Mnacho Echenim
small evolutions : Patrick Reignier

Grenoble INP-Ensimag

2025-2026

Outline

Foreword

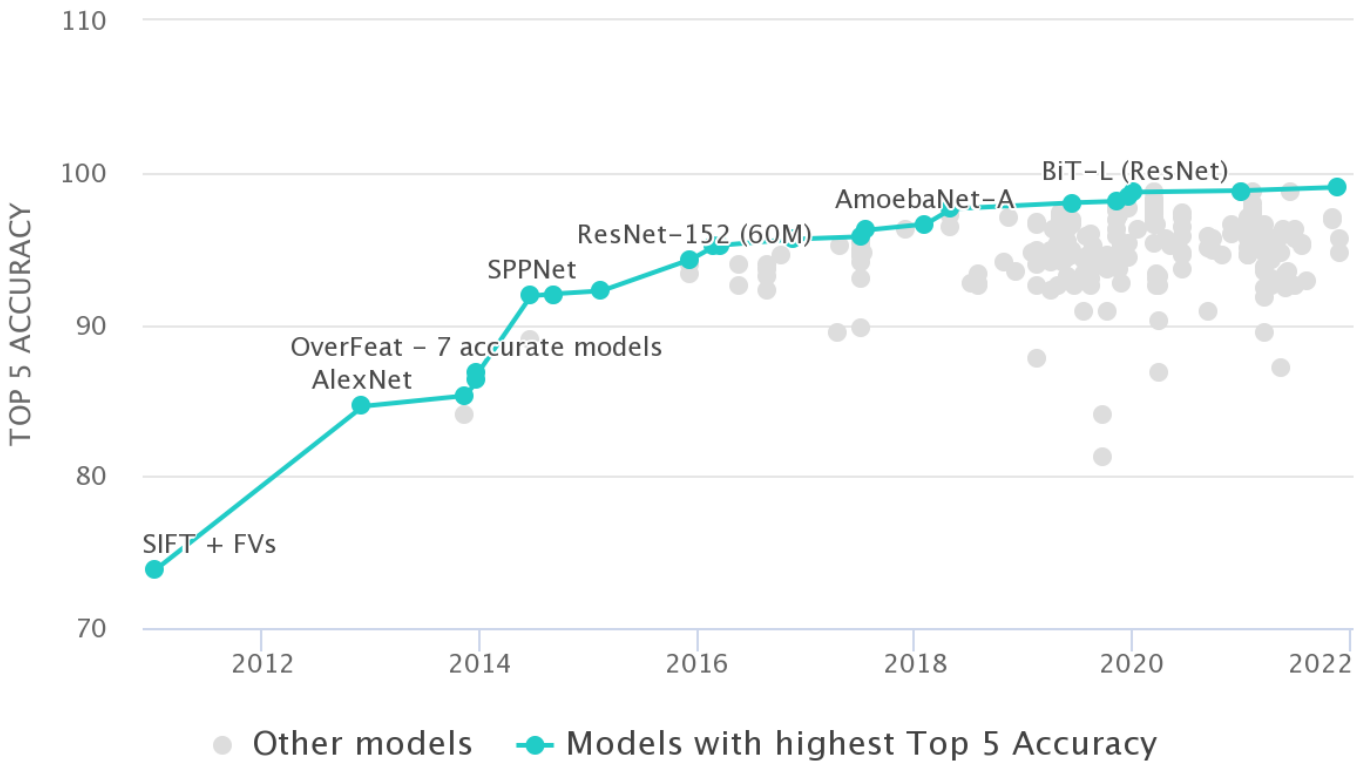
- There is an increasing concern on the environmental impact of deep learning
- Hardware concerns will not be addressed here
  - `https://bouveres.pages.ensimag.fr/jied/ressources/2021-2022/20211109-ImpactNumerique.pdf` (in french)
- We will focus on energy consumption

A focus on accuracy

- Image recognition: improve top N error
  - Principle: the correct category should be among the N guesses with the highest priority

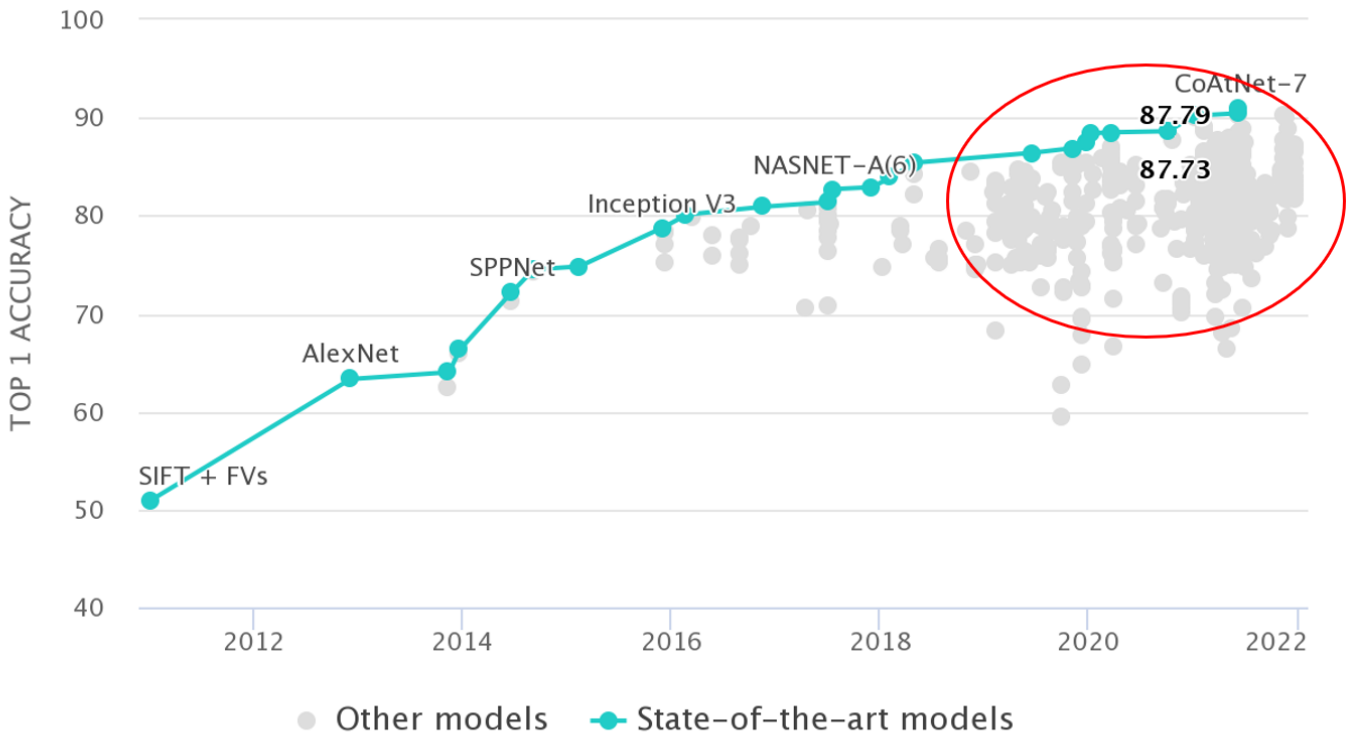
A focus on accuracy

- Image recognition: improve top N error
  - Principle: the correct category should be among the N guesses with the highest priority
- Results in December 2021 (`https://paperswithcode.com/sota/image-classification-on-imagenet`):



A focus on accuracy (2)

- Still a lot of ongoing research



The requirements of an improved accuracy

A better accuracy is obtained using more complex models

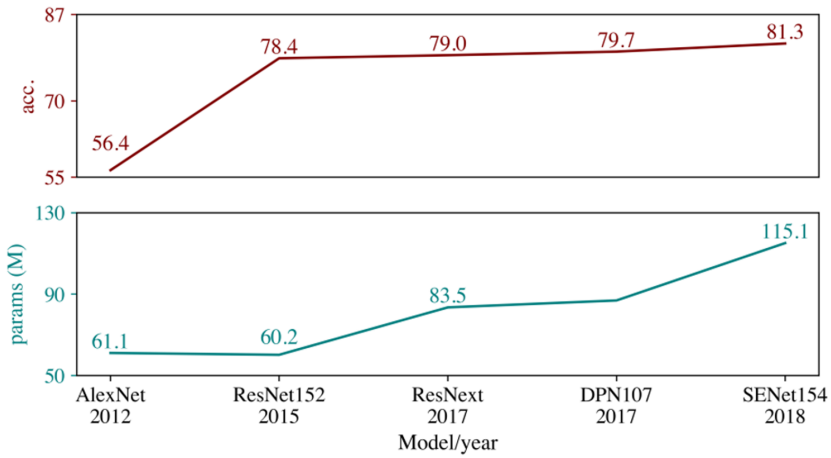
- Models with billions of parameters...
  - Up to 40 billion in NLP models or LLMs

The requirements of an improved accuracy

A better accuracy is obtained using more complex models

- Models with billions of parameters...
  - Up to 40 billion in NLP models or LLMs

Source: Green IA (https://arxiv.org/pdf/1907.10597.pdf)

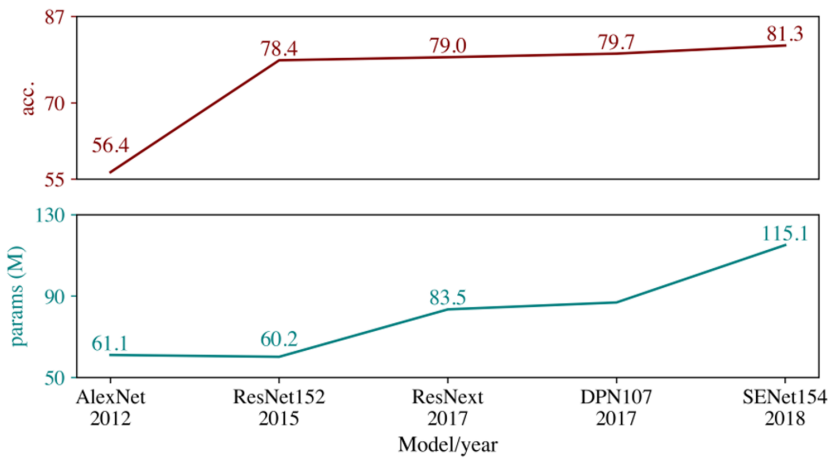


The requirements of an improved accuracy

A better accuracy is obtained using more complex models

- Models with billions of parameters...
  - Up to 40 billion in NLP models or LLMs
- ... that require larger training sets
  - 3.5 billion images from Instagram for a model trained by Facebook

Source: Green IA (https://arxiv.org/pdf/1907.10597.pdf)

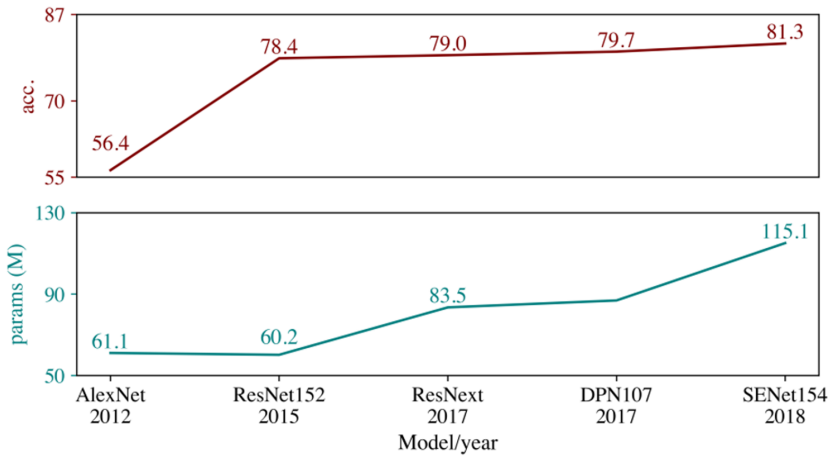


The requirements of an improved accuracy

A better accuracy is obtained using more complex models

- Models with billions of parameters...
  - Up to 40 billion in NLP models or LLMs
- ... that require larger training sets
  - 3.5 billion images from Instagram for a model trained by Facebook
- Models require hyperparameter tuning
  - Sometimes thousands of hyperparameters are tested

Source: Green IA (https://arxiv.org/pdf/1907.10597.pdf)

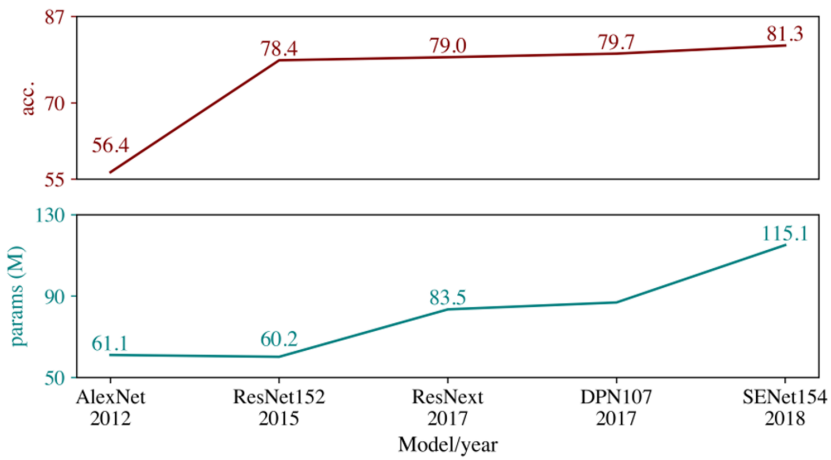


The requirements of an improved accuracy

A better accuracy is obtained using more complex models

- Models with billions of parameters...
  - Up to 40 billion in NLP models or LLMs
- ... that require larger training sets
  - 3.5 billion images from Instagram for a model trained by Facebook
- Models require hyperparameter tuning
  - Sometimes thousands of hyperparameters are tested
- This has to be done for each tested network architecture

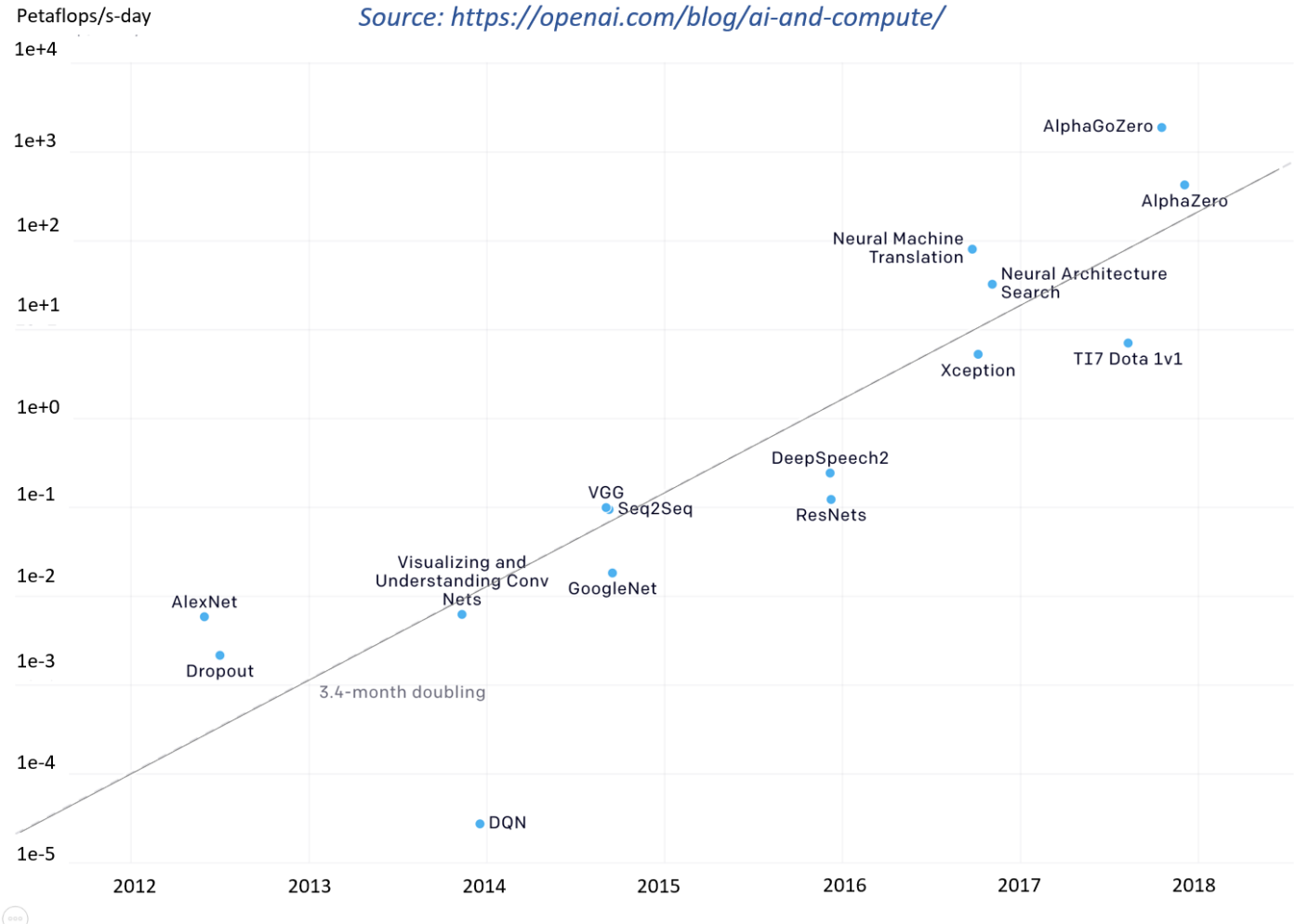
Source: Green IA (https://arxiv.org/pdf/1907.10597.pdf)



The computational cost of an improved accuracy

The search for more accurate models has entailed an **exponential growth** in **computational costs**

Source: https://openai.com/blog/ai-and-compute/



Towards efficiency

Suggestion (`https://arxiv.org/pdf/1907.10597.pdf`): we should distinguish Red AI and Green AI

- Goal of Red AI: improve the state of the art
  - Most common research domain
  - Little to no consideration for environmental impact

Towards efficiency

Suggestion (`https://arxiv.org/pdf/1907.10597.pdf`): we should distinguish Red AI and Green AI

- Goal of Red AI: improve the state of the art
  - Most common research domain
  - Little to no consideration for environmental impact
- Goal of Green AI: obtain new results without increasing computational costs
  - Ideally reducing computational costs
  - Improve global efficiency (model size, training effort, necessary amount of training data...)

Towards efficiency

Suggestion (`https://arxiv.org/pdf/1907.10597.pdf`): we should distinguish Red AI and Green AI

- Goal of Red AI: improve the state of the art
  - Most common research domain
  - Little to no consideration for environmental impact
- Goal of Green AI: obtain new results without increasing computational costs
  - Ideally reducing computational costs
  - Improve global efficiency (model size, training effort, necessary amount of training data...)

How can we evolve towards more Green AI?

- By fully reporting efforts to obtain a model
  - Number of epochs
  - Number of hyperparameters to tune...

Towards efficiency

Suggestion (`https://arxiv.org/pdf/1907.10597.pdf`): we should distinguish Red AI and Green AI

- Goal of Red AI: improve the state of the art
  - Most common research domain
  - Little to no consideration for environmental impact
- Goal of Green AI: obtain new results without increasing computational costs
  - Ideally reducing computational costs
  - Improve global efficiency (model size, training effort, necessary amount of training data...)

How can we evolve towards more Green AI?

- By fully reporting efforts to obtain a model
  - Number of epochs
  - Number of hyperparameters to tune...
- By making the trained models publicly available

Transfer learning

- Some data sets have features that are reusable on other data sources

Transfer learning

- Some data sets have features that are reusable on other data sources
  - Text words
  - Images

Transfer learning

- Some data sets have features that are reusable on other data sources
  - Text words
  - Images
- Principle: use a pretrained model on another data set as a starting point

Transfer learning

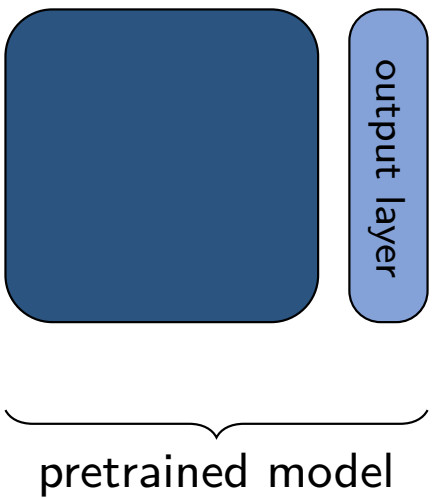
- Some data sets have features that are reusable on other data sources
  - Text words
  - Images
- Principle: use a pretrained model on another data set as a starting point
  - No longer necessary to train a model from scratch

Transfer learning

- Some data sets have features that are reusable on other data sources
  - Text words
  - Images
- Principle: use a pretrained model on another data set as a starting point
  - No longer necessary to train a model from scratch
  - *Zoo of pretrained PyTorch models*: `https://pytorch.org/serve/model_zoo.html`

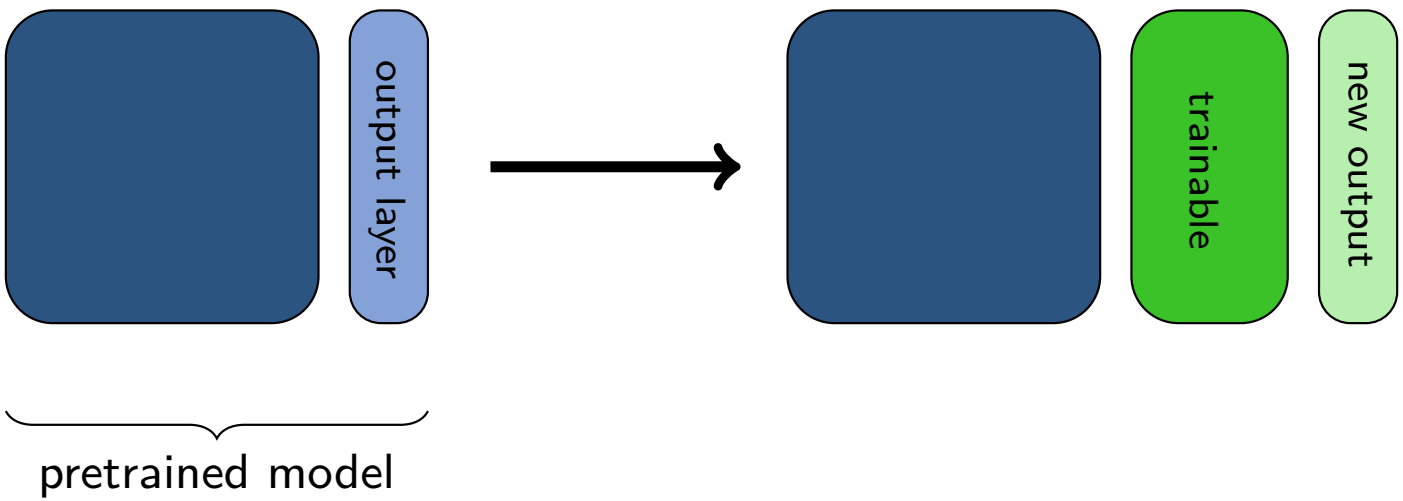
Transfer learning

- Some data sets have features that are reusable on other data sources
  - Text words
  - Images
- Principle: use a pretrained model on another data set as a starting point
  - No longer necessary to train a model from scratch
  - *Zoo of pretrained PyTorch models*: `https://pytorch.org/serve/model_zoo.html`



Transfer learning

- Some data sets have features that are reusable on other data sources
  - Text words
  - Images
- Principle: use a pretrained model on another data set as a starting point
  - No longer necessary to train a model from scratch
  - *Zoo of pretrained PyTorch models*: `https://pytorch.org/serve/model_zoo.html`



Conclusion

- This is a new topic, with questions that don't have a clear answer yet
  - What is the best measure of efficiency to use?
  - Is there even a single best measure of efficiency?

Conclusion

- This is a new topic, with questions that don't have a clear answer yet
  - What is the best measure of efficiency to use?
  - Is there even a single best measure of efficiency?
- It is a topic to be aware of
  - Especially during internships on applications of Deep Learning

Conclusion

- This is a new topic, with questions that don't have a clear answer yet
  - What is the best measure of efficiency to use?
  - Is there even a single best measure of efficiency?
- It is a topic to be aware of
  - Especially during internships on applications of Deep Learning
- Additional resources
  - Lecture notes by Anne-Laure Ligozat (in french): `https://ecoinfo.cnrs.fr/wp-content/uploads/2021/04/Cours_CentraleSupelec_IA_responsable-1.pdf`

Outline