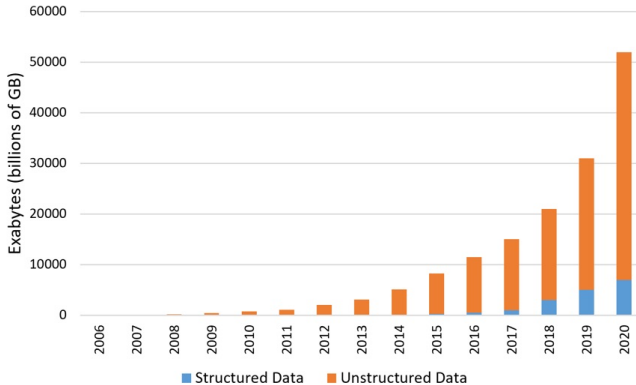# Big data or frugal AI: where can optimization techniques help?

DSAIDIS optimization and neural networks workshop
15 November 2022

Big data

The Cambrian Explosion...of Data

The tradeoffs of large scale learning

Study 2 famous algorithms in the big data regime
[Bottou & Bousquet, 2007]

estimation error	$\mathcal{E}_{\text{est}} \leq c\sqrt{\frac{d}{N}}$	
	gradient descent	stochastic gradient
step size	$\gamma = 1/L$	$\gamma = \frac{a}{\sqrt{k}}$
optimization error after k iterations	$\mathcal{E}_{\text{opt}} \leq \frac{C_1}{k}$	$\mathcal{E}_{\text{opt}} \leq \frac{C_2}{\sqrt{k}}$
cost for 1 iteration	Nd	d
total cost for $\mathcal{E}_{\text{opt}} \approx \mathcal{E}_{\text{est}}$	$C_3 Nd\sqrt{\frac{N}{d}}$	$C_4 d(\sqrt{\frac{N}{d}})^2$

$O(N\sqrt{N})$ vs $O(N)$: a constant number of passes over the data is enough
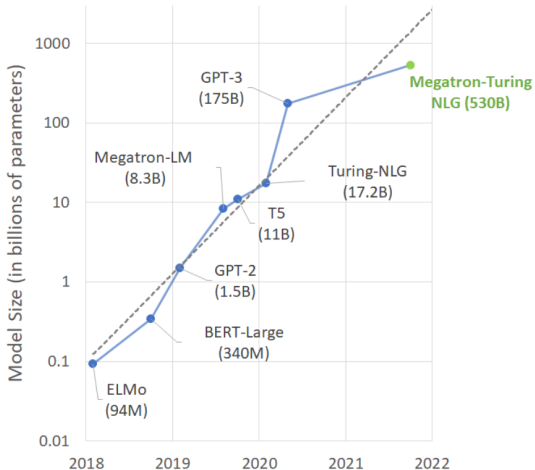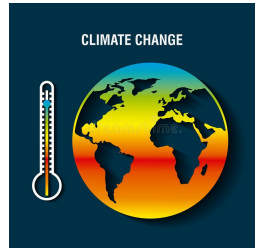
Big models

Figure 1: Trend of sizes of state-of-the-art NLP models with time.

The computational cost of artificial intelligence

Big data $\times$ Big models $=$

Components of frugal artificial intelligence

- Evaluation metrics
    - Measurement methodology $+$ shared tools
    - Operational metrics (can serve as objective functions)
- Data-efficiency
- Frugality in Computation and Memory
    - A posteriori frugality
    - Frugality by Design
    - Efficient hardware
- Frugality as one criterion of Responsible AI
    - Combine & satisfy several criteria (robustness, fairness, privacy, explainability)