "soaring spending" refers to the act of spending a large amount of money. The word "soaring" suggests that the spending is increasing rapidly or at a high rate.



The sentence mentions that the U.S. is pouring money into chips, which means that it is investing a large amount of money in the development and production of chips. However, the sentence also suggests that this spending has limits, which means that there is a maximum amount of money that the U.S. is willing or able to invest in chips.

Overall, the phrase "soaring spending" refers to the high level of investment that the U.S. is making in chips, but it also implies that this level of investment is not unlimited.

Définition

Reformulation
complète du titre

(d'ailleurs fausse en
partie)



Détection du plagiat

[S. Pajak]

L'externalité fait référence au fait qu'une activité économique d'un agent peut avoir un impact sur d'autres personnes sans qu'il y ait de compensation financière. Cela peut être bénéfique pour les autres, comme offrir une utilité gratuitement, ou nuisible, comme causer des dommages à l'écosystème économique ou qui ne sont pas compensés par le coût, mais...

L'externalité caractérise le fait qu'un agent économique crée, par son activité, un effet externe en procurant à autrui, sans contrepartie monétaire, une utilité ou un avantage de façon gratuite, ou au contraire une nuisance, un dommage sans compensation (coût social, coût écosystémique, pertes de ressources pas, peu, difficilement, lentement ou coûteusement renouvelables...).

De la sorte, un agent économique se trouve en position d'influer consciemment ou inconsciemment sur la situation d'autres agents, sans que ceux-ci soient parties prenantes à la décision : ces derniers ne sont pas forcément informés et/ou n'ont pas été consultés et ne participent pas à la gestion de ses conséquences par le fait qu'ils ne reçoivent (si l'influence est négative), ni ne paient (si l'influence est positive) aucune compensation.

En résumé : « Tout coûte mais tout ne se paie pas »

Reformulation par chatGPT

Tout cocher

Trier les documents par Date de dépôt

1 - 2 sur 2

<input type="checkbox"/>		Plagiat Def 2 #4483eb 📅 07/01/2023 19:18 par vous 122 mots 19,47 ko Plus d'infos		● 0%	<div style="width: 100%; height: 10px; background-color: gray;"></div>	Rapport	⋮
<input type="checkbox"/>		Plagiat Def 1 #f90ff3 📅 07/01/2023 19:16 par vous 135 mots 16,78 ko Plus d'infos		● 100%	<div style="width: 100%; height: 10px; background-color: red;"></div>	Rapport	⋮

Définition de Wikipedia



Hallucination



Peux tu me donner une bibliographie des VAE sur les séries temporelles?



Je m'excuse pour la confusion. Voici donc une sélection de références sur les Variational Autoencoders (VAE) pour les séries temporelles :

1. "Variational Autoencoder for Deep Learning of Images, Labels and Sequences" par Zhao et al. (2017) - Cette étude présente une approche VAE pour la modélisation de séquences temporelles en utilisant une architecture de réseaux de neurones récurrents.
2. "Variational Recurrent Autoencoders" par Chung et al. (2016) - Cette étude introduit une approche VAE pour la modélisation de séquences temporelles en utilisant des réseaux de neurones récurrents et une

3. "Generative Modeling for Time Series:

Variational recurrent auto-encoders

[O Fabius, JR Van Amersfoort](#) - arXiv preprint arXiv:1412.6581, 2014 - arxiv.org

In this paper we propose a model that combines the strengths of RNNs and SGVB: the **Variational Recurrent** Auto-Encoder (VRAE). Such a model can be used for efficient, large scale ...

☆ Enregistrer Citer Cité 302 fois Autres articles Les 2 versions

► Souvent sur des connaissances rares ⇒ difficile à mémoriser



Assistant personnel intelligent

[S. Pajak]

SE

Ecris un mail à l'administration pour décaler mon cours d'économie du lundi 13h au jeudi 9h.

1



Objet : Demande de décalage de cours d'économie



Monsieur ou Madame,

Je vous écris aujourd'hui pour vous demander de bien vouloir décaler mon cours d'économie du lundi 13h au jeudi 9h.

En effet, j'ai un empêchement le lundi à cette heure-là et je suis dans l'impossibilité de suivre mon cours à cette heure. Le jeudi 9h serait une plage horaire beaucoup plus adaptée à mon emploi du temps.

Je vous remercie par avance pour votre compréhension et reste à votre disposition pour toutes informations complémentaires.

Dans l'attente de votre réponse, veuillez agréer, Monsieur ou Madame, l'expression de mes salutations distinguées.

Cordialement,

[Votre nom]



Assistant personnel intelligent

[S. Pajak]

SE Ecris un mail à l'administration pour décaler mon cours d'économie du lundi 13h au jeudi 9h. 1

1

 Objet : Demande de décalage de cours d'économie  

Monsieur ou Madame,

Je vous écris aujourd'hui pour vous demander de bien vouloir décaler mon cours d'économie du lundi 13h au jeudi 9h.

En effet, j'ai un empêchement le lundi à cette heure-là et je suis dans l'impossibilité de suivre mon cours à cette heure. Le jeudi 9h serait une plage horaire beaucoup plus adaptée à mon emploi du temps.

Je vous remercie par avance pour votre compréhension et reste à votre disposition pour toutes informations complémentaires.




Dans l'attente de votre réponse, veuillez agréer, Monsieur ou Madame, l'expression de mes salutations distinguées.

Cordialement,

[Votre nom]

SE Ecris un mail aux étudiants pour décaler mon cours d'économie du lundi 13h au jeudi 9h. 2

2

 Objet : Modification de l'horaire du cours d'économie  

Chers étudiants,

Je vous informe que le cours d'économie prévu le lundi 13h a été décalé au jeudi 9h.

Je vous prie de bien vouloir noter cette modification dans vos agendas et de vous présenter au cours à cette nouvelle heure.

Je vous remercie pour votre compréhension et vous dis à très bientôt.

Cordialement,

[Votre nom]



Assistant personnel intelligent

[S. Pajak]

SE

Ecris un mail de candidature spontanée à un stage d'assistant comptable en cabinet.



Objet : Candidature pour un stage d'assistant comptable



Monsieur ou Madame,

Je vous écris aujourd'hui pour vous présenter ma candidature pour un stage d'assistant comptable au sein de votre cabinet.

Actuellement étudiant(e) en comptabilité et finance, je suis à la recherche d'un stage de fin d'année afin de mettre en pratique les connaissances acquises lors de mes études et de découvrir le monde de la comptabilité en cabinet.

Je suis particulièrement intéressé(e) par votre cabinet pour sa réputation et la qualité de ses services. Je suis persuadé(e) que cette expérience serait pour moi une occasion unique de développer mes compétences et de m'épanouir professionnellement.

Je suis rigoureux, organisé et j'aime travailler en équipe. Je suis également doté(e) d'un excellent relationnel et je suis persuadé(e) que je pourrais apporter une valeur ajoutée à votre équipe.

Je vous remercie par avance pour l'attention que vous porterez à ma candidature et me



Assistant personnel intelligent

[S. Pajak]

SE Ecris un mail de candidature spontanée



Objet : Candidature pour un stage d'ass

Monsieur ou Madame,

Je vous écris aujourd'hui pour vous prés
comptable au sein de votre cabinet.

Actuellement étudiant(e) en comptabili

fin d'année afin de
de découvrir le m

Je suis particulièr
ses services. Je s



unique de dévelop


Je suis rigoureux,
excellent relation
votre équipe.

Je vous remercie
tiens à votre dispo



SE Propose moi 3 phrases différentes pour dire ce que l'entreprise gagnerait à m'accepter en stage.



1. "Je suis convaincu que ma motivation et mon dynamisme pourraient être un atout pour votre entreprise et que je pourrais apporter une valeur ajoutée à votre équipe."  
2. "Ma rigueur et mon organisation me permettraient de m'intégrer rapidement et de prendre en main efficacement les tâches qui me seraient confiées."
3. "Je suis passionné par mon domaine d'études et j'aimerais mettre mes connaissances

SE Donne une phrase professionnelle pour dire que j'ai déjà travaillé comme assistant comptable l'année dernière et que j'ai aimé 



Voici une phrase professionnelle qui exprime que vous avez déjà travaillé comme assistant comptable l'année dernière et que vous avez apprécié cette expérience :  

"J'ai eu l'opportunité de travailler comme assistant comptable l'année dernière et j'ai vraiment apprécié cette expérience. Cela m'a permis de mettre en pratique mes connaissances et de développer de nouvelles compétences, tout en travaillant dans un environnement stimulant et professionnel."



Assistant personnel intelligent

[S. Pajak]



propose moi un titre pour une présentation qui traite des conséquences de l'IA dans la pédagogie et les enseignements à l'université.



Voici quelques suggestions de titres pour une présentation sur les conséquences de l'IA dans la pédagogie et l'enseignement à l'université :



1. "L'IA et la pédagogie universitaire : nouvelles opportunités pour l'apprentissage"
2. "Comment l'IA peut transformer l'enseignement universitaire"
3. "L'IA et l'université : les défis de la pédagogie de demain"
4. "L'IA et l'enseignement supérieur : une nouvelle ère pour l'apprentissage"
5. "L'apport de l'IA à l'enseignement universitaire : quelles conséquences pour les étudiants et les enseignants?"



SE

Les titres doivent tenir sur une ligne

Enchaînement à partir du
résultat précédent

Voici quelques suggestions de titres pour une présentation sur les conséquences de l'IA dans la pédagogie et l'enseignement à l'université, qui tiennent sur une seule ligne :



1. "L'IA et l'enseignement universitaire: opportunités et défis"
2. "Transformer l'enseignement universitaire avec l'IA"
3. "L'IA et la pédagogie de demain: enjeux et perspectives"
4. "L'IA et l'apprentissage universitaire: nouvelles possibilités"
5. "L'IA et l'enseignement supérieur: révolution ou évolution?"



SE

des titres plus créatifs stp



Prompt en langage naturel



Voici quelques propositions de titres créatifs pour une présentation sur les conséquences de l'IA dans la pédagogie et les enseignements à l'université :



- "Le cerveau augmenté : l'IA au service de l'apprentissage universitaire"
- "Apprendre autrement grâce à l'IA : nouvelles perspectives pour l'enseignement universitaire"
- "L'IA, une supercherie pour l'enseignement universitaire ?"
- "L'IA, un atout ou un piège pour l'enseignement universitaire ?"
- "AI University : quand l'IA révolutionne l'enseignement universitaire"
- "L'IA, un allié ou un rival pour l'enseignement universitaire ?"



Assistant personnel intelligent

[S. Pajak]

SE

turn the following citation into BibTex format: Anne Boring. (2017) Gender Biases in Student Evaluations of Teaching, Journal of Public Economics, 145, 27-41.



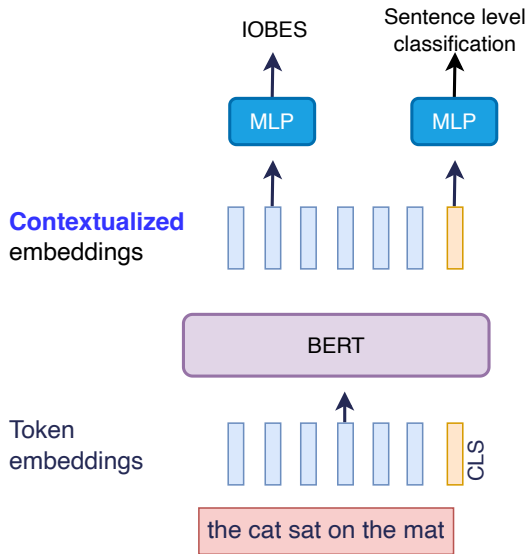
```
@article{boring2017gender,  
title={Gender Biases in Student Evaluations of Teaching},  
author={Boring, Anne},  
journal={Journal of Public Economics},  
volume={145},  
pages={27-41},  
year={2017}  
}
```



NLP *vs* GÉNÉRATIF



Extraction de connaissances



Words	BIOES Label
Jane	B-PER
Villanueva	E-PER
of	O
United	B-ORG
Airlines	I-ORG
Holding	E-ORG
discussed	O
the	O
Chicago	S-LOC
route	O
.	O

- ▶ Etat de l'art en NER
- ▶ Efficace sur l'extraction de relation



Extraction de connaissances

V

Extract entities from the following wikipedia text:

Mozart naît le 27 janvier 1756 à 8 heures du soir au numéro 9 de la Getreidegasse à Salzbourg. Il est le fils de Léopold Mozart, musicien, compositeur et pédagogue originaire d'Augsbourg, ville libre d'Empire, qui occupe alors la fonction de vice-maître de chapelle à la cour du prince-archevêque de Salzbourg, et d'Anna Maria Pertl, sa femme⁴, fille d'un fonctionnaire de la cour de Salzbourg.



From the given text, the entities that can be extracted are:



- Person: Mozart, Léopold Mozart, Anna Maria Pertl
- Location: Salzbourg, Augsbourg, Empire
- Date: 27 janvier 1756
- Time: 8 heures du soir
- Organization: la cour du prince-archevêque de Salzbourg, la cour de Salzbourg



Extraction de connaissances

V

can you give me the relations between thos entities?



Sure, here are the relationships between the entities extracted from the text:



- Mozart is the son of Léopold Mozart and Anna Maria Pertl
- Léopold Mozart is a musician, composer, and pedagogue from Augsburg and was working as the Vice-Kapellmeister at the court of the Prince-Archbishop of Salzburg
- Anna Maria Pertl was the wife of Léopold Mozart and daughter of a court official in Salzburg
- Mozart was born on 27th January 1756 at 8 PM at number 9 of Getreidegasse in Salzburg
- Léopold Mozart and Anna Maria Pertl lived in Salzburg, which was the capital of the Prince-Archbishopric of Salzburg
- Augsburg was Léopold Mozart's birthplace and was a free city of the Empire
- The Prince-Archbishopric of Salzburg was the political jurisdiction where Salzburg was located.



Question Answering

Les modèles de langue savent répondre aux questions... Dans un certain contexte :)

Passage Sentence

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.

Question

What causes precipitation to fall?

Answer Candidate

gravity

- ▶ Est-il possible/facile d'extraire des passages avec un moteur de recherche?
- ▶ Est-ce une alternative aux bases de connaissances?
- ▶ Quelles différences avec chatGPT?



Rajpurkar et al., ACL 2018

Know What You Don't Know: Unanswerable Questions for SQuAD



Question Answering & chatGPT



how old is Obama



Barack Obama was born on August 4, 1961, making him 61 years old as of February 2, 2023.



what is your source?



I am a machine learning model created by OpenAI, I was trained on a diverse range of texts from the internet. The information I provide is based on the patterns and relationships in that text, and is not based on any one specific source.





Question Answering & chatGPT

V how old is Obama



Barack Obama was born on August 4, 1961, making him 61 years old as of February 2, 2023.



V how old is obama?



As of 2021, Barack Obama was born on August 4, 1961, so he is 60 years old.



V and today?



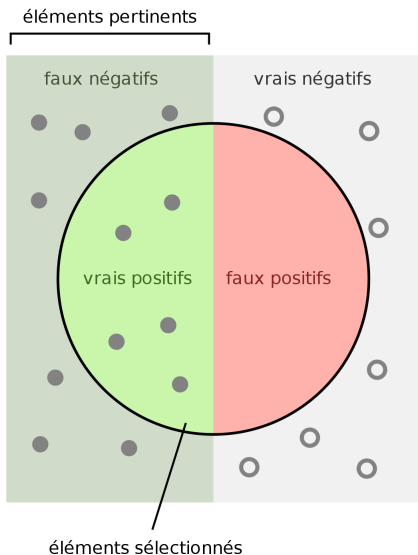
As a language model AI, I don't have real-time access to current dates. However, Barack Obama was born on August 4, 1961, so if today's date is after August 4th, 2021, he would



MÉTRIQUE



Précision et rappel



Combien de candidats sélectionnés sont pertinents ?

$$\text{Précision} = \frac{\text{vrais positifs}}{\text{vrais positifs} + \text{faux positifs}}$$

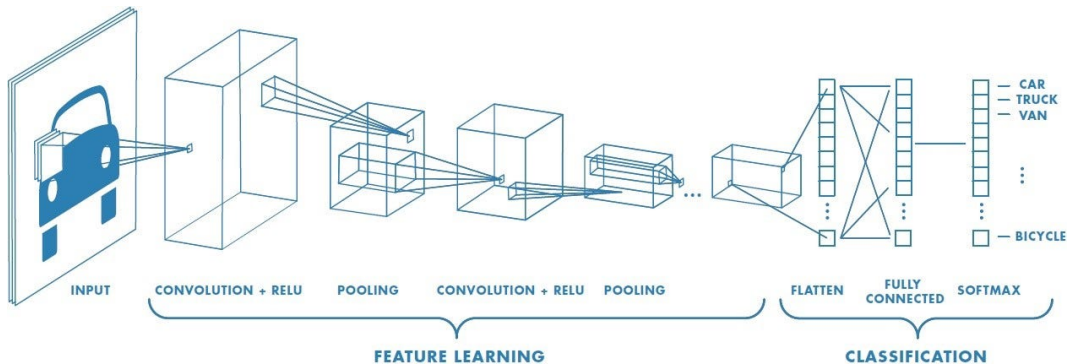
Combien d'éléments pertinents sont sélectionnés ?

$$\text{Rappel} = \frac{\text{vrais positifs}}{\text{vrais positifs} + \text{faux négatifs}}$$

- ▶ Précision ↗ = ne pas dire de bêtise
- ▶ Rappel ↗ = couvrir les informations pertinentes
- ▶ F1 = moyenne (harmonique) de la précision et du rappel

GPT-4: MODÈLE MULTI-MODAUX

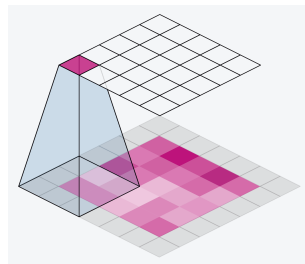
Apport de l'architecture transformer en vision



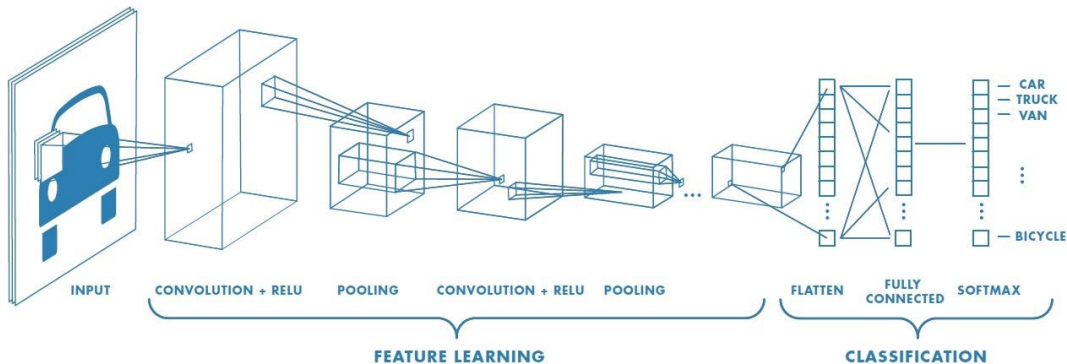
1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



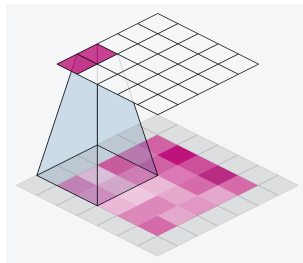
Apport de l'architecture transformer en vision



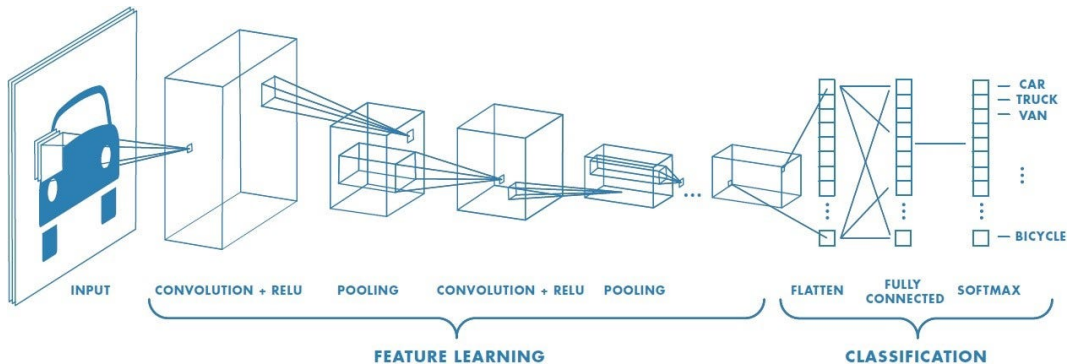
1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



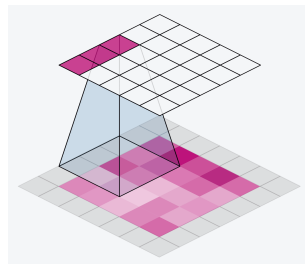
Apport de l'architecture transformer en vision



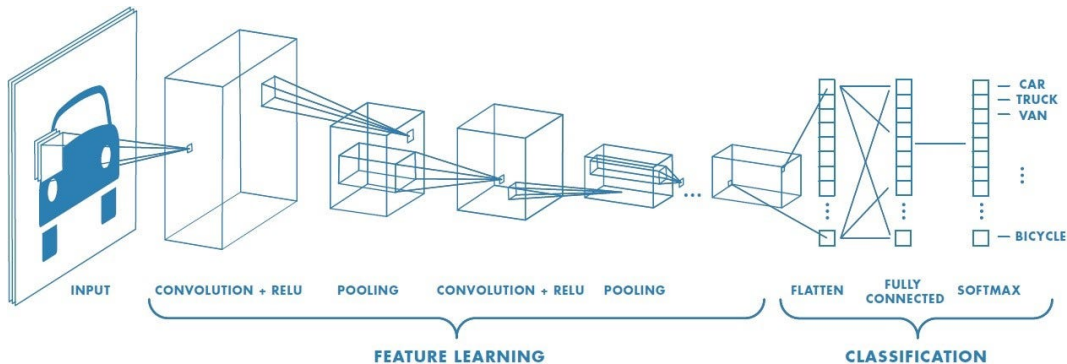
1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



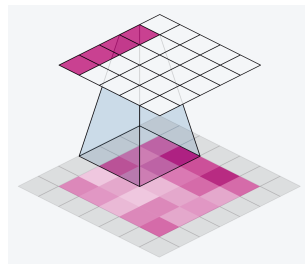
Apport de l'architecture transformer en vision



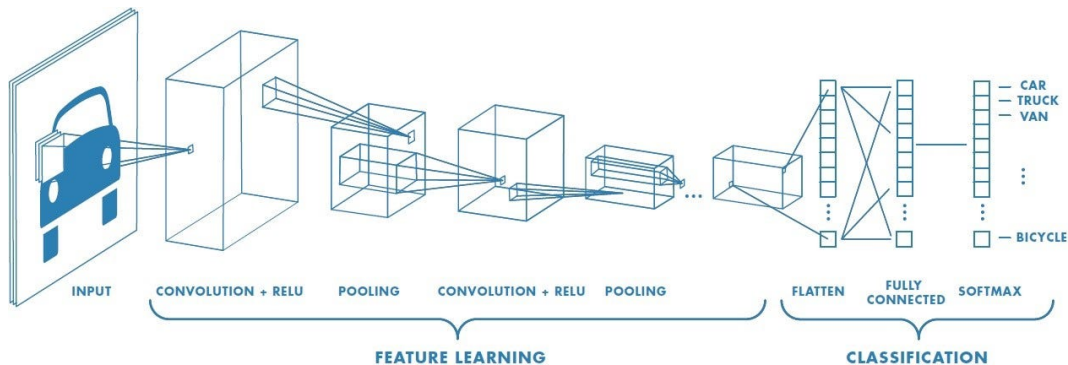
1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



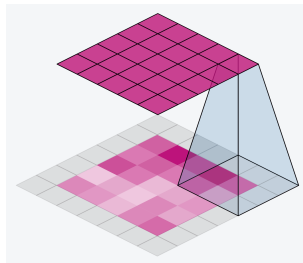
Apport de l'architecture transformer en vision



1. Convolution

Filtre d'analyse

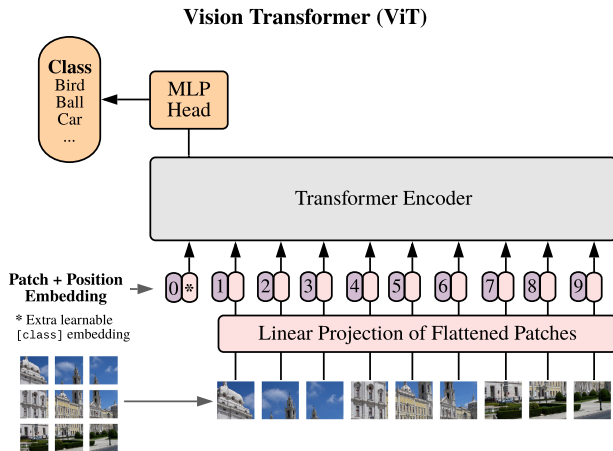
- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



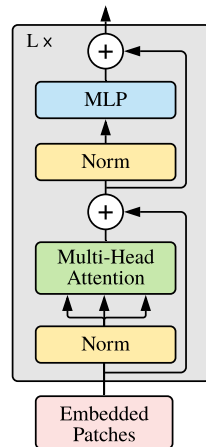
Apport de l'architecture transformer en vision

2. Transformer

- ▶ Apprentissage des motifs (locaux) à extraire
- ▶ Analyse des interactions entre les portions d'images



Transformer Encoder





Mieux comprendre les images, mieux modéliser le texte

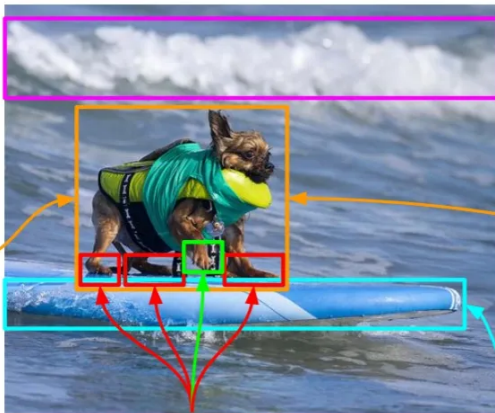
Visual Grounding

Where does this scene take place?

- A) In the sea. ✓
- B) In the desert.
- C) In the forest.
- D) On a lawn.

What is the dog doing?

- A) Surfing. ✓
- B) Sleeping.
- C) Running.
- D) Eating.



Which paw is lifted?

Why is there foam?

- A) Because of a wave. ✓
- B) Because of a boat.
- C) Because of a fire.
- D) Because of a leak.

What is the dog standing on?

- A) On a surfboard. ✓
- B) On a table.
- C) On a garage.
- D) On a ball.

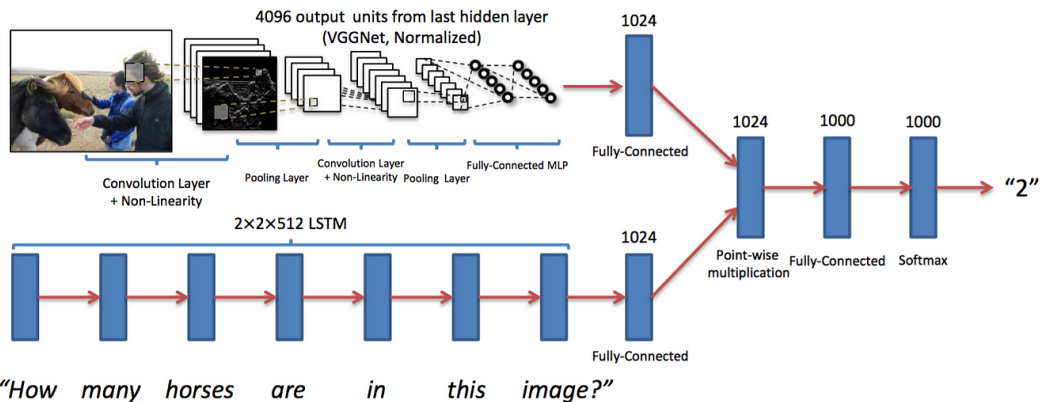
Comprendre les éléments d'une image
Intégrer du bon sens (common sense reasoning)

Modéliser le contexte des mots



Apprendre un système multimodal

L'exemple du VQA: visual question answering



⇒ Rétro-propager l'erreur ⇒ modifier les représentations de mots + analyse de l'image

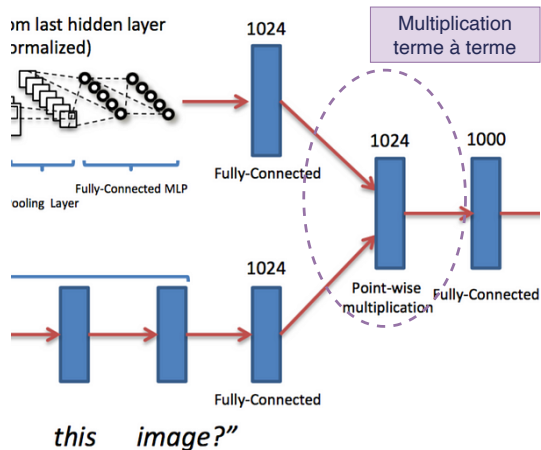
Compréhension globale / approfondie

Risque de spécialisation /
Oubli de la connaissance générale



Processus de fusion multi-modale

- ▶ Simple multiplication terme à terme des représentations latentes
- ▶ Concaténation et mixture complexe





Opportunité multi-modale

Attention multi-modale:

Mot / Portion d'image

⇒ Modélisation fine

⇒ Distinction éléments /
background

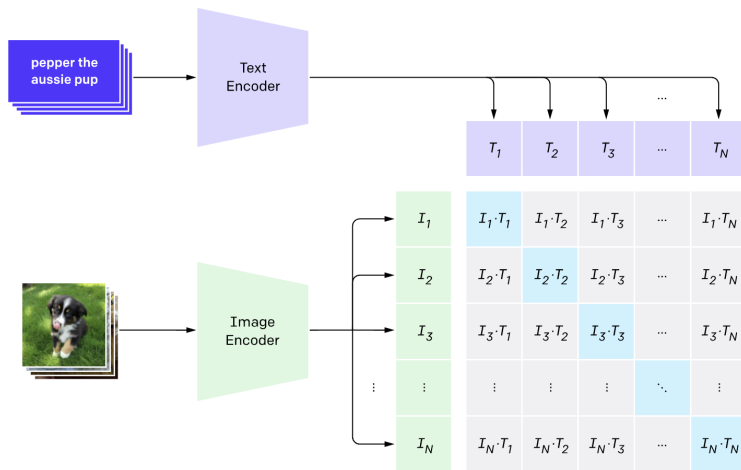


Illustration = encodeur seul (compréhension)

⇒ La même architecture permet d'obtenir du texte ou des images en sortie