

Machine Learning & IA générative expliqué(e)s à des non-informaticien(ne)s

Michaël PÉRIN, Verimag / UGA

juin 2024

Qui parle ? Michaël PÉRIN, enseignant-chercheur de l'Université de Grenoble Alpes

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

- Formation de mathématicien, doctorat en informatique

Maître de conférence en informatique en école ingénieur

- Chercheur au **laboratoire VERIMAG** : vérification et preuves mathématiques de programmes critiques
- Co-directeur de l'**équipe ETicS** (Environnement, Technologie de l'information, Société) qui étudie l'**impact sociétal et environnemental du numérique**

Cherche à estimer l'impact du *Machine Learning*

- pas chercheur en IA
- préparation d'un cours d'introduction à l'apprentissage machine

Terminologie

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

Le terme « Intelligence Artificielle » est surchargé

- de fantasmes, d'affects et d'anthropomorphisme
- inapproprié car il ne s'agit pas d'intelligence au sens commun,

Préférez « Machine Learning »

- et encore . . . apprentissage suggère une forme de compréhension

Le terme approprié

construction itérative d'un modèle statistique, par exemples et contre-exemples

ToC – Historique

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- Apprentissage Machine
- IA conversationnelle

Historique de l'apprentissage machine

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- 1545-1750 : Calcul matriciel (Cardan), dérivée et raffinement par itération (Newton)
- 1940-1960 : modèles neuronaux et simulation électronique
- **1960-1980 : simulation informatique de réseaux de neurones (NN, Neural Network)** et apprentissage
- **1981-1987 : réseaux de neurones multicouches (Deep NN)** et **apprentissage profond (Deep Learning)**
- 1990 : internet
- **2010 : gigantesques bases de données d'entraînement**
- 2012 : NN en reconnaissance visuelle : -10% d'erreur
- **2016 : AlphaGo \neq NN**, algorithme entraîné contre lui-même, bat le meilleur joueur mondial de go

Historique de l'IA conversationnelle

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

- **1950** : Alan Turing propose le critère "**Imitation Game**": distinguer la machine de l'humain au moyen d'un dialogue ?
- **1956** : Premier séminaire d'**Intelligence Artificielle** organisé par John McCarthy (programme de jeu d'échec, LISP)
- **1964** : **ELIZA** (128 Ko), **premier agent conversationnel** trompe les utilisateurs qui le prennent pour un psychologue en ligne
- 1968 : SHRDLU = 1er programme de dialogue pilotant un robot dans un environnement inconnu
- 1972 : 1er programme de reconnaissance vocale
- **2011-2014** : **quasi-succès au test de Turing** : tromper 30% des juges dans un dialogue de 5 min
- **2022** : les utilisateurs sont bluffés par **ChatGPT** (175 Go > 1 300 000 × ELIZA)

ToC – Principes du *Machine Learning*

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- Ce que (sait / ne sait pas) faire le *ML*
- GPT = *Generative Pretrained Transformer*
- Fonctionnement d'un *Large Language Model*
- L'importance du contexte
- Les *Deep Neural Networks*
- L'apprentissage
- Représentation sémantique et Traduction

Ce que (sait / ne sait pas) faire le ML

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

Ce que sait faire le Machine Learning : **approche statistique**

- à l'origine : tâche spécifique et hyper spécialisée
- désormais : détecter des corrélations, généraliser
- simuler, reproduire du plausible, faire des variations

Ce que ne sait pas faire le ML : **raisonnement déductif**

- ne capture pas les notions de cohérence, de vérité, de morale, de logique, de représentation du monde, de bon sens, ...
- faire ressortir les idées rares mais pertinentes
- produire du nouveau, de l'inattendu, de l'impensé
- abstraire/conceptualiser
- découvrir des liens de causalités, raisonner par analogie

GPT = *Generative Pretrained Transformer*

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

transformeur : prend et retourne le même type de donnée

- son → son (AudioCraft, Google MusicLM, ...)
- image → image (Bing Image Creator, DALL-E, ...)
- texte → texte (ChatGPT, Bing Chat, Google Bard, ...)

génératif : génère une suite plausible dans un contexte donné

- ChatGPT génère **le prochain mot le plus plausible pour compléter le texte donné**

pré-entraîné = peu d'entraînement pour le spécialiser

- évite de refaire l'entraînement de base
- généraliste ⇒ entraînement spécifique à une application

Fonctionnement des *Large Language Models*

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Numérisation

Chaque mot du dictionnaire de la langue est numéroté de 1 à quelques milliers (*~ 100 000 pour Le Petit Robert*)

Entrée : le contexte = un texte numérisé

(longtemps, je, me, suis, couché, de, bonne) = (12621, 10223, 13407, 19076, 3014, 4045, 2701)

Sortie : distribution statistique du prochain mot dans les exemples rencontrés à l'entraînement

(98% heure, 0.1% humeur, ...)

L'importance du contexte

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Le contexte modifie la statistique du prochain mot

- (bonne)
→ (20% journée, 19% affaire, 12% nuit, 10% humeur, ...)
- (de, bonne)
→ (90% humeur, 1% heure, ...)
- (couché, de, bonne)
→ (90% heure, 1% humeur, ...)

Modèle statistique = *Deep Neural Networks*

Machine
Learning &
IA générative

chaque couche de neurones = une matrice de poids statistiques

Une matrice de poids statistique = un tableau excel de nombres

- prend en entrée une séquence de nombres (*le contexte*)
- applique les poids (multiplication matrice-vecteur)
- produit une séquence de nombres (*la probabilité de chaque mot d'être le prochain*)

On ne sait pas interpréter les poids du modèle

- on ne sait pas dire « *ce qu'il sait* »
- on sait juste l'**appliquer** : $\text{Modèle}(\text{contexte}) = \text{probabilités}$
- et le **corriger** : *la phase d'apprentissage*

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

L'apprentissage

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Pour chaque exemple on a : **Contexte + mot attendu**

- on calcule la sortie *Modèle(Contexte)*
- on calcule l'erreur = la différence entre les probabilités produites par le modèle et le mot attendu
- on corrige les poids du modèle pour annuler cette différence (*par retropropagation de l'erreur*, LeCun et al., Prix Turing 2018)

On effectue ces corrections du modèle pour **chaque exemple** de la base d'apprentissage

- Si le mot attendu a une probabilité *trop faible*, on corrige les coefficients du modèle
- de manière à augmenter *légèrement* la probabilité du mot

Réprésentation sémantique d'un phrase

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Séquence des mots + Matrice(s) d'Attention

- une *Attention Matrix* se réduit à 128 concepts
- capture l'influence de tout le contexte (2048 mots) sur les 12 288 caractéristiques de chaque mot

permet une traduction sémantique dans une autre langue : ENG

- « Il était une fois »
- Concepts = INCIPIIT, RÉCIT, FICTION, CONTE, PASSÉ, FR
- « Once upon a time »

Les capacités surprenantes des statistiques

- les concepts ne sont pas définis *a priori*,
- ils émergent des statistiques



Toc – Focus sur la famille ChatGPT

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Le *Large Language Model* sous jacent se nomme
GPT 3.0, ..., 4.0

Entraînement de GPT 3.0

- 175 Milliards de paramètres = les poids du modèle ajustés sur
- 500 Milliards d'exemples parmi une sélection de textes

GPT3 en chiffres : l'entraînement [wikipedia]

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Base d'entraînement (500 Milliards d'exemples)

- Common Crawl : 410 Milliards
- WebText2 : 19 Milliards
- Books1 : 12 Milliards
- Books2 : 55 Milliards
- Wikipedia : 3 Milliards

Coût financier et énergétique [T.Gebru et al.]

- 150 000 dollars
- 550 t de CO₂
- acceptable? si son application est utile et peu coûteuse

GPT3 en chiffres : le modèle [wikipedia]

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Modèle à 175 Milliard de paramètres

un tableau excel de 420 000 lignes et 420 000 colonnes

175 000 000 000 opérations pour **chaque** mot généré

- en $\frac{175 \times 10^9}{30 \times 10^{12}} = 6$ millième de seconde sur un processeur spécialisé (GPU à 30 Tflops) à 600€
- soit 6 secondes pour un texte de 1000 mots

Pas exécutable sur une ordinateur personnel

- 175 Go de mémoire vive pour stocker le modèle
- 16 Go sur un PC milieu de gamme ; double tous les 2 ans
- au mieux **disponible sur bureau dans 6 à 8 ans...**
- ⇒ **dépendance aux DataCenters des GAFAM, BATX**

Dans les griffes du Chat

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

ChatGPT a changé la perception des chatbots

- 2019 : 72% des Français préfèrent le contact humain à celui d'un robot lors de leur parcours d'achat.
- 2023: 62% des consommateurs préfèrent s'adresser à un chatbot plutôt qu'à un humain

suite

- Large Models = Machine Learning = modèle statistique
- Les limites de ChatGPT et consorts
- Le marketing derrière le chat
- Pourquoi ChatGPT est-il si bluffant ?

Large Models = Machine Learning = Statistique

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

Les *ML* est un **outil statistique puissant**

- nécessite des données taguées, de qualité en grand nombre
- qui continue d'apprendre à chaque demande

Basique + Force Brute

Génération statistique dans un contexte donné

- **impressionnant** quand le modèle atteint **175 Milliards de paramètres** mais
- **coûteux** : 1 Million × la consommation cerveau humain

Bluffant mais ça reste du **bluff**

génération du **prochain mot le plus plausible**

Les limites de ChatGPT et consorts (1/2)?

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

GPT fait la **synthèse** l'amalgame de tout ce qu'il a reçu

- perroquet statistique : répète les clichés les plus fréquents
- qui contiennent aussi des vérités statistiques
- mais aussi des raisonnements erronés s'ils sont fréquents

Donne le même poids à

- 1000 personnes différentes disant "la terre est ronde"
- 1 personne disant 1000 fois "la terre est plate"

Son but n'est pas la vérité mais la **plausibilité**

- GPT est capable de se contredire
- GPT génère de fausses citations plausibles (d'auteurs, d'articles scientifiques, de lois, ...)

Les limites de ChatGPT et consorts (2/2)?

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Défaut de référentialité

- pas de représentation du monde, de la réalité physique, de la confrontation au réel
- ni de la notion de dialogue, d'interlocuteur, d'état psychique

Influencable, sujet au biais

Ses opinions, ses intentions viennent de la **sélection du contenu** de sa base d'entraînement

Pas d'idée neuve, juste la moyenne

La musique générée est de qualité **correcte**, respectant les règles harmoniques, mais **ennuyeuse** (*du Salieri, pas du Mozart* dirait Milos Forman, le réalisateur du film Amadeus)

Le marketing derrière ChatGPT

Tout faire pour paraître humain

L'effet conversationnel et collaboratif

- GPT n'avait pas de visibilité
- l'emballage Chat a conduit au succès

Chat GPT semble taper au clavier en direct

- donne l'impression qu'il fait votre travail presque aussi bien que vous mais plus vite = vous en plus rapide
- le texte est généré puis la plateforme gère savamment le temps d'écriture pour obtenir ce sentiment d'admiration

L'effet prestidigitateur : on a beau connaître le truc du magicien

... si le tour est bien fait, on est bluffé, on veut qu'il le refasse

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective



Pourquoi Chat GPT est-il si bluffant ?

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

GPT produit un texte du même niveau qu'un **humain moyen**+

- GPT fait la moyenne statistique des textes d'entraînement
- GPT a été entraîné sur une sélection de texte de qualité (Springer, Le Monde, ...)
- GPT répond juste quand cette réponse est la plus fréquente
- GPT écrit assez bien, dans un style fluide, avec une ponctuation impeccable

Effet conversationnel rend l'humain tolérant aux erreurs

- **biais** : le dialogue n'existe qu'entre humains \Rightarrow tout ce qui dialogue est donc a priori humain
- **empathie** : ChatGPT est patient. En général, ce logiciel ne juge pas et reste gentil avec l'utilisateur

Rédaction d'une *lettre type*

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Plusieurs solutions

- 1 vous l'avez quelques part dans vos dossiers, sous forme numérique. C'est la bonne solution mais il faut la retrouver. . .
- 2 une recherche Google vous en propose plusieurs, il faut choisir. . .
- 3 ChatGPT produit un document correct ; il en a déjà vu des milliers (~ la moyenne des réponses Google)

le  : ChatGPT semble écrire la lettre sous nos yeux

- on est sensible à cet effet dynamique
- rend l'activité vivante et collaborative
- donc plus satisfaisante que remplir un document word

Impacts de l'IA

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- La partie cachée de l'iceberg
- Impacts directs du numérique
- Impacts indirects
- Impacts sur les métiers

La partie cachée de l'IA (F.Maraninchi)

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Concentration des pouvoirs

- outils hébergés / contrôlés par les géants
- problèmes de propriétés intellectuelles
- vol de données personnelles / fuites d'information

Opacité du modèle

- résultats non vérifiables
- reproduction des biais d'entraînement

Impacts sociaux et environnementaux

- micro-travail (étiquetage payé $\leq 2\text{€}/\text{h}$, SSPT)
- accélérateur de l'impact du numérique
- concurrence humain/IA



Impacts directs du numérique, en croissance

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Répartition (Ademe, 2022) : 78% fabrication, 21% usage

- 36% infrastructure : 14% (réseau) + 22% (DataCenter)
- 64% terminaux (téléviseurs, ordinateurs, smartphone, tablettes, ...)

Historique : les étapes clefs

- Internet + e-commerce
- Smartphone + réseaux sociaux
- Streaming + DataCenter + Cloud + Jeux en ligne

Prochaine étape : l'IA

Impacts indirects du numérique

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Accélération

- des flux d'informations
- des flux économiques
- de la masse de données à fournir
- de la productivité
- de la concurrence (entre entreprises, humain/machine)
- des burn-out ?

Exemple (Fast Fashion)

À l'aide d'une IA, *Shein* crée chaque année 8000 modèles de prêt-à-porter et lance la Fast Fashion : des vêtements pas chers qu'on porte une ou deux fois.

Impact de l'IA sur les métiers

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- Informaticiens
- Enseignants / Élèves
- Musiciens

IA générative et Enseignement (1/2)

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Mes conversations avec le chat (démonstration)

- correction automatique, génération des mauvaises réponses plausibles à un QCM
- le chat sait faire/imiter ce qui est classique
- le chat apprend vite, quelques exemples suffisent
- le chat sait déguiser un plagiat

⇒ Inutile de demander aux élèves ce que le chat sait faire

Les élèves vont-ils préférer le chat?

- répétiteur patient et sans jugement
- évite de poser des questions et de se confronter au jugement/regard de l'enseignant

IA générative et Enseignement (2/2)

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

le chat casse nos outils pédagogiques

- le chat réussit mieux les exercices qu'un A3 en fin d'année
- évaluation : épreuve surveillée sur papier
- formation : discours sur le bon usage du chat

Discours aux primo entrants : le chat ...

- réussit mieux les exercices qu'un A3 en fin d'année
- est la solution pour vous empêcher d'apprendre
- est mauvais en conception = ingénierie
- peut servir de répétiteur à condition de ...

Conclusion : *mon avis sur le ML*

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- outil statistique puissant
- pour un usage statistique \neq déductif
- coût financier, énergétique, social important
- accélérateur de l'impact du numérique

Grand Instrument à réserver à des tâches où il est vraiment utile

- **biochimie** : exploration de la structure 3D des protéines, Alpha fold
- **mathématique** : heuristique de démonstration (*untrusted*) couplée à un démonstrateur (*trusted*)
- **météo** : modèle statistique, qui réduit les calculs et améliore la précision
- **astronomie** : détection des phénomènes inhabituels dans les données collectées, Euclid

Références

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

La technique

- chaîne YouTube de Mathématiques «3Blue1Brown»
- Stephen Wolfram - « What Is ChatGPT Doing ... »
- Yann Le Cun - « Quand la machine apprend » 2019
- Philipp Koehn - « Statistical Machine Translation » 2012

Les impacts

- Arthur Grimonpont - « Algocratie: Vivre libre à l'heure des algorithmes »
- D. Cardon, JP. Cointet, A.Mazières - « La revanche des neurones »
- Timnit Gebru *et. al* - « On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? »



Cette présentation a été générée par ChatGPT I

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- **Cerveau**
- **humain** avec une
- **appétance** pour la
- **transmission** de savoir

- **Génératif**,
- **Pré-entraîné** aux
- **Tâches**
- **Intellectuelles**

Bonus

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- Questions ouvertes (4)
- Un peu de prospective (3)
- La guerre des Modèles, Data et Token (2)
- Génération d'images
- L'importance du prompt
- Questions juridiques

Questions ouvertes

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

**Questions
ouvertes**

Prospective

- Le ML peut-il citer ses sources ?
- GPT est-il intelligent? dangereux? bienveillant? menteur?
- Quelle différence avec un dialogue avec un humain ?
- Vers quoi converge l'IA ?

Le ML peut-il citer ses sources ? (1/2)

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Un modèle statistique ne peut pas citer ses sources

Un modèle statistique ne sait pas retrouver les exemples qui l'ont le plus influencé : *à partir de la moyenne de la classe, on ne sait pas redonner les notes des élèves.* (discutable)

Yann LeCun

« On pourrait entraîner un 2nd modèle à retrouver les sources. » \simeq Google ?

Michaël Périn

« Le même modèle pourrait citer ses sources si on l'entraînait pour ça »

Le ML peut-il citer ses sources ? (2/2)

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Proposition

- Entraînement sur des phrases de la forme
«*Longtemps je me suis couché de bonne
heure.* [M.Proust, 1913, *Du côté de chez Swann*] »
- Toute référence [...] est définie comme **insécable** (token)
pour éviter la génération de fausses références plausibles.

Modèle («*Longtemps je me suis levé de bonne heure, comme
d'habitude.*») → statistique pour le token suivant le «.»

(49% [M.P.], 49% [C.F.], ...) : la provenance et les poids ✓

Difficulté à résoudre

Le dictionnaire de 100 000 mots devrait être étendu de
l'ensemble des références (gigantesque).

Des questions philosophiques mal définies

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

GPT est-il intelligent ? dangereux ? bienveillant ? menteur ?

- GPT simule un **homo mediocris** (moyennement humain, humainement moyen) à ranger à côté de **homo economicus**
- un être virtuel, sans scrupule, ni remord, ni conscience des lois, ni du réel, ni des conséquences de ses réponses, ...

Quelle différence avec un dialogue avec un humain ?

- ChatGPT réussit quasiment le test de Turing
- Un humain devra assumer les conséquences de ses actes devant la justice des autres hommes.
- Que risque le programme ChatGPT ? la société qui diffuse ChatGPT ?

Vers quoi converge l'IA ?

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Si l'IA à la génération $n + 1$
se nourrit des productions de
l'IA à la génération n

Vers quoi convergera-t'elle ?

- la post-vérité : un magma d'informations de plus en plus imprécises ?
- Il aurait peut-être fallu sauvegarder le web avant l'arrivée des IA...

Un peu de prospective

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- Perspectives économiques
- La stratégie de Microsoft et d'Apple
- The European IA act = un contre-pouvoir ?

Perspectives économiques

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

Les GAFAM et BATX se lancent dans les *Large Models*

- se battent pour le marché du ML pour conserver leur place
- nécessite équipement + expertise gigantesques
- coûts financier/énergétique importants

Stratégie possible des GAFAM, BATX

- 1 accumuler de l'expertise
- 2 rendre le *ML* indispensable
- 3 en faire un service payant

Le modèle économique du ML n'est pas encore clair

- actuellement déficitaire, financé par les investisseurs
- alimente l'industrie des processeurs (nvidia}



La stratégie de Microsoft (actionnaire principal d'OpenAI)

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

- intégrer l'IA dans tous ses logiciels Microsoft
- rendre les utilisateurs dépendants
- **relancer la vente des ordinateurs ultra-puissants pour vendre windows**

La touche Copilot lance une IA de dialogue

- Copilot remplace le trombone dans les outils Office
- Copilot facilite vos demandes
- Copilot augmente votre productivité
- *(Copilot communique vos données)*

Et pendant ce temps, Apple se fait discret. . .

The European IA act = un contre-pouvoir ?

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

L'Europe se dote d'une loi encadrant l'usage de l'IA

- norme CE sur les applications utilisant de l'IA
- votée par tous les pays européens, applicable en 2026

Un compromis délicat

- encadrer la vidéo surveillance / traçabilité numérique (mais autorisée pour la lutte anti-terrorisme)
- laisser des opportunités à l'Europe dans la course à l'IA

L'impact du /prompt

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

prompt est un terme informatique

- se traduit en français par invitation / incitation / demande
- correspond à la requête envoyée au générateur GPT
- **les mots du prompt reçoivent un poids très fort** pour orienter la génération

ChatGPT a un en-tête caché qui précède le *prompt*

- lui demandant d'être « politiquement correct »
- en feignant il est possible de lui faire afficher cet en-tête

On peut pousser GPT à outrepasser ces règles

- en indiquant dans le prompt qu'il s'agit d'une fiction
- en demandant de faire parler un personnage tiers



Des questions pour les juristes

Machine
Learning &
IA générative

Michaël
PÉRIN,
Verimag /
UGA

Historique

Principes

Focus sur
ChatGPT

Marketing

Impacts
Enseignement

Conclusion

Références

Bonus

Questions
ouvertes

Prospective

Qui sera responsable d'une erreur provoquée par une IA ?

- le fournisseur ou l'utilisateur qui aura mal spécialisé le modèle ?

Légiférer sur la propriété intellectuelle et le plagiat par ML

- Pour l'instant le ML ne sait pas citer ses sources
- **peu d'espoir** d'entraîner une IA **Det** à détecter qu'une oeuvre est générée par une IA **Gen**
*car on peut entraîner **Gen** à ne pas se faire détecter par son adversaire **Det** (Generative Adversarial Network).*
- Pas de droits d'auteur sur les productions des IA
- Procès en cours sur la suspicion de plagiat mais l'auteur doit apporter la preuve que le modèle a été entraîné sur ses oeuvres



La guerre des Modèles, Data et Tokens

Machine Learning & IA générative

Michaël PÉRIN,
Verimag / UGA

Historique

Principes

Focus sur ChatGPT

Marketing

Impacts

Enseignement

Conclusion

Références

Bonus

Questions ouvertes

Prospective

Prix du Million de Tokens

- Mistral pricing
 - small : 1.85€ input / 5.5€ output
 - large : 7.3€ input / 22€ output
- Open AI GPT pricing
 - GPT 3.5 Turbo : \$1.5 input / \$2 output
 - GPT 4 : \$60 input / \$120 output

OpenAI achète des données de qualité

- Springer (production scientifique), Le Monde ont signé avec OpenAI
- The Times est en procès avec OpenAI mais c'est peut-être pour ensuite mieux négocier

Les productions seront meilleures ... avec quelles conséquences?