

Reinforcement with Deep Learning

a Discussion

Guillaume.Lozenquez

@imt-nord-europe.fr



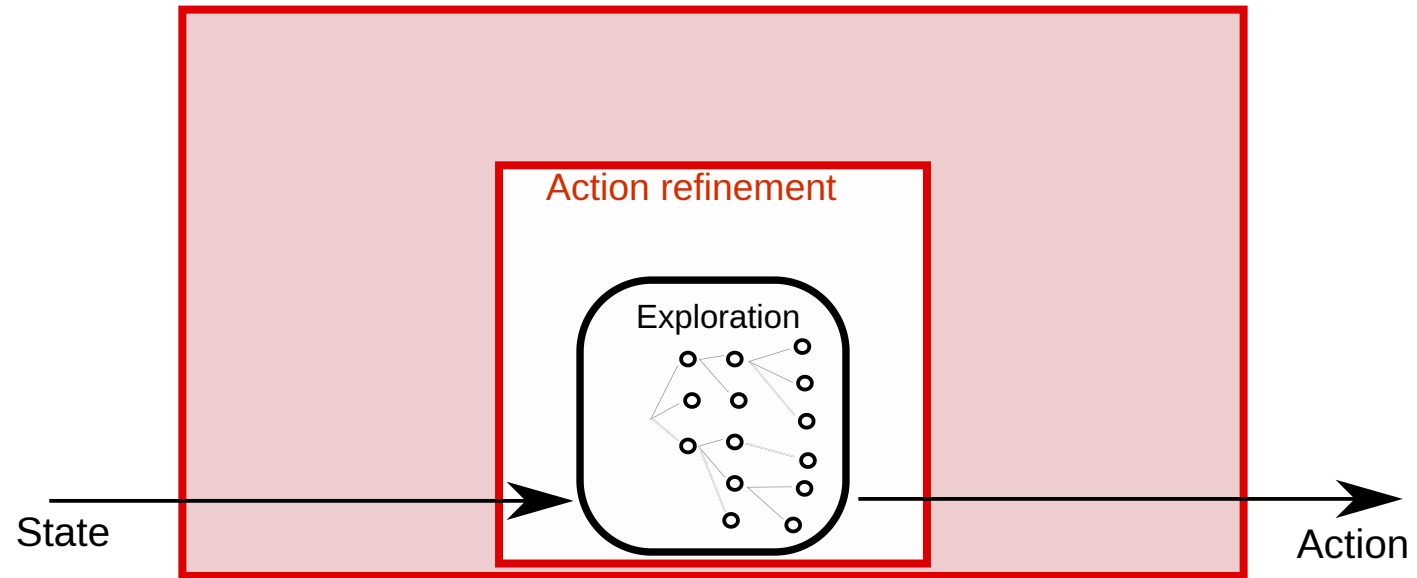
IMT Nord Europe

École Mines-Télécom

IMT-Université de Lille

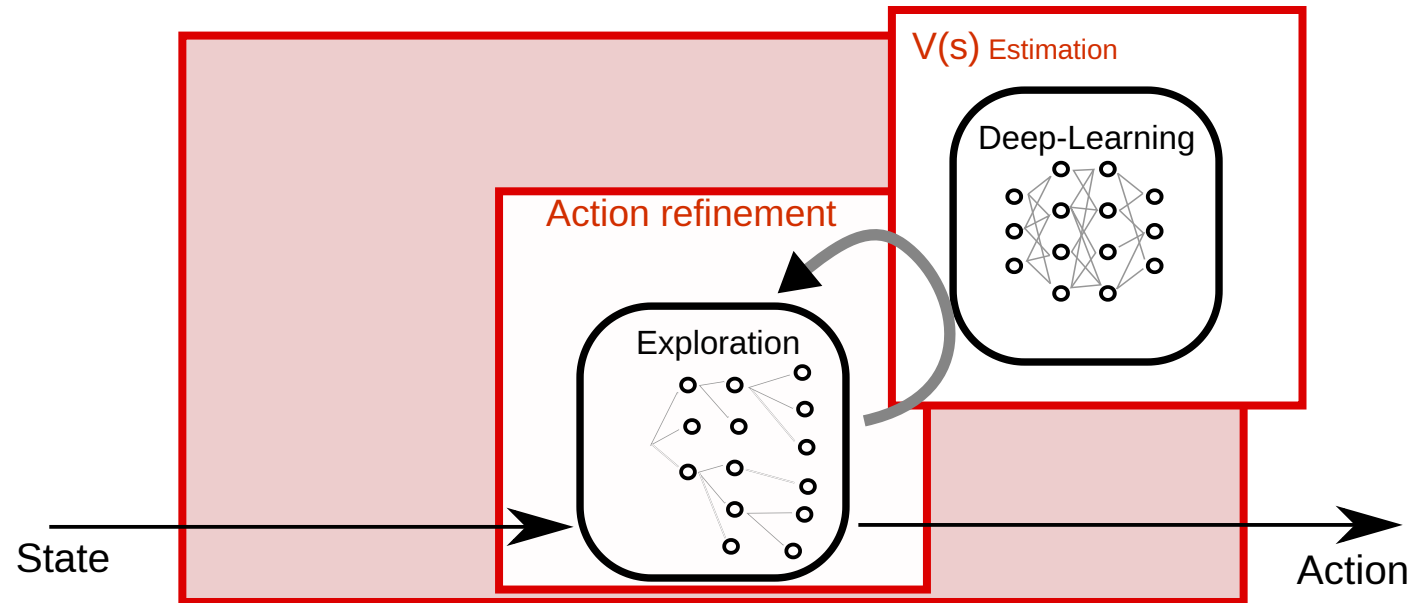
Basic approach - Random Exploration Tree

Explore locally possible evolution from the current state.



Basic approach - Incorporate Neural Networks

Incorporate Intuition as Deep Learning based Estimation



- Cut exploration branches based on a heuristic value function

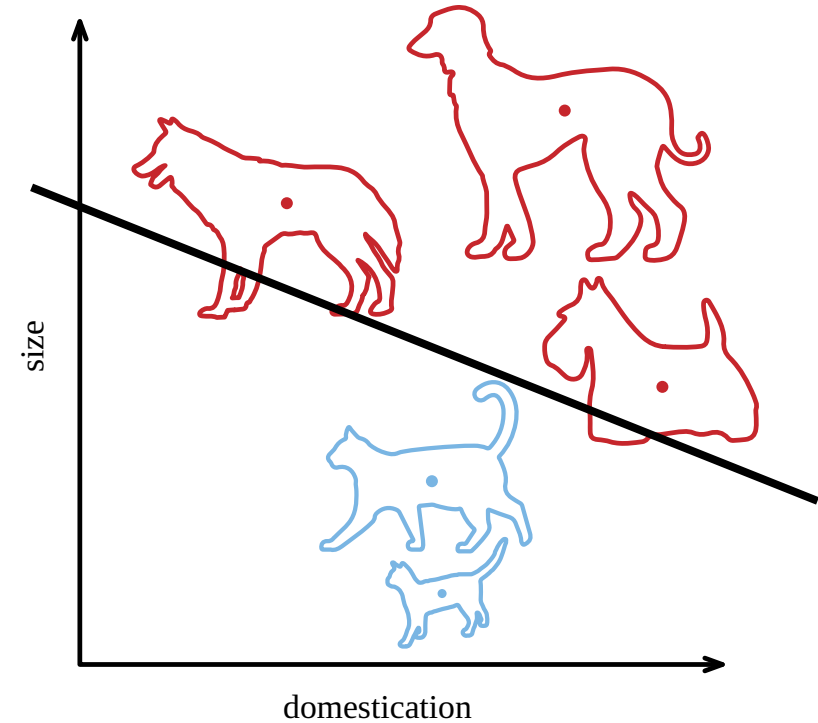
Le Deep-Learning: un cerveau à un neurone

Le perceptron:

$$si : (p_1 \cdot x + p_2 \cdot y > b)$$

Les paramètres:

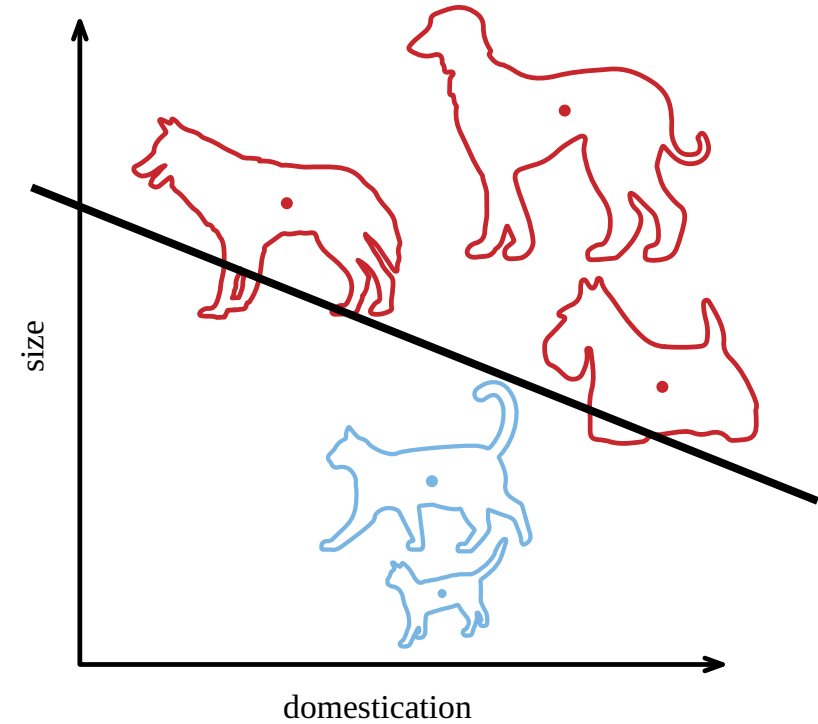
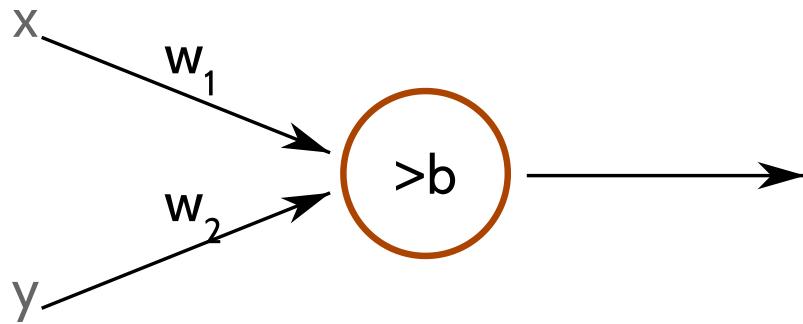
- ▶ les poids: p_1, p_2
- ▶ un seuil: b



Le grand problème des data-sciences: Classifier

Le Deep-Learning: un cerveau à un neurone

Le perceptron:



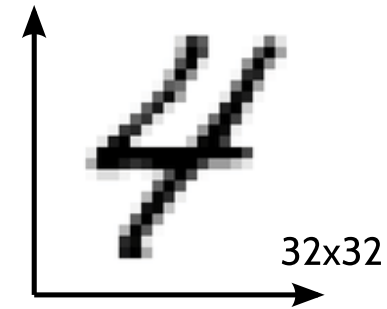
Le grand problème des data-sciences: Classifier

Le Deep-Learning: des réseaux de neurones profonds

► La rétropropagation:

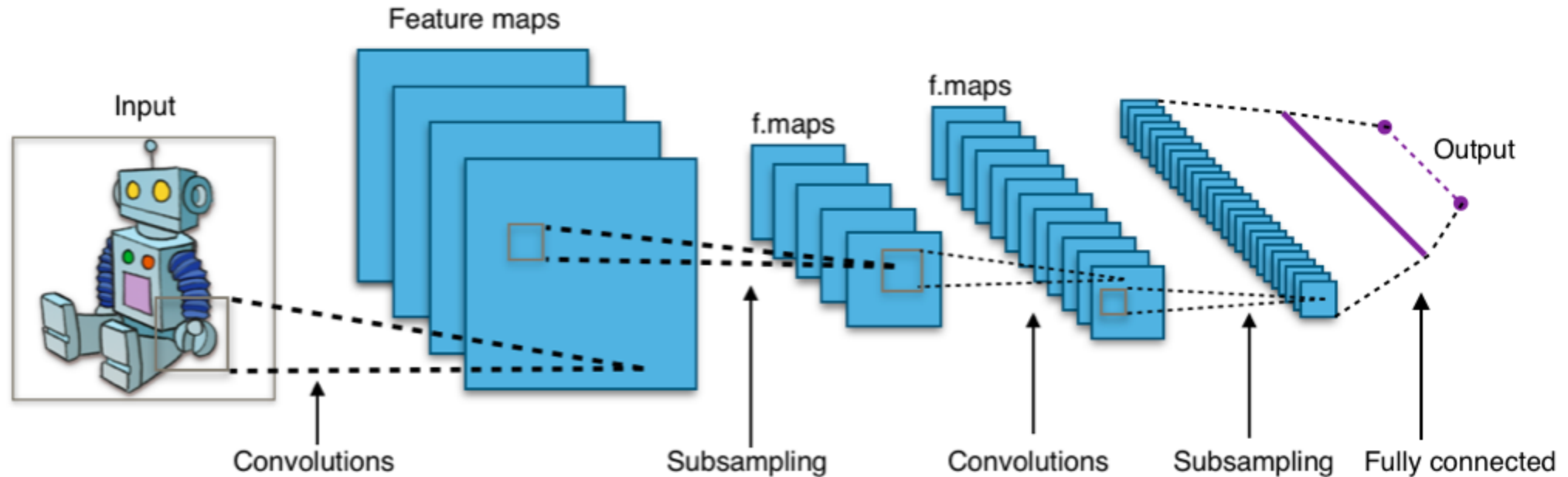
LeNet (1989-1998) *Yan LeCun et al.*

Livre: *Quand la machine Apprend*



Le Deep-Learning: des réseaux de neurones profonds

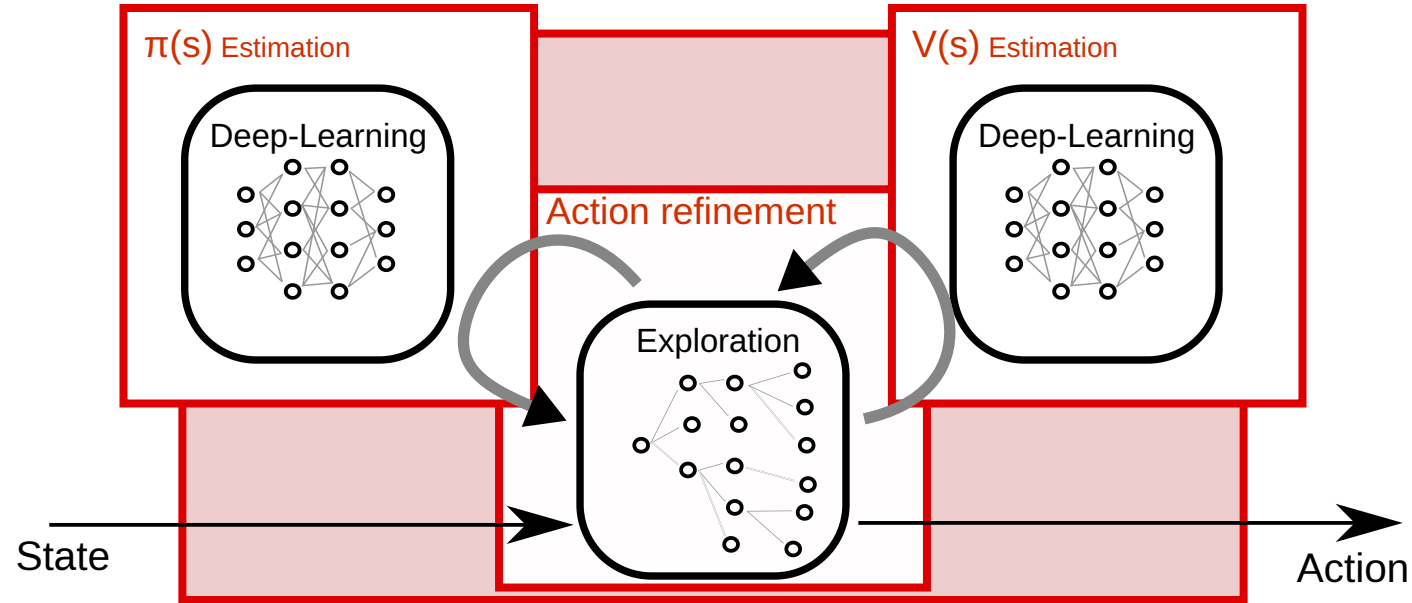
Une démocratisation de l'usage des CNN



La classification: pour n'importe qui, avec des donnés.

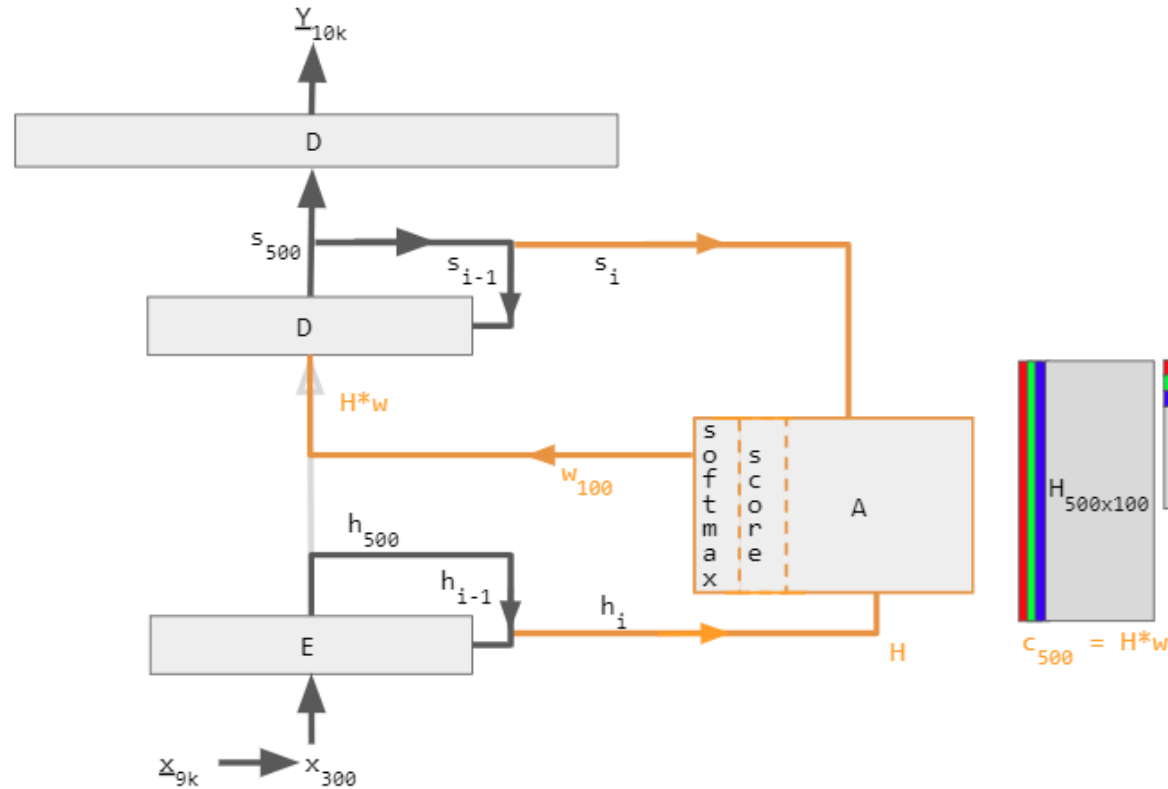
La "killing Architecture" - AlphaZero

Use both policy and value estimation.



1. Iterative process: (Play phase // Analysis phase)
2. Suppose a model of the systems (game)

Le Deep-Learning: IA générative... (Transformer)



Processus *Iteratif* couplé à un mécanisme d'*Attention* (poids dans le contexte)

IA générative: du texte à l'image

- ▶ Génération d'image sur la base d'un *prompt*.
- ▶ Un astronaute montant un cheval.

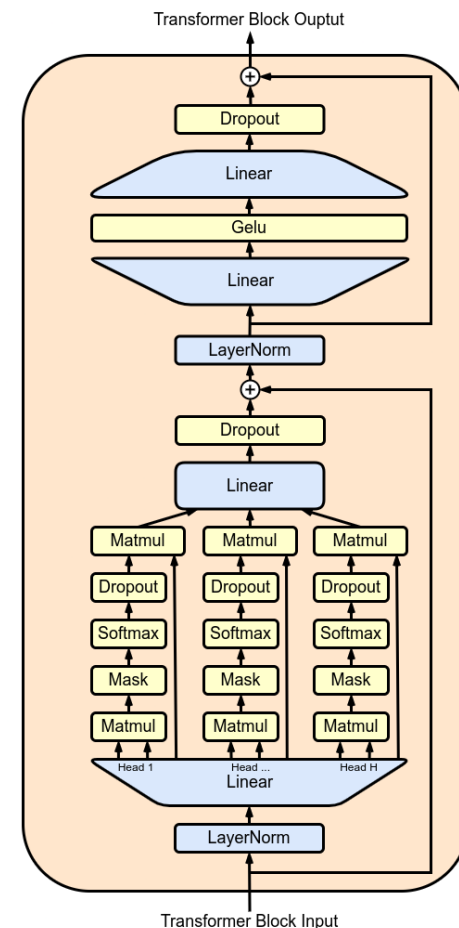
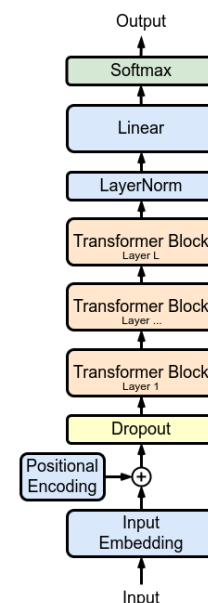
Generated by: [Stable Diffusion](#)



Le Deep-Learning: et bien sûr ChatGPT

Chat GPT-3 (nov. 2022):

- **Objectif:** (re)produire du texte
- *175 milliards* de paramètres
(avec une augmentation itérative)
- *Les données:* le web (80%), des livres (16%) et Wikipédia (3%)
- **Résultat:** une réponse par association de mots (le + disant) bien construite



Demian ?

Au cours de bourse actuel:

- ▶ **Classification** - coût (*financié/environnemental/societal*) négligeable
 - Massification: partout et pour n'importe quoi.
 - Principalement du Deep (si données en masse).
 - Alternative: Decision Tree (Gradient Boost) sur des niches.
- ▶ **Prise de decision sous incertitude**
 - Nouri par des outils de classification
 - Model: Exploration guidée
 - Sans-Model: Mechanisme d'attention
- ▶ **Apprendre le modèle ?**