



Plan de la présentation

- IA et science des textes
- Récentes avancées en TAL et LLMs
- Application à la science des textes
- Conclusion

2

IA et science des textes

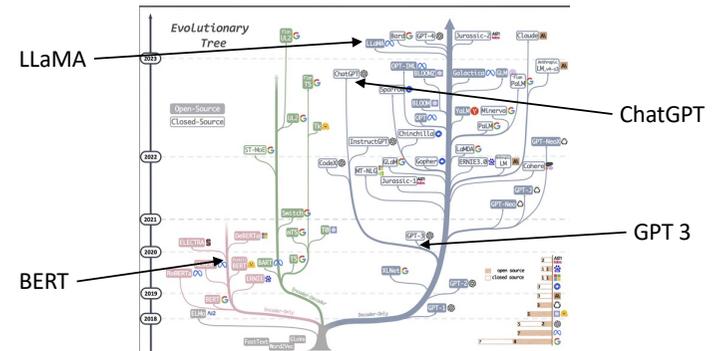
Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2



Récentes avancées en TAL et LLMs

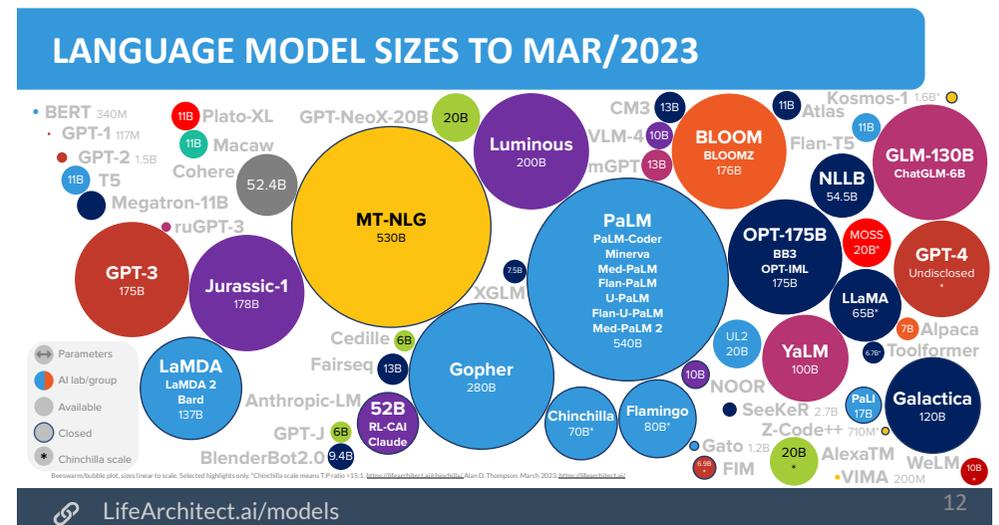
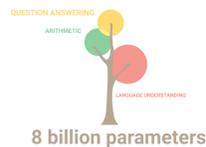
Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2

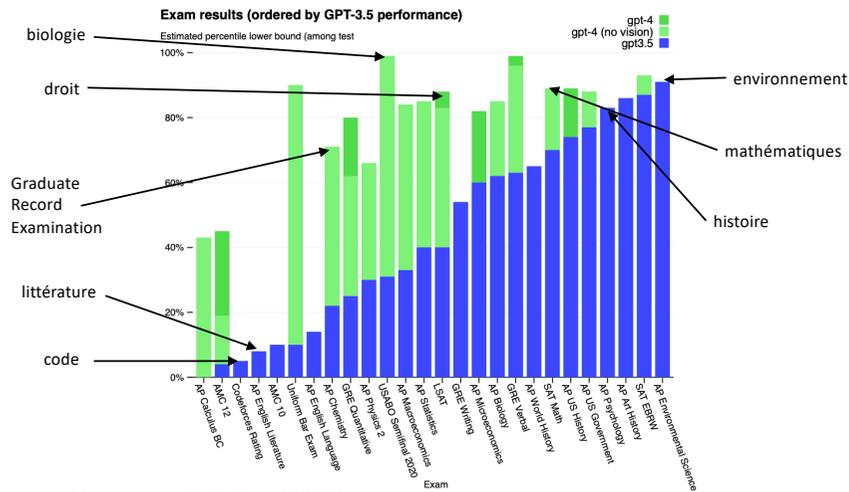
Succès des grands modèles de langue (LLMs)



Note : ce graphique est en constante évolution

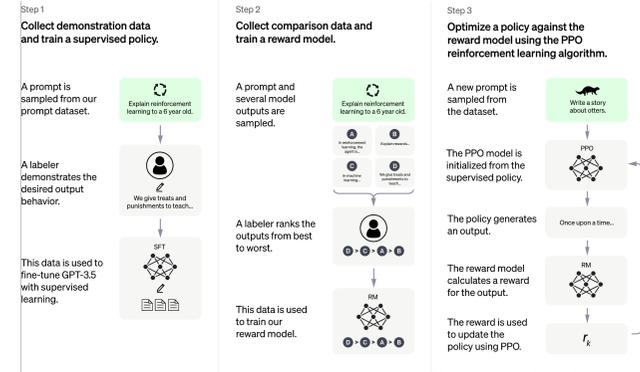
Des modèles de plus en plus larges





extrait du rapport technique de GPT-4 (OpenAI, 2023)

ChatGPT (Ouyang et al., 2022)



<https://openai.com/blog/chatgpt>

L'essor des modèles libres

- LLaMA est proposé par Meta en février 2023 (Touvron et al., 2023), disponible pour la communauté, entraînés sur données publiques, nouvelle version avec LLaMA 2 en juillet 2023
- Mistral AI, fondé par des anciens de Google DeepMind et Meta, propose un modèle libre à 7B de paramètres, Mistral-7B, qui dépasse LLaMA 2 en septembre 2023
- Ces modèles peuvent être affinés (*fine-tuned*) avec des instructions pour les utiliser comme *chatbots*
- Disponibles « facilement » sur le hub d'HuggingFace

Utiliser des LLMs aujourd'hui

Installation des bibliothèques

Initialisation d'un pipeline

LLM

recours à un GPU

```

pip install transformers>=4.0
pip install sentencepiece
import tensorflow as tf
assert tf.__version__ >= "2.0"
from transformers import pipeline
nlp = pipeline('question-answering', model='etalab-ia/camembert-base-squadFR-fquadv-piaf', tokenizer='etalab-ia/camembert-base-squadFR-fquadv-piaf')

question = "Comment s'appelle le portail open data du gouvernement ?"
context = "Etab est une administration publique française qui fait notamment office de Chief Data Officer."
answer = nlp({
    'question': question,
    'context': context
})
answer

{'score': 0.9958766102790833,
 'start': 409,
 'end': 423,
 'answer': ' data.gouv.fr.'}

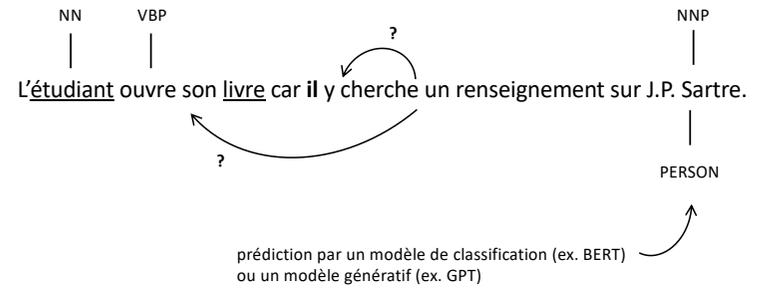
```

<https://colab.research.google.com>

Application à la science des textes

Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2

Enrichissement du texte



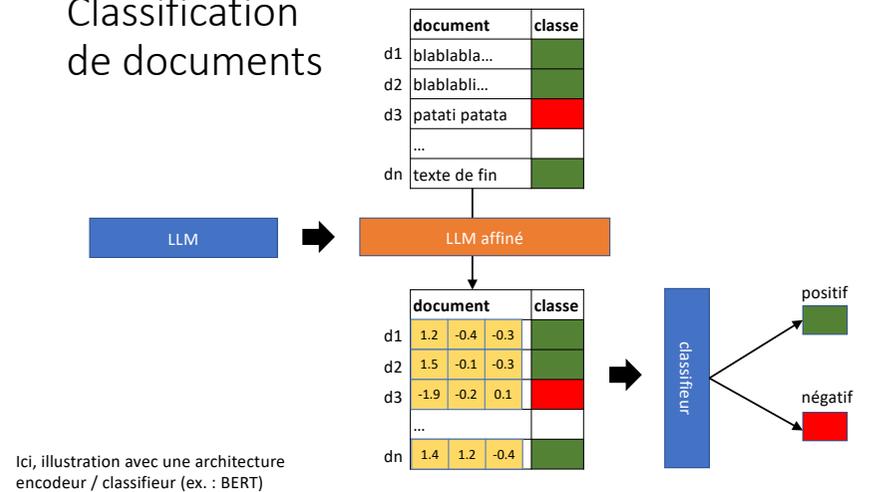
18

Recherche d'information



19

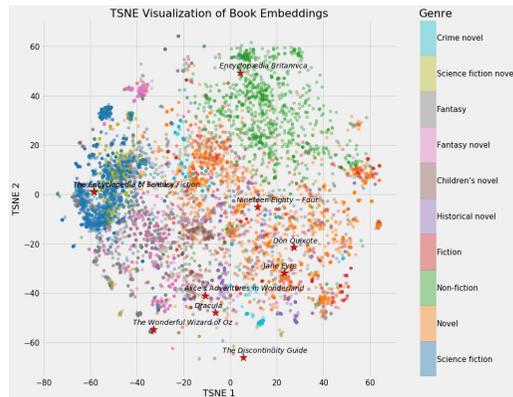
Classification de documents



20

Plongement de textes et représentations

- La plupart des applications nécessitent de « plonger » (*embed*) les textes dans des espaces vectoriels. Ces vecteurs sont une **représentation** qui vise à capturer la sémantique des textes.

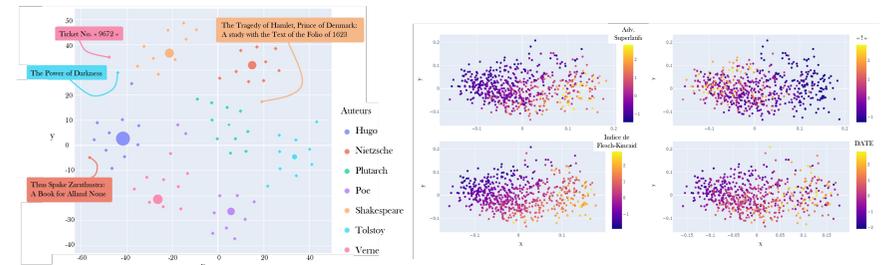


<https://towardsdatascience.com/neural-network-embeddings-explained-4d028e6f0526>

21

Application à l'analyse du style littéraire

- Projet ANR LIFRANUM, thèse d'E. Terreau
<https://marge.univ-lyon3.fr/projet-lifranum>



22

Identification et réduction des biais

- Projet ANR DIKé, thèse d'I. Proskurina sur les biais dans les modèles réduits (ou compressés).
- Pour donner un exemple simple :

```
fill_masker("The editor stopped the driver and asked [MASK] for a ride")
```

[MASK] Bert Predictions with confidence:

- 90% confidence: [MASK] = him
- 10% confidence: [MASK] = her

Language Models can predict HIM instead of HER given the context and [MASK]

23

Conclusion

Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2

Quelques conclusions personnelles

- Beaucoup de ressources « prêtes à l'emploi »
- La plupart sont issues des travaux en TAL et en machine learning
- Nécessité de travailler avec des experts pour l'évaluation et l'affinage des modèles (mais attention aux coûts, aux biais...)
- Difficulté de trouver des problématiques de recherche en informatique (ou mathématiques)
- Toujours un chaînon manquant entre les développeurs de modèles et d'algorithmes et les chercheurs/utilisateurs en SHS

25

Références

- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Prépublié en 2018 mais finalement accepté à NAACL 2019 (~80k citations)
- Ouyang, Long, et al. "Training language models to follow instructions with human feedback." NeurIPS 2022 (~2300 citations)
- Irina Proskurina, Guillaume Metzler, Julien Velcin: The Other Side of Compression: Measuring Bias in Pruned Transformers. IDA 2023
- Irina Proskurina, Guillaume Metzler, Julien Velcin: nouveau papier accepté au BabyLM Challenge, ACL 2023
- Enzo Terreau, Antoine Gourru, and Julien Velcin. Writing Style Author Embedding Evaluation. Workshop Evaluation and Comparison of NLP Systems, co-situé avec EMNLP 2021
- Touvron, Hugo, et al. "Llama: Open and efficient foundation language models." arXiv preprint arXiv:2302.13971, 2023, ~1300 citations

26