

# Introduction à l'Intelligence Artificielle

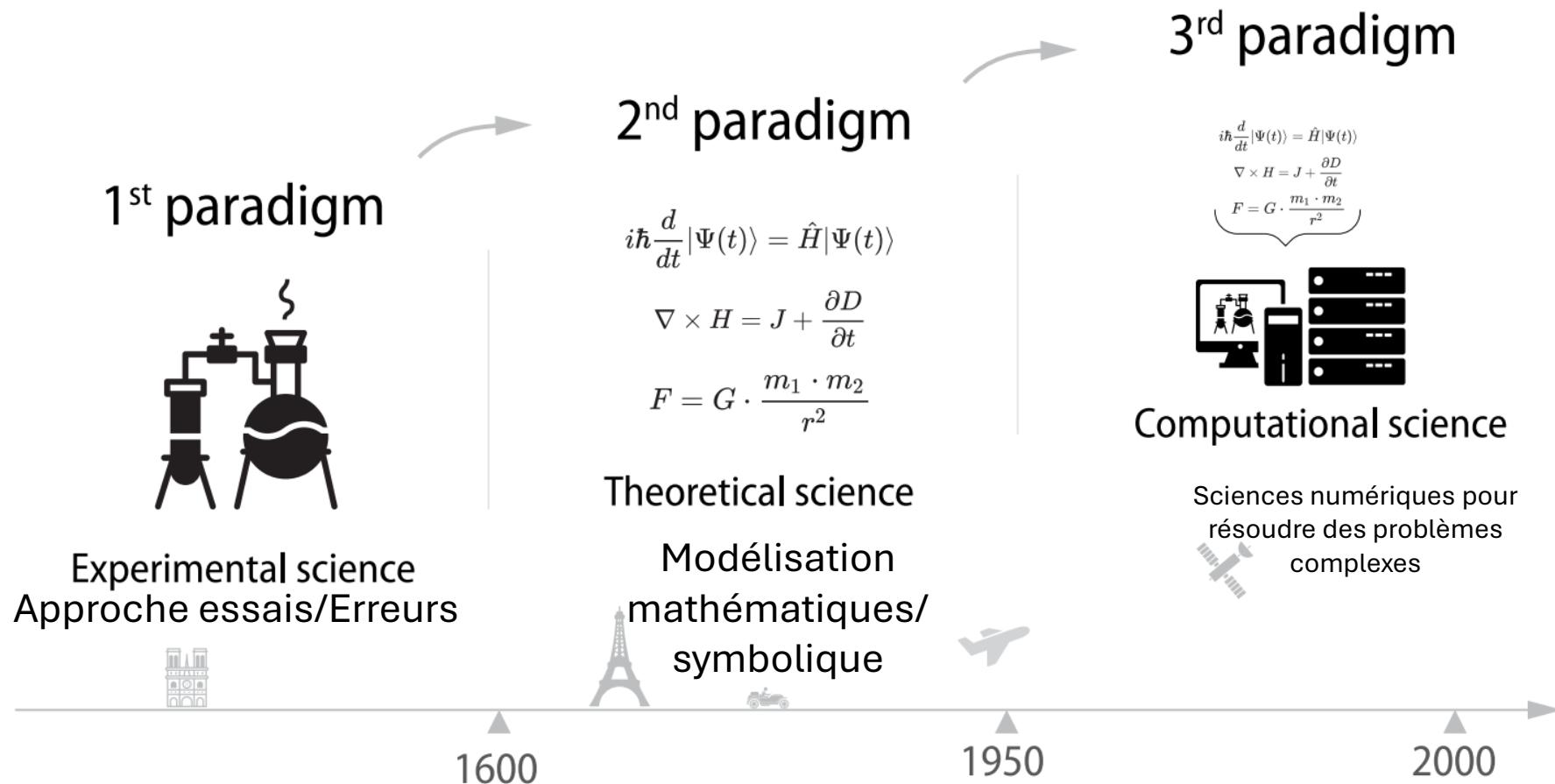
## Bases, concepts et histoire

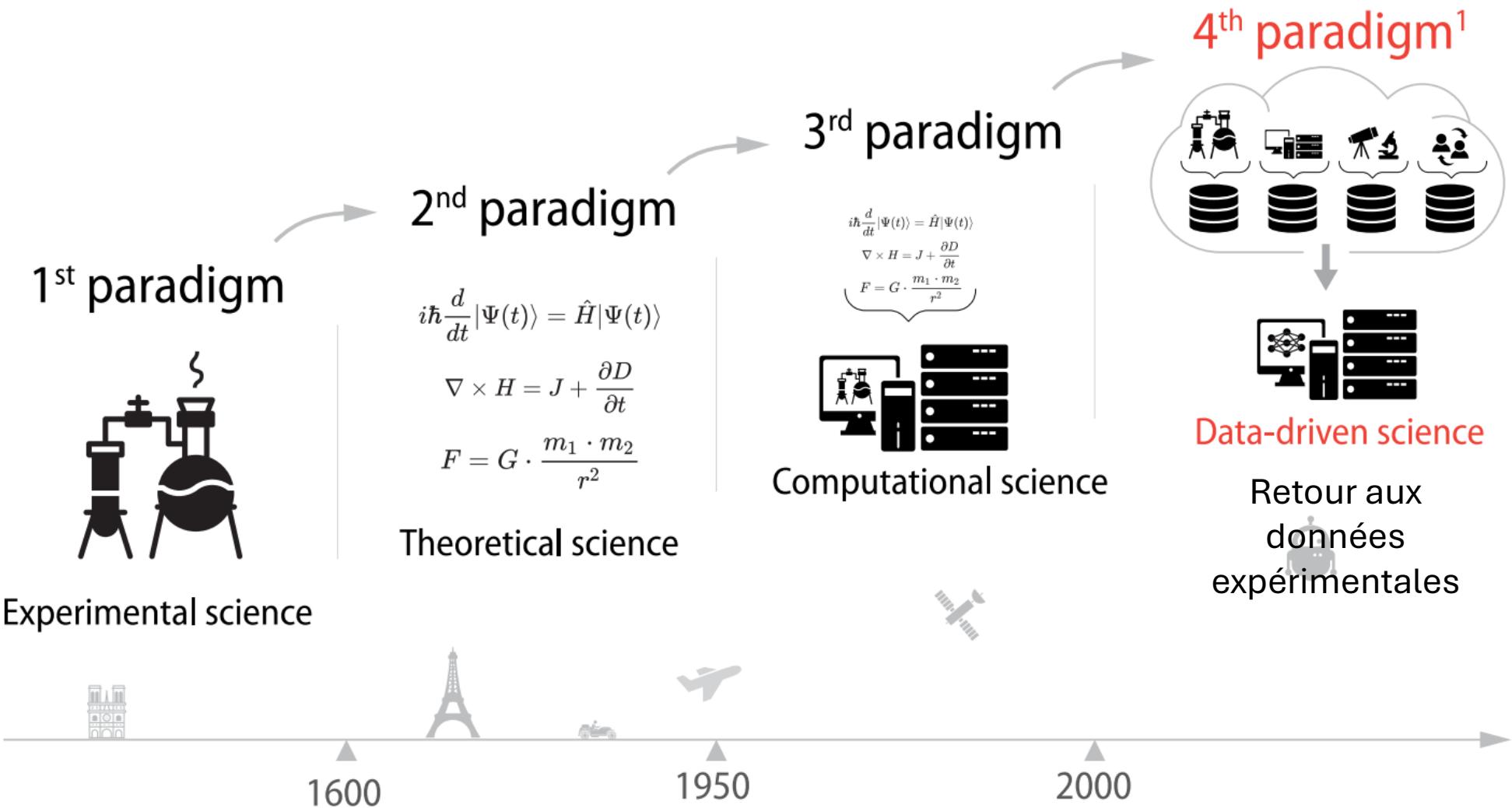


## • C'est quoi l'IA pour vous ?

- Tour de table rapide : "Citez une IA que vous connaissez ou utilisez au quotidien." (Ex: GPS, Siri/Gemini, Netflix, ChatGPT, Midjourney...).
- **Discussion** : Qu'est-ce que ces outils ont en commun ? Quelles sont vos premières impressions (enthousiasme, méfiance) ?
- **Et pourquoi faire ?**

# Comment appréhender un problème ?







# Une brève histoire de l'IA



Années 1830 : Le Rêve Mécanique

Charles Babbage conçoit la Machine Analytique. Ada Lovelace écrit le premier algorithme.

**Symboles  
Règles explicites**

Systèmes symboliques  
(1950 -> 1980)

**Données**

Systèmes statistiques  
et apprentissage  
automatique  
(1980 -> 2010)

**Apprentissage profond**  
(2010 -> auj.)

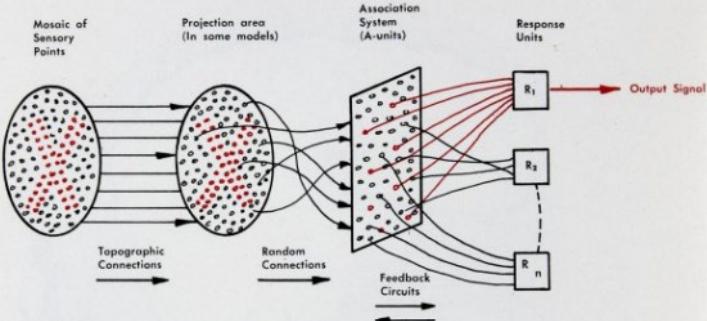
**Perception  
Interaction**



# Un Voyage dans le Temps

Découvrez les moments clés qui ont façonné l'Intelligence Artificielle.

**FIG. 1 — Organization of a biological brain. (Red areas indicate active cells, responding to the letter X.)**



**FIG. 2 — Organization of a perceptron.**

## L'Âge d'Or et le 1er Hiver (1956-1980)

### L'Enthusiasme et la Douche Froide

Les premières IA comme \*\*ELIZA\*\* (1966) et le \*\*Perceptron\*\* (1957) suscitent un immense espoir. On distingue l'approche \*\*connexionniste\*\* (apprendre par l'exemple) et \*\*symbolique\*\* (programmer des règles).

Le \*\*"Premier Hiver de l'IA"\*\* arrive : les promesses sont trop grandes, la puissance de calcul trop faible, les limites techniques (ex: incapacité du Perceptron à résoudre des problèmes simples comme le XOR) et des rapports critiques (Rapport Lighthill, 1973) coupent les financements.



1



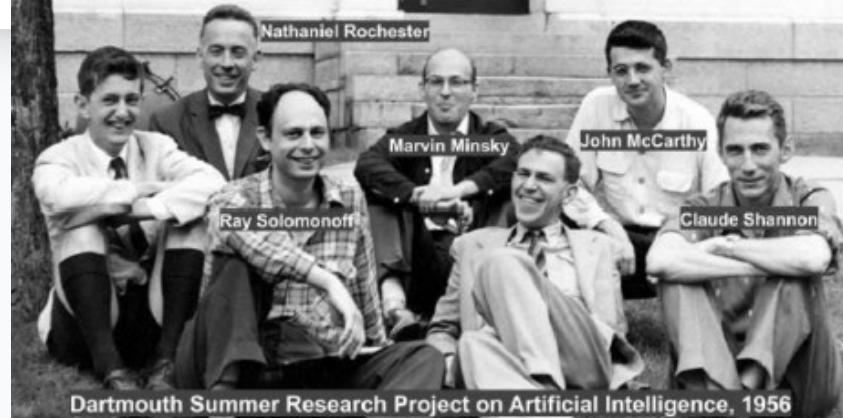
## La Genèse (1940-1956)

### Le Rêve des Machines Pensantes

**Alan Turing (1950)** : Pose la question fondatrice "Les machines peuvent-elles penser ?" et propose le célèbre Test de Turing.

**McCulloch & Pitts (1943)** : Créent le premier modèle mathématique de neurone artificiel, fondant l'idée que la pensée peut être modélisée comme un calcul logique.

**Dartmouth Workshop (1956)** : Naissance officielle du terme "Intelligence Artificielle" avec l'ambition de simuler l'intelligence humaine.



## L'Ère du Deep Learning (1993-Aujourd'hui)

### La Révolution des Données et du Calcul

La convergence de 3 facteurs déclenche la révolution :

1. **Algorithmes performants :** En 2012, **AlexNet** (un CNN) gagne la compétition ImageNet. Le **Deep Learning** devient la norme.
2. **Big Data :** Internet fournit une quantité massive de données pour l'entraînement.
3. **Puissance de calcul :** Les **GPU** des jeux vidéo sont détournés pour accélérer les calculs de l'IA.

3

## Le Printemps et le 2ème Hiver (1980-1993)

### Le Retour des Experts

L'IA symbolique connaît son heure de gloire avec les **Systèmes Experts** (ex: MYCIN), efficaces dans des domaines très spécialisés.

En parallèle, l'approche connexioniste fait un bond en avant avec la popularisation de la **rétropropagation du gradient** (1986), la clé manquante pour entraîner des réseaux de neurones profonds.

Le **"Deuxième Hiver de l'IA"** survient car les systèmes experts sont trop rigides et coûteux à maintenir.

4

[intelligence]



# D'autres formes d'intelligence



## Science

### Social signal learning of the waggle dance in honey bees

Science, 9 mars 2023

<https://www.science.org/doi/10.1126/science.adc1702>

### Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions

Science, 30 novembre 2018

<https://doi.org/10.1126/science.aat1590>

Apprentissage du signal social de la danse frétilante chez les abeilles domestiques.

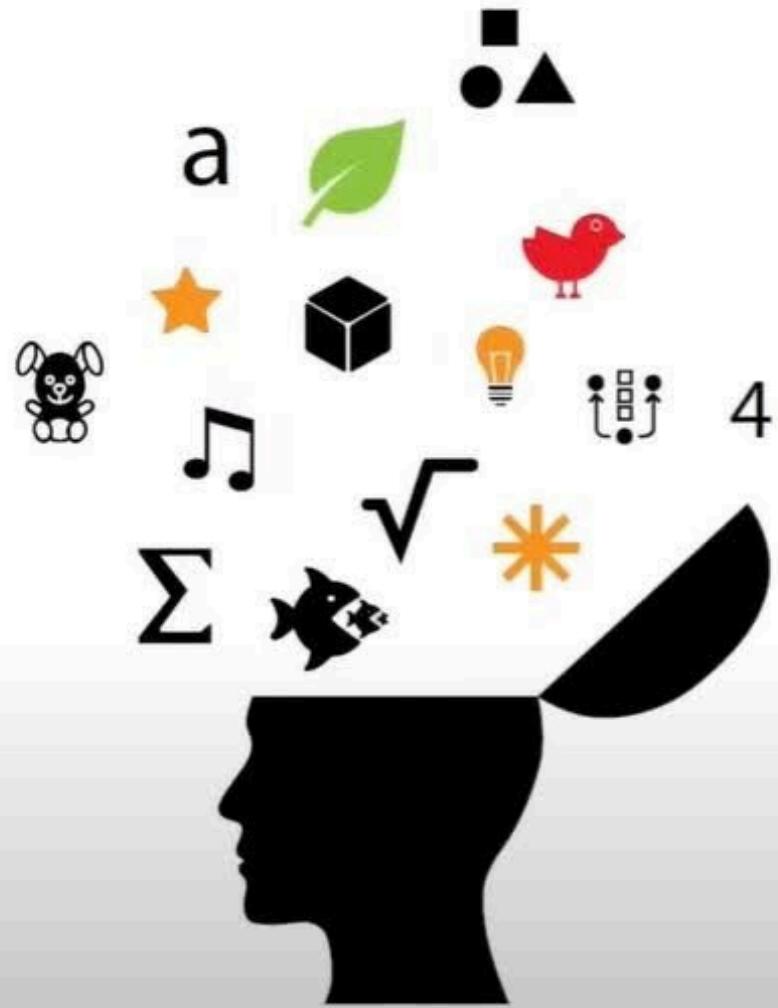
« Le conformisme de l'apprentissage social chez les mouches prédit des traditions durables en matière de choix de partenaire »

# [ intelligence ]

« Capacité de percevoir ou d'inférer l'information, et de la conserver comme une connaissance à appliquer à des comportements adaptatifs dans un environnement ou un contexte donné »

« Ability to perceive or infer information, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context »\*





# [ intelligence ]

« Ensemble des **fonctions** mentales ayant pour objet la connaissance **conceptuelle** et **rationnelle** »\*

« *Set of mental functions aimed at conceptual, rational knowledge* »

D'un point de vue plus scientifique, **Howard Gardner** (*Gardner, H. Intelligences multiples -1993*) à travers sa théorie des intelligences multiples, discute aujourd'hui de 8 types d'intelligence :

- L'intelligence linguistique (ou capacité à s'exprimer)
- l'intelligence logico-mathématique (ou capacité à penser de manière abstraite et à raisonner)
- l'intelligence intrapersonnelle (ou capacité à se connaître soi-même)
- l'intelligence interpersonnelle (ou capacité à interagir les uns avec les autres)
- l'intelligence kinesthésique (ou capacité à utiliser son corps)
- l'intelligence musicale (ou capacité à créer, reproduire, reconnaître des sons, des rythmes, des mélodies)
- l'intelligence visuelle et spatiale (ou capacité à se repérer dans l'espace)
- l'intelligence naturelle (ou capacité à comprendre et à interagir avec la nature)

<https://synapses-lamap.org/2020/04/02/que-peut-on-dire-de-lidee-dintelligences-multiples-et-de-son-application-en-classe/>

## L'intelligence artificielle, de quelle forme d'intelligence s'agit-il ?

l'intelligence artificielle est « **l'ensemble des théories et des techniques mises en œuvre pour créer des machines capables de simuler l'intelligence humaine** »

les grands principes de la logique et du raisonnement, que sont les principes de déduction, d'induction ou d'inférence que nous utilisons dans notre raisonnement quotidien pour prendre des décisions, sont également les grands principes des familles de méthodes de l'IA.

Niveau 1

# Intelligence Artificielle

## IA Symbolique

L'IA de la logique et des règles.

- Raisonnement
- Connaissance explicite
- Systèmes experts

Utilise des symboles et des règles de traitement de l'information pour effectuer des tâches, à la manière d'un expert humain qui suit une procédure.

## IA Connexionniste

Données

Niveau 2

## Machine Learning

Apprentissage

Niveau 3

## Deep Learning

Profondeur / Réseaux

Niveau 4

## Architectures Clés Transformers

Attention

Niveau 5

## IA Générative

Création

Niveau 6

## LLM (Grands Modèles de Langage)

Langage

Niveau 7

## GPT

Famille de modèles

Niveau 8

## ChatGPT

Application

# principales méthodes de machine learning

## supervisé

### classification

données avec label  
(pixels) -> (nombre)



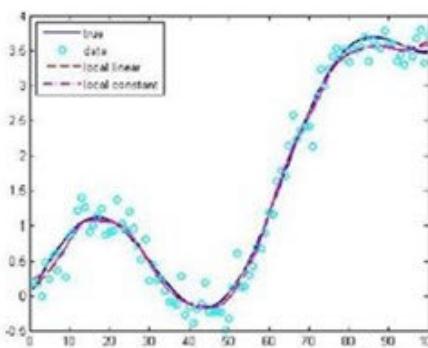
**regrouper** automatiquement les objets en classes, et prédire l'appartenance d'un nouvel objet à une classe identifiée

**type d'objet** : complexes (image, voix, ...)

**exemples** : prédire si une tumeur est cancéreuse en fonction de critères multiples, identifier un spam

### régression

données avec label  
prévoir (y) en fonction de (x)



**prédire** une valeur en fonction de données d'entraînement multidimensionnelles

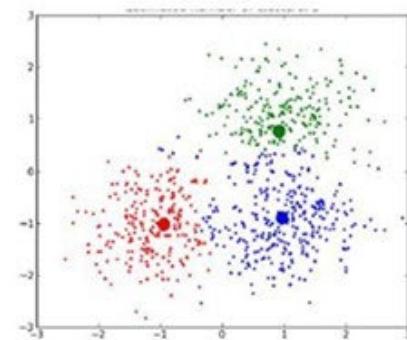
**type d'objet** : valeurs numériques continues

**exemples** : anticipation de churn client, de demande client, évaluation de pipe client, prévision de panne, prévision de récidive

## non supervisé

### clustering

données sans label  
(x, y, z, ...)



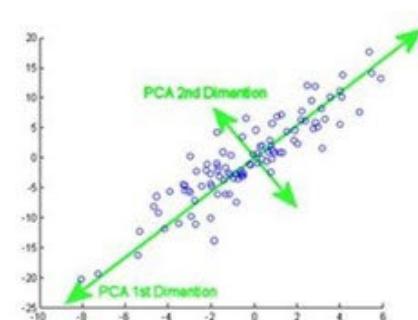
**segmenter** automatiquement un jeu de vecteurs (x, y, z, etc)

**type d'objet** : ensemble de n-uplets de valeurs numériques

**exemples** : détection de fraude, blanchiment d'argent sale, détection de faille de sécurité

### réduction dimensions

données sans label  
(x, y, z, ...)



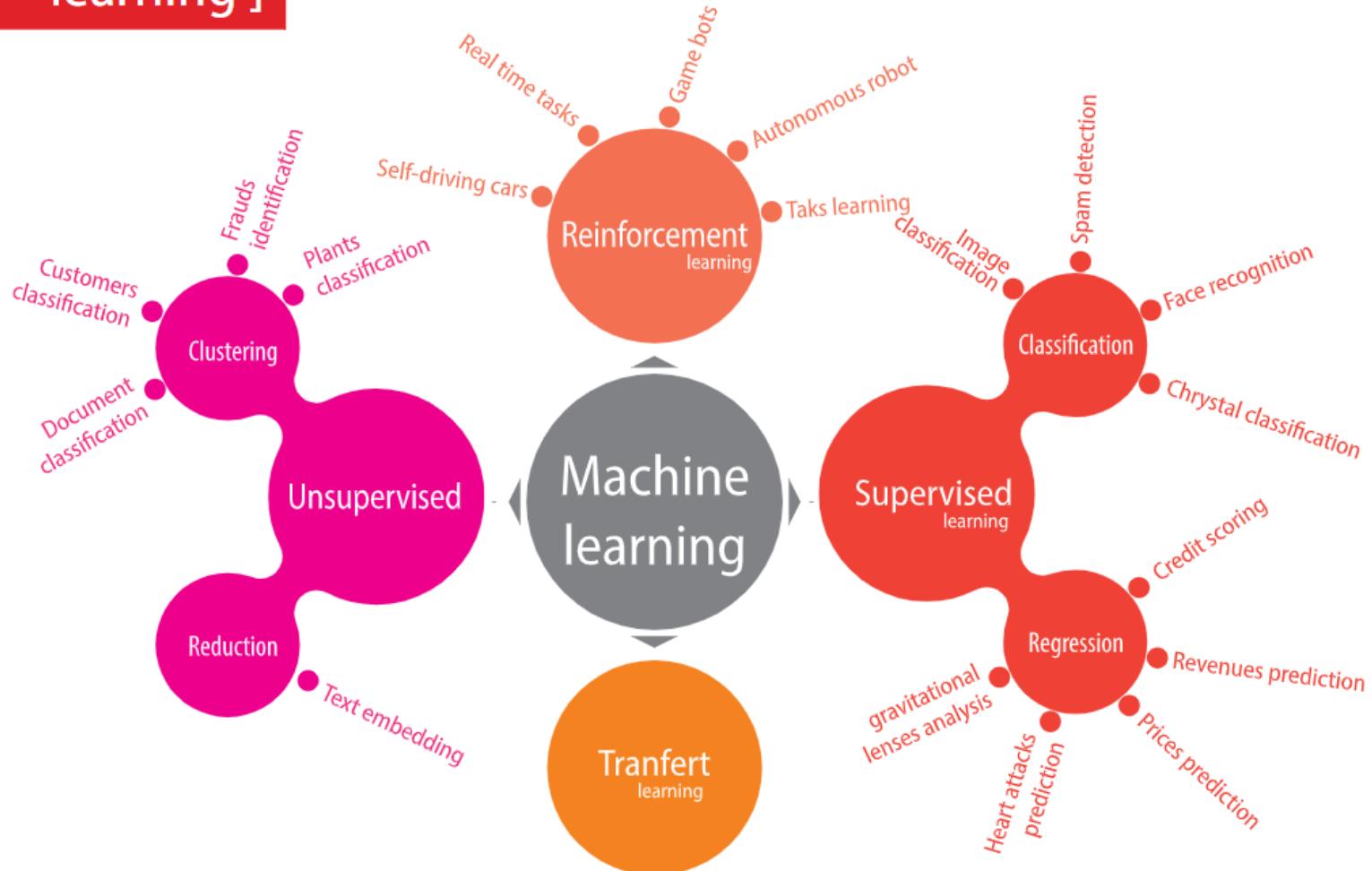
**déterminer** automatiquement les paramètres discriminants d'un jeu de données par rapport à une variable cible

**type d'objet** : ensemble de n-uplets de valeurs numériques

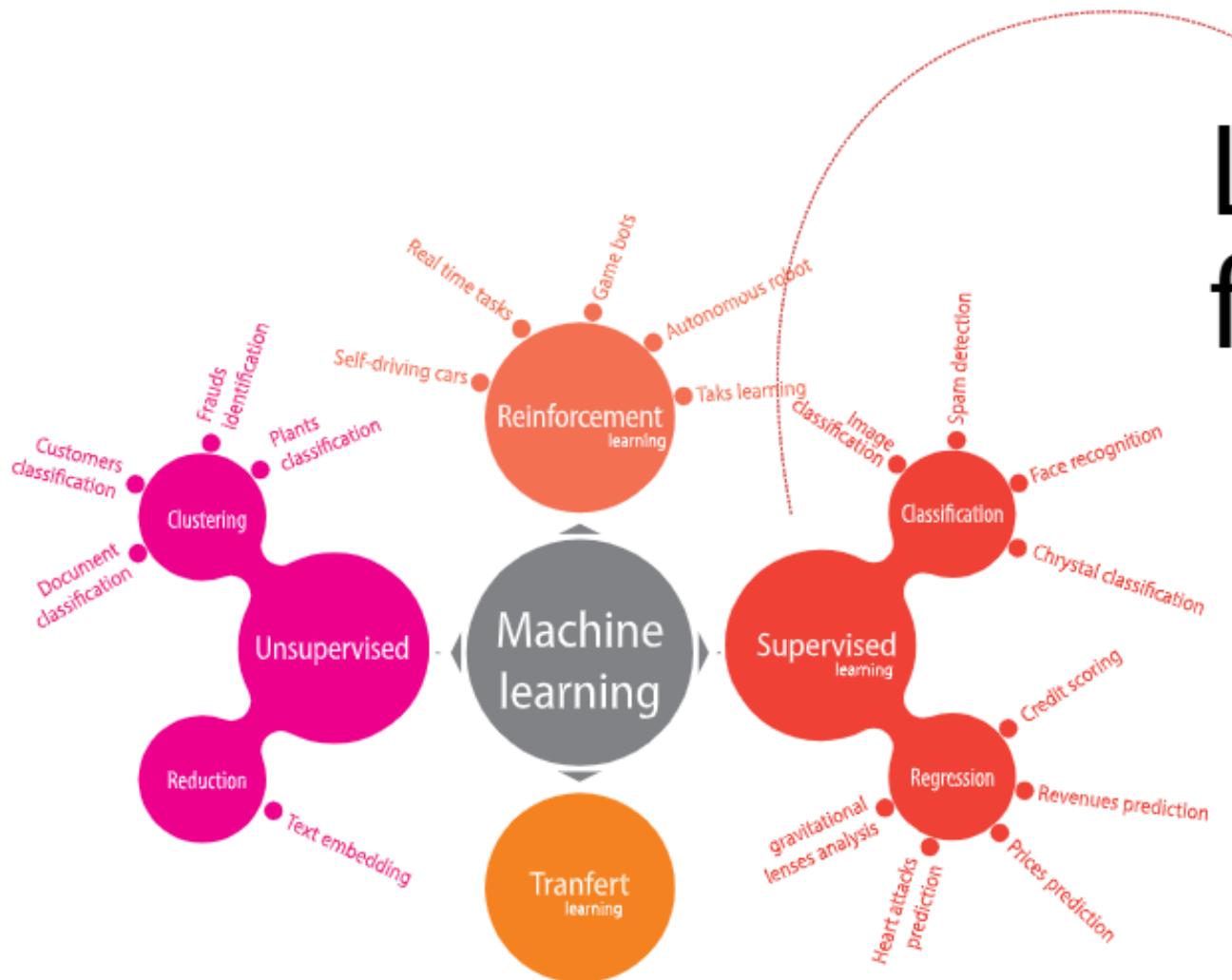
**exemples** : identifier les paramètres déterminants la corrélation entre des paramètres clients et leur comportement futur

# Mais ce n'est pas que ça !

[ \*-learning ]



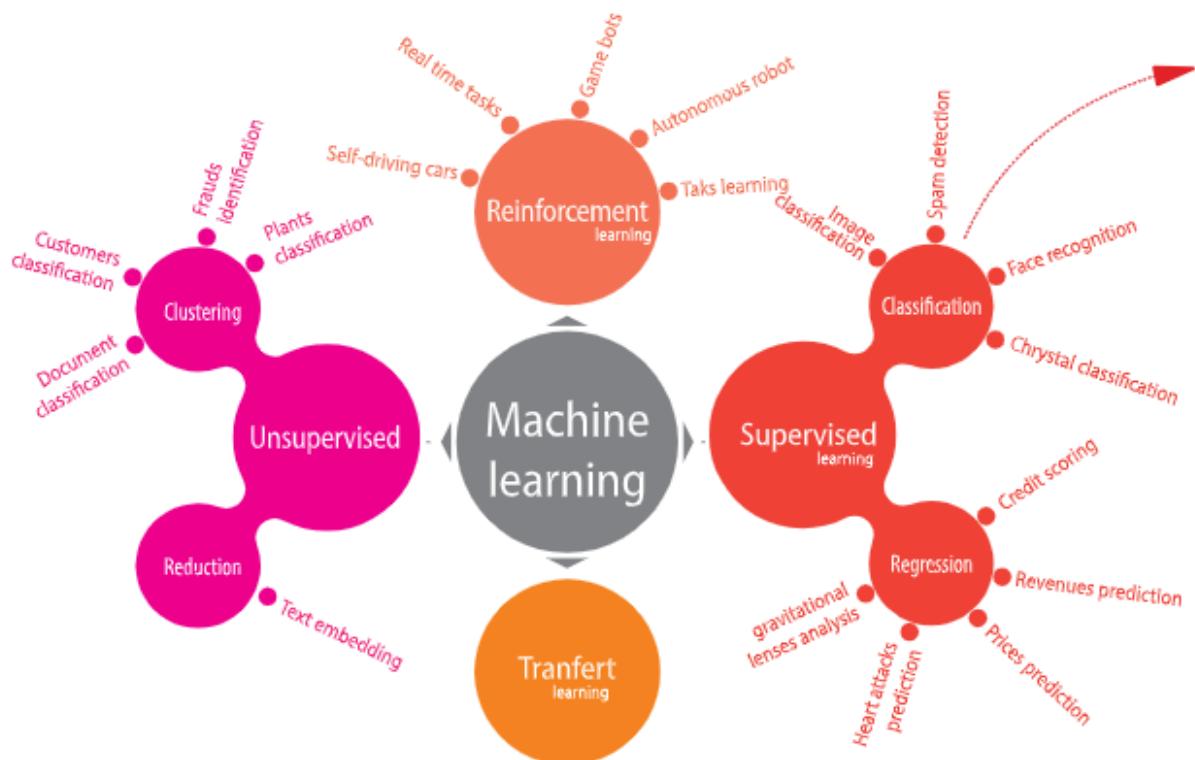
# Supervised learning



# Learning from labeled data

# Supervised learning

Learning from labeled data



**Classification :**  
Predict qualitative informations



This is a cat



This is a rabbit

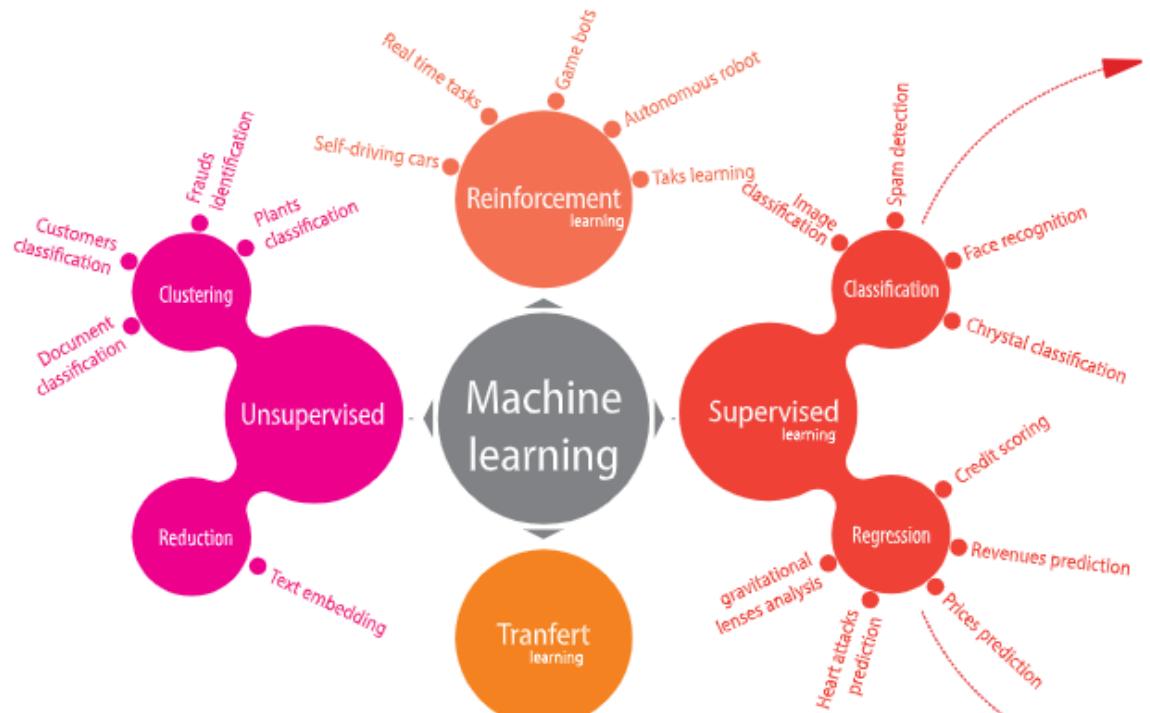


Tell me,  
what is it ?



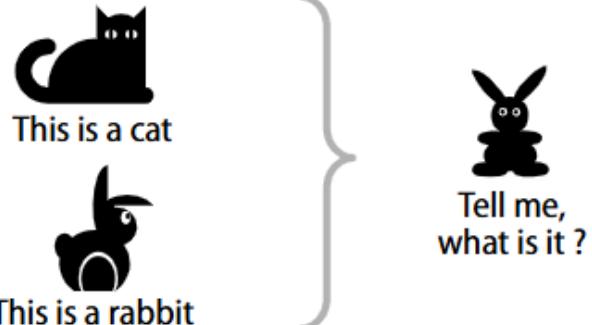
# Supervised learning

## Learning from labeled data



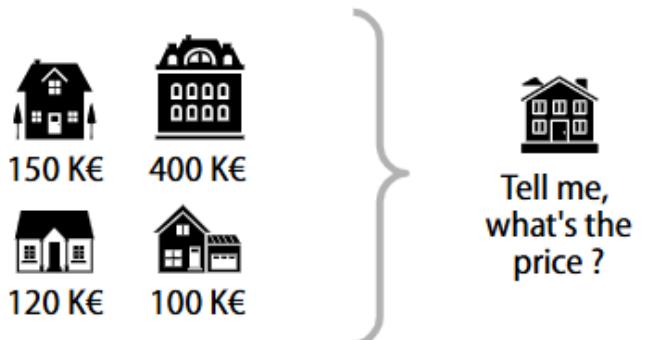
### Classification :

Predict qualitative informations



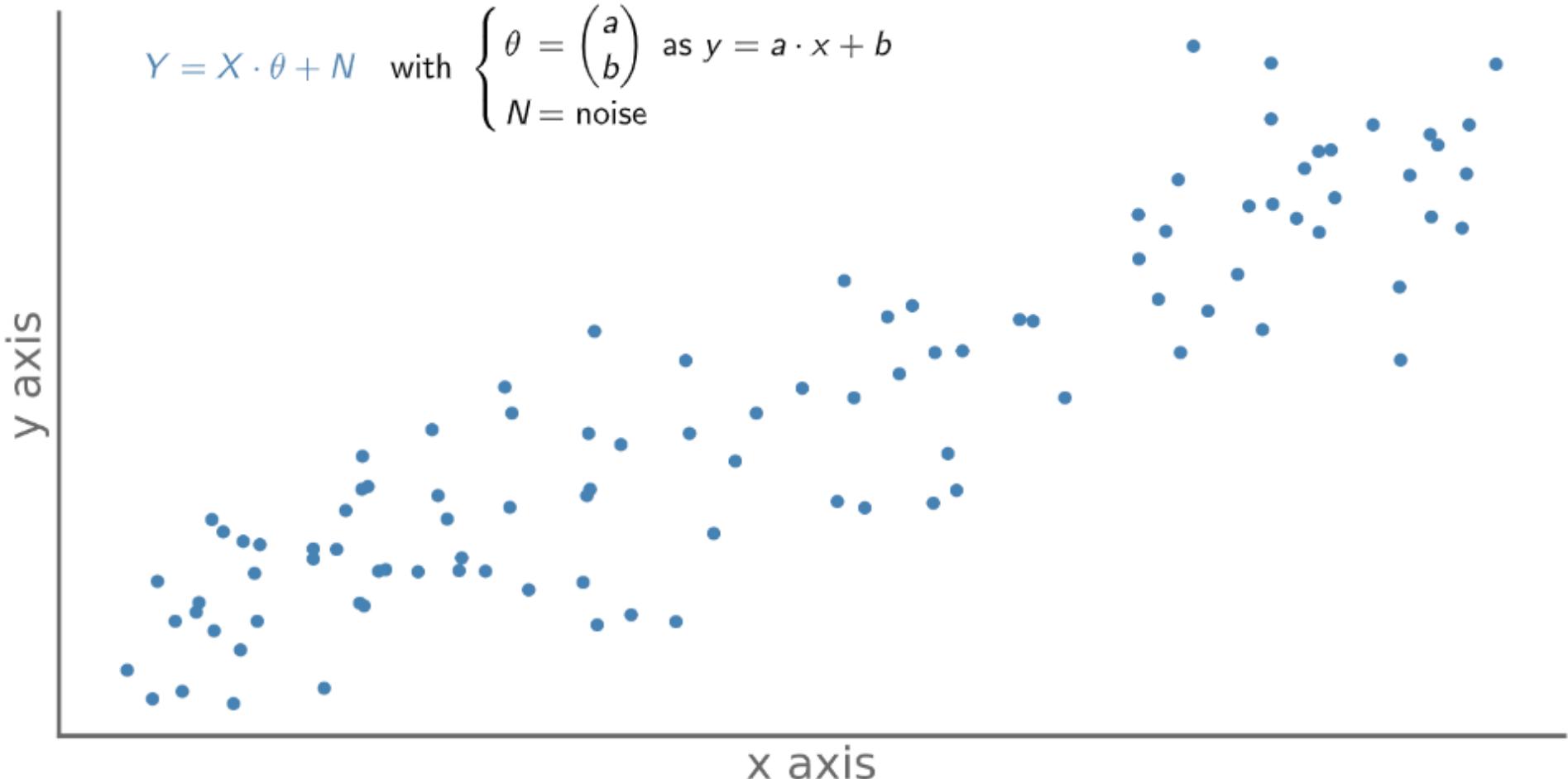
### Régression :

Predict quantitative informations



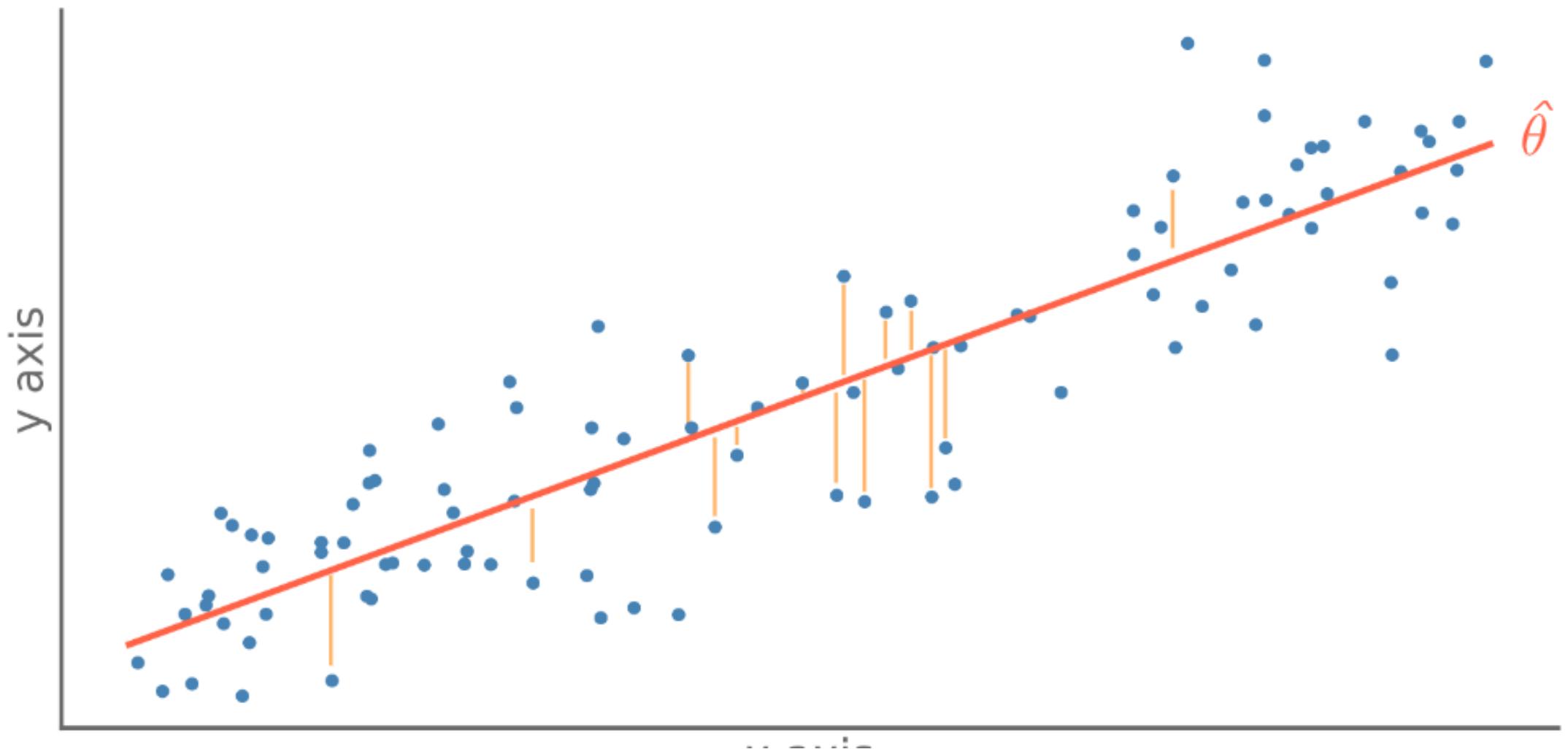
# Linear regression

We have a phenomenon, for which we have observations



# Linear regression

"As close as possible" means "minimize the distance" between the line and our points (observations).



# Linear regression

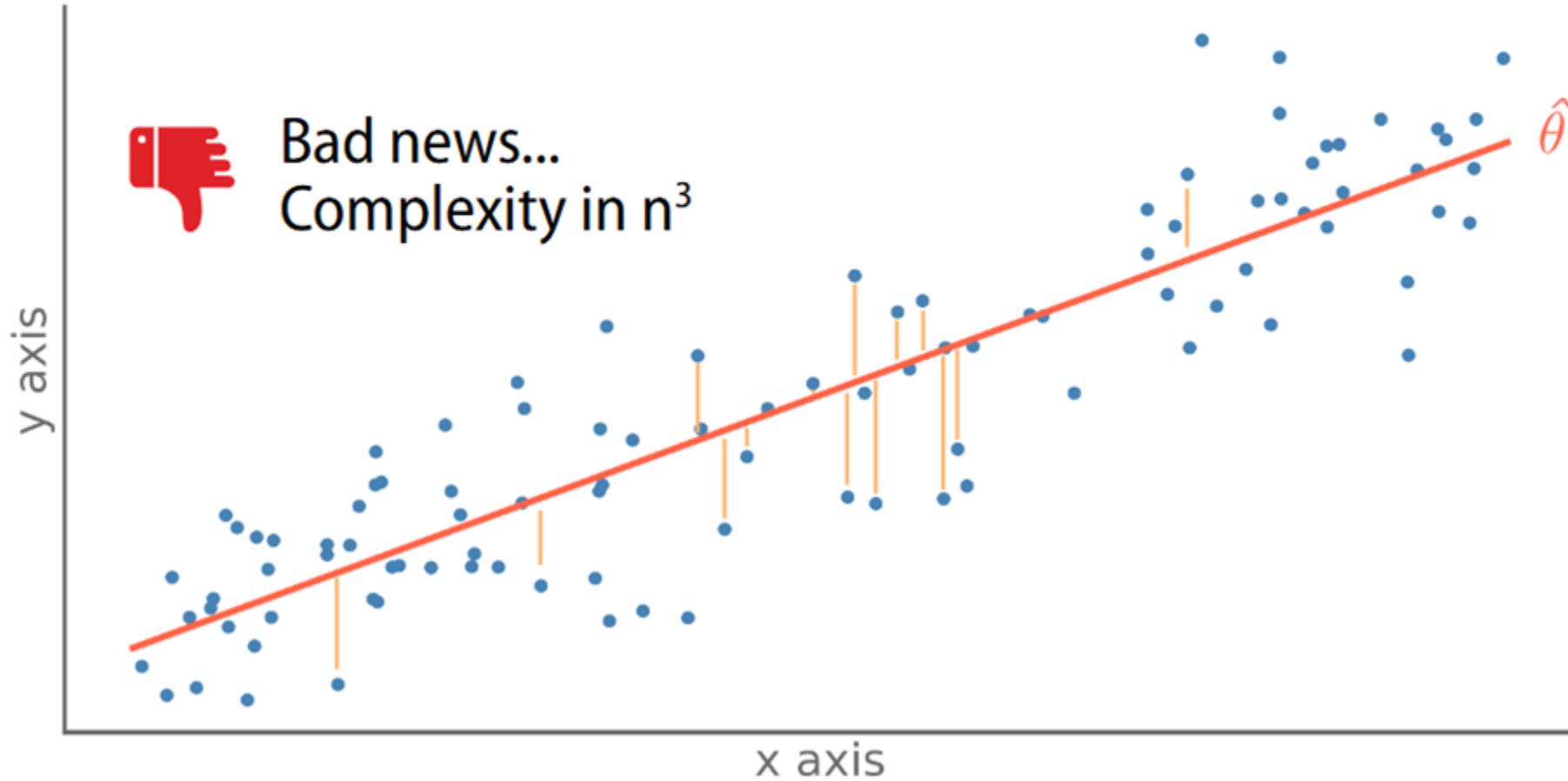
$$\hat{\theta} = (X^T X)^{-1} X^T Y$$



Good news !  
We have a direct solution !



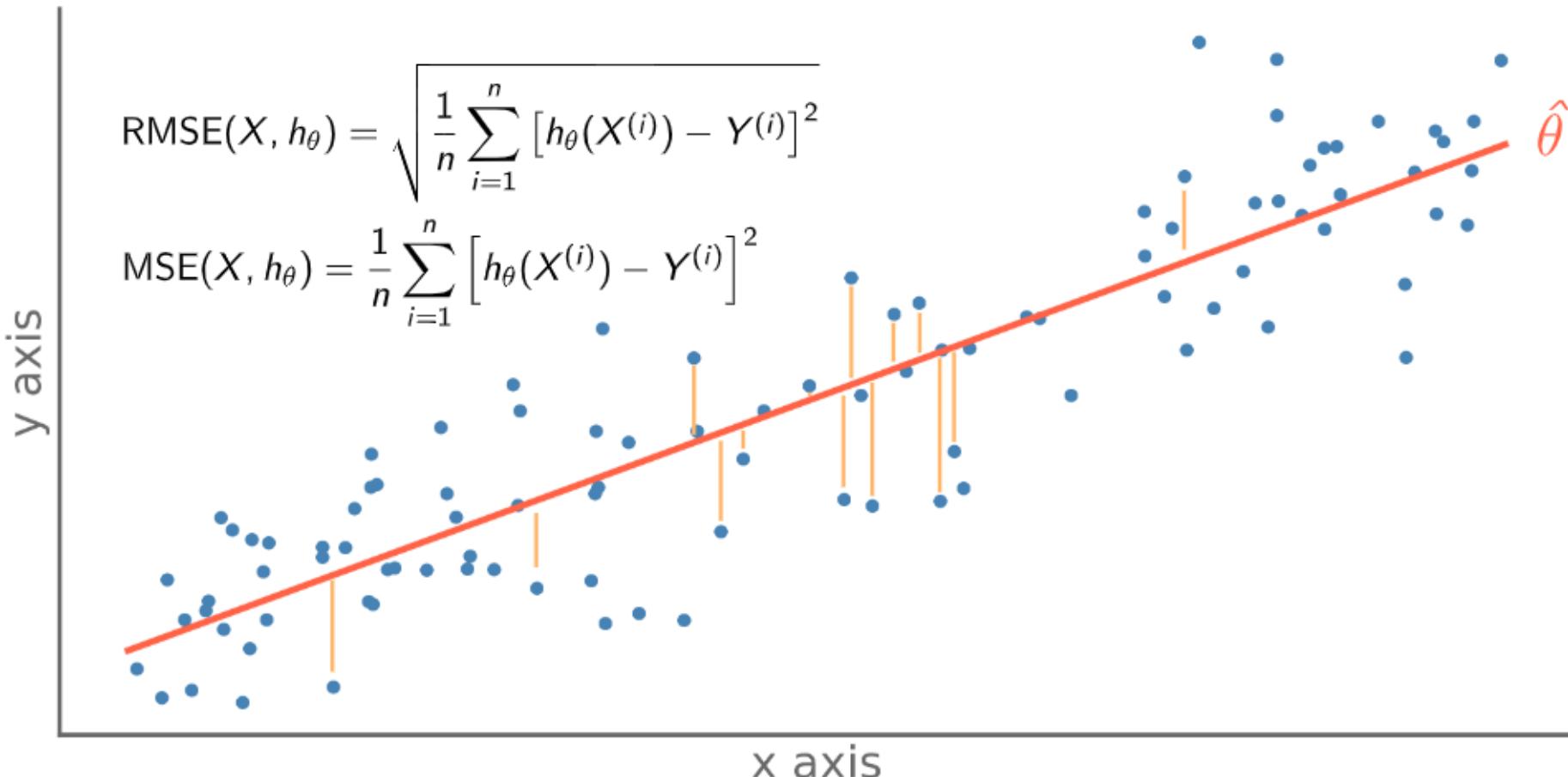
Bad news...  
Complexity in  $n^3$



# Loss function



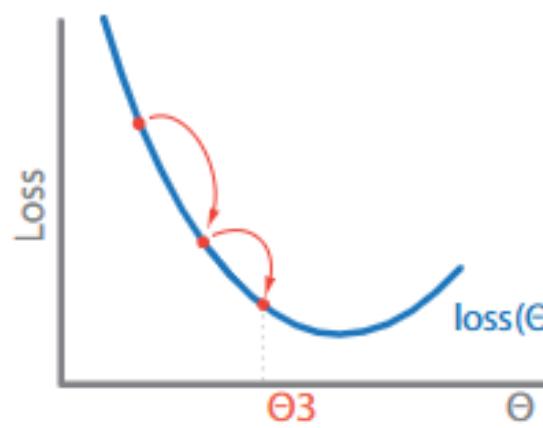
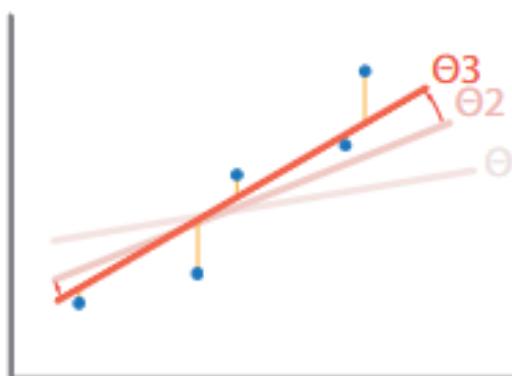
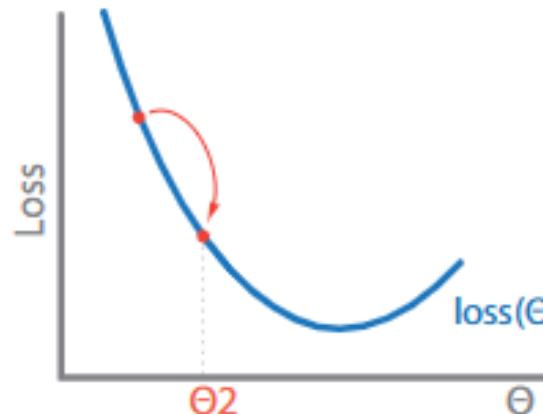
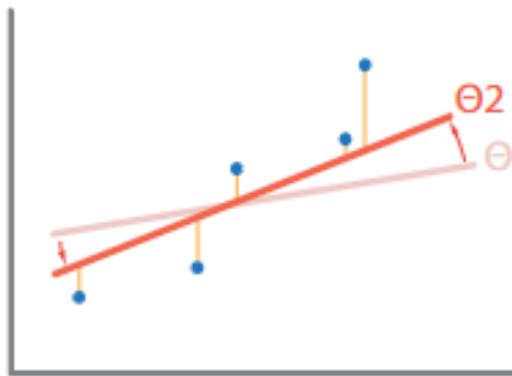
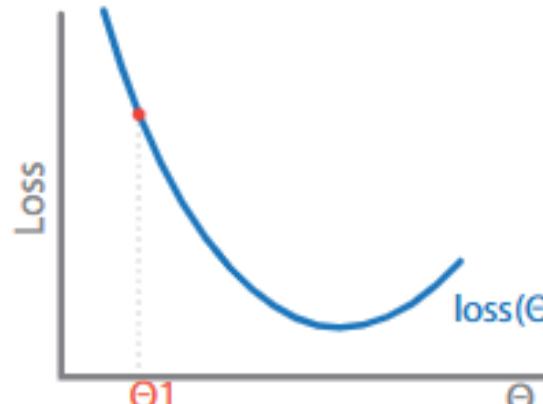
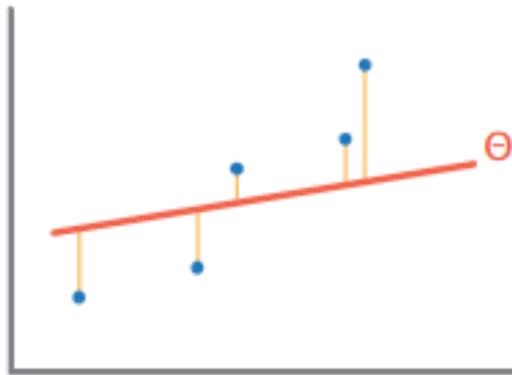
We will use an «loss function», which we will try to minimize.



RMSE : Root Mean Square Error - Erreur quadratique moyenne

MSE : Mean Squared Error - Moyenne du carré des erreurs

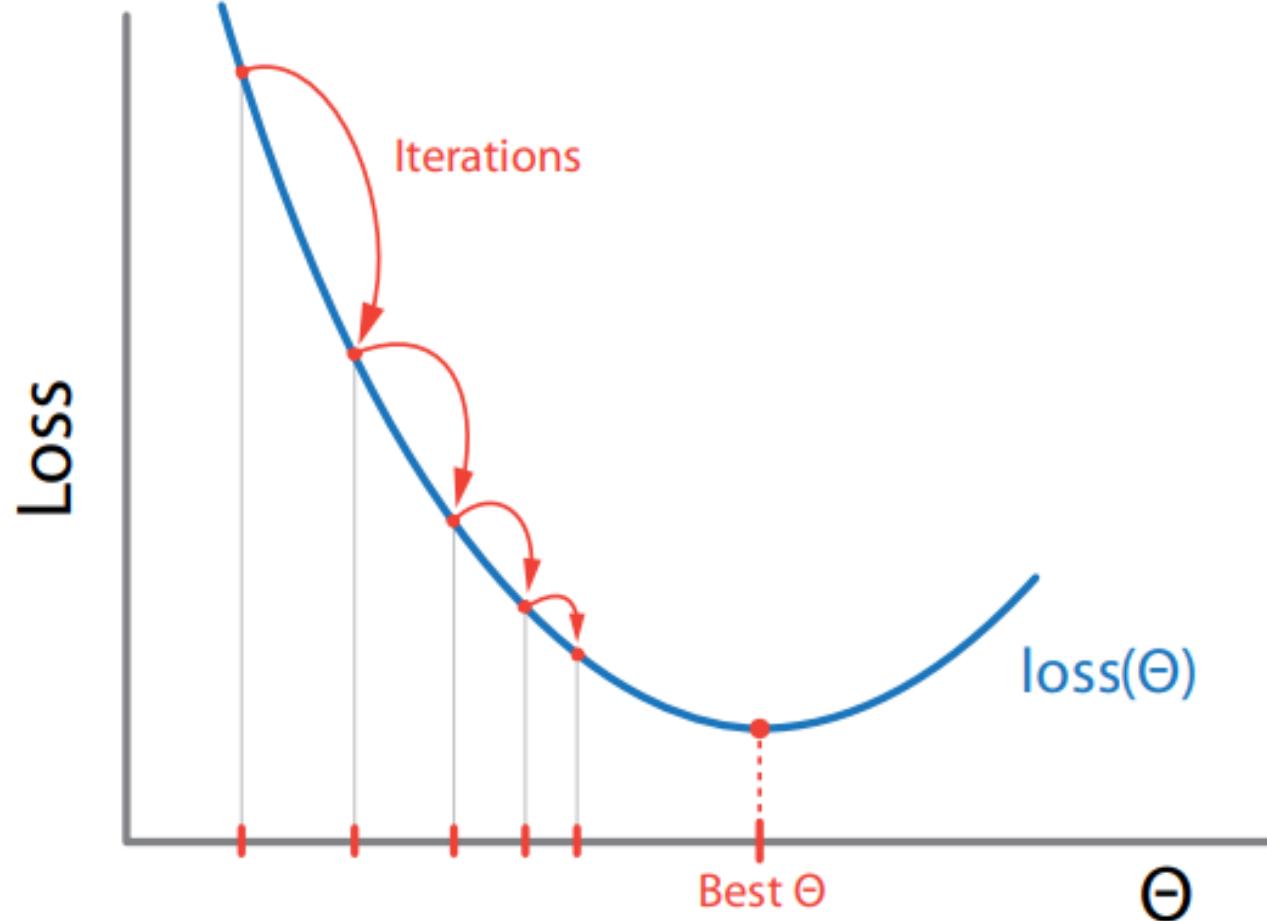
# Optimization



We will iteratively look for the best position of our line, by varying its parameters ( $\Theta$ ).

This is called **optimization**

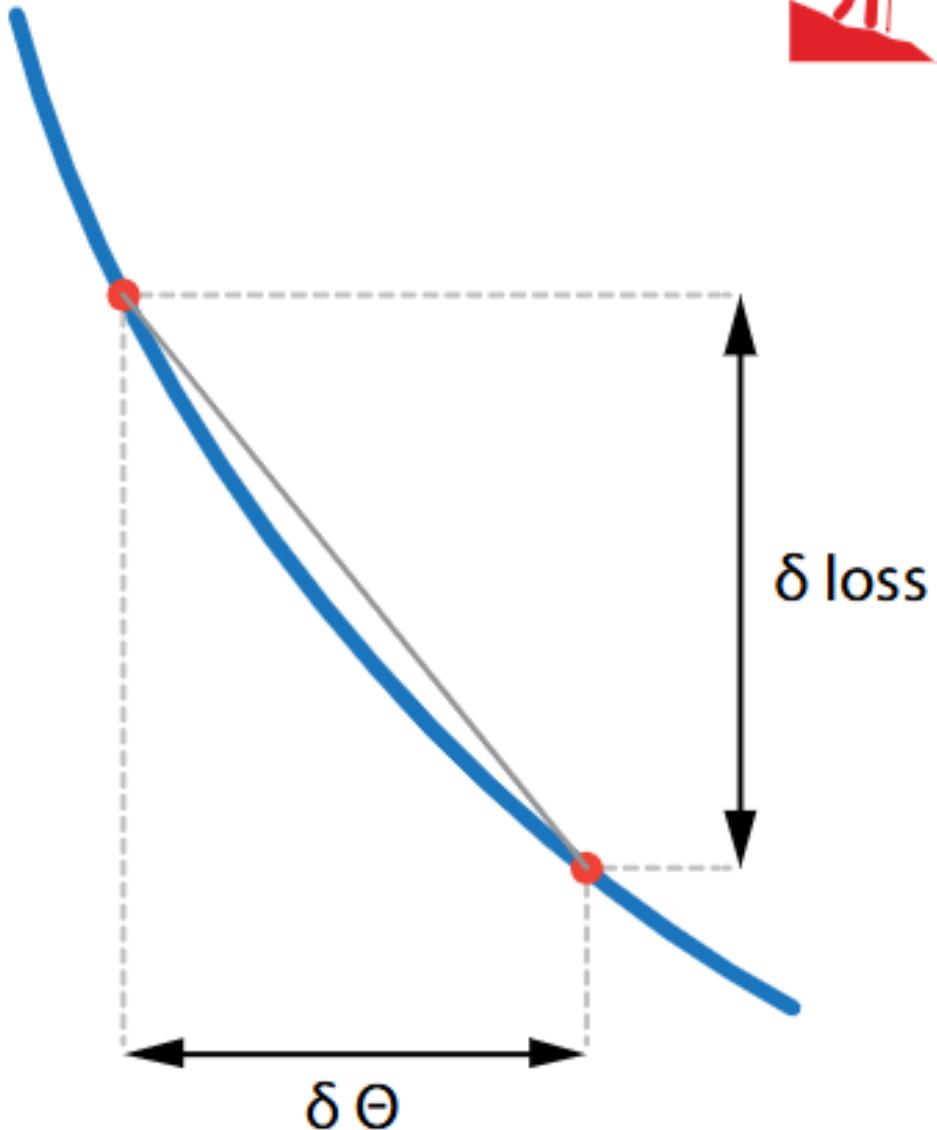
# Optimization



We will iteratively look for the best position of our line, by varying its parameters ( $\Theta$ ).

This is called **optimization**

# Gradient descent



By changing  $\Theta$  from  $\delta\Theta$   
We improve  $\text{loss}(\Theta)$  of  $\delta\text{loss}$

The gradient is the slope we will  
follow to minimize our loss function.

$$\text{gradient} = \frac{\delta \text{loss}}{\delta \theta}$$

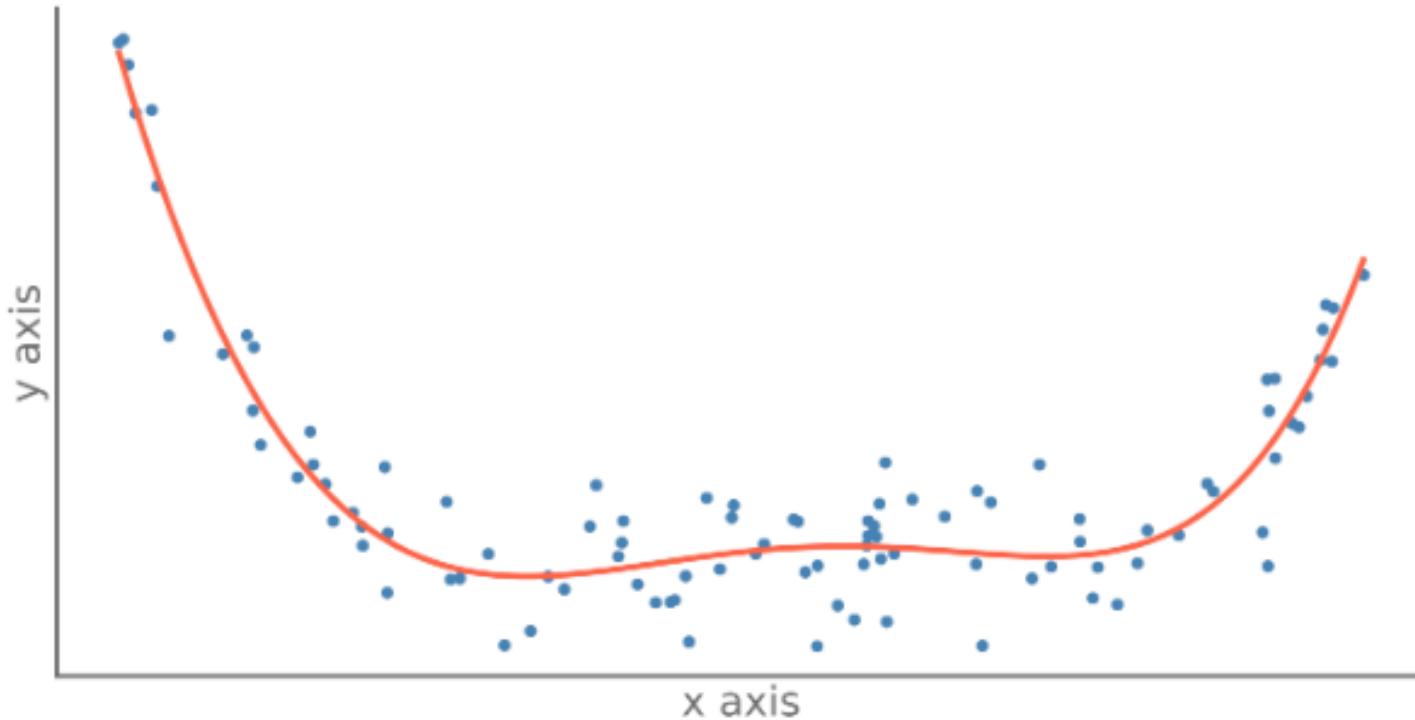
One iterative solution is :  $\theta \leftarrow \theta - \eta \cdot \frac{\delta \text{loss}}{\delta \theta}$

where  $\eta$  is the learning rate

This process is called  
**gradient descent**

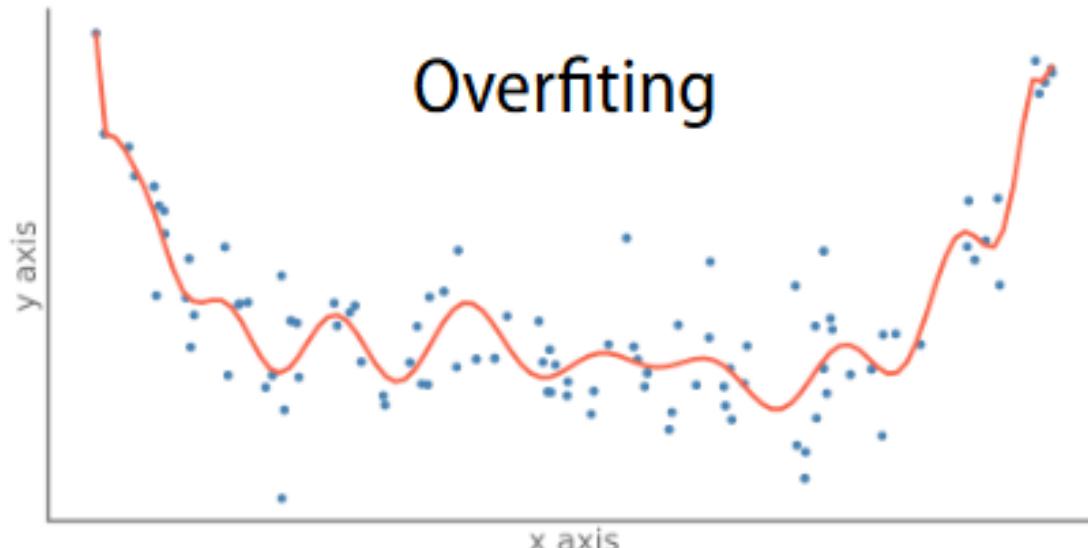
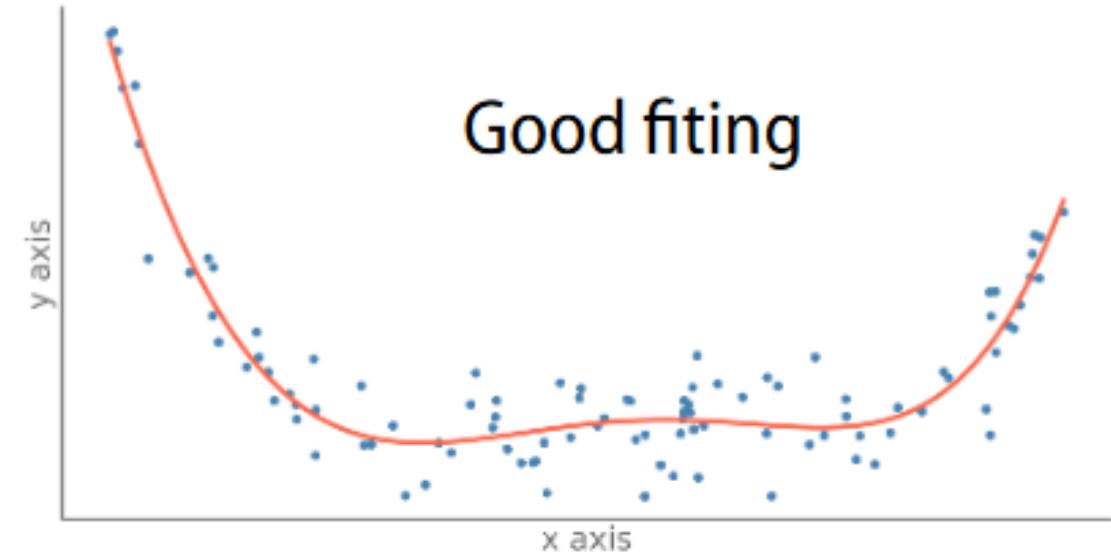
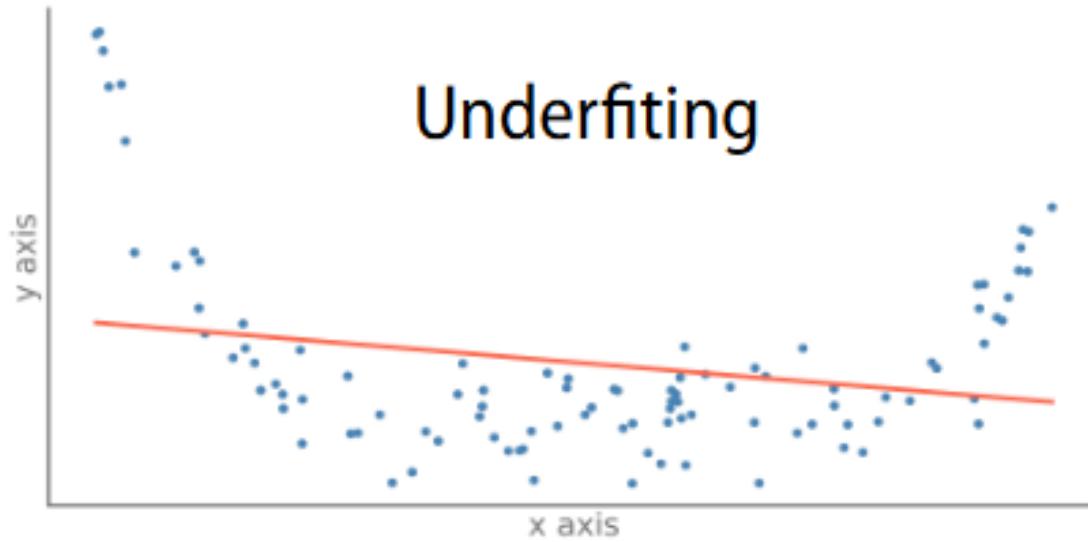
# Tout n'est pas linéaire

## Polynomial regression



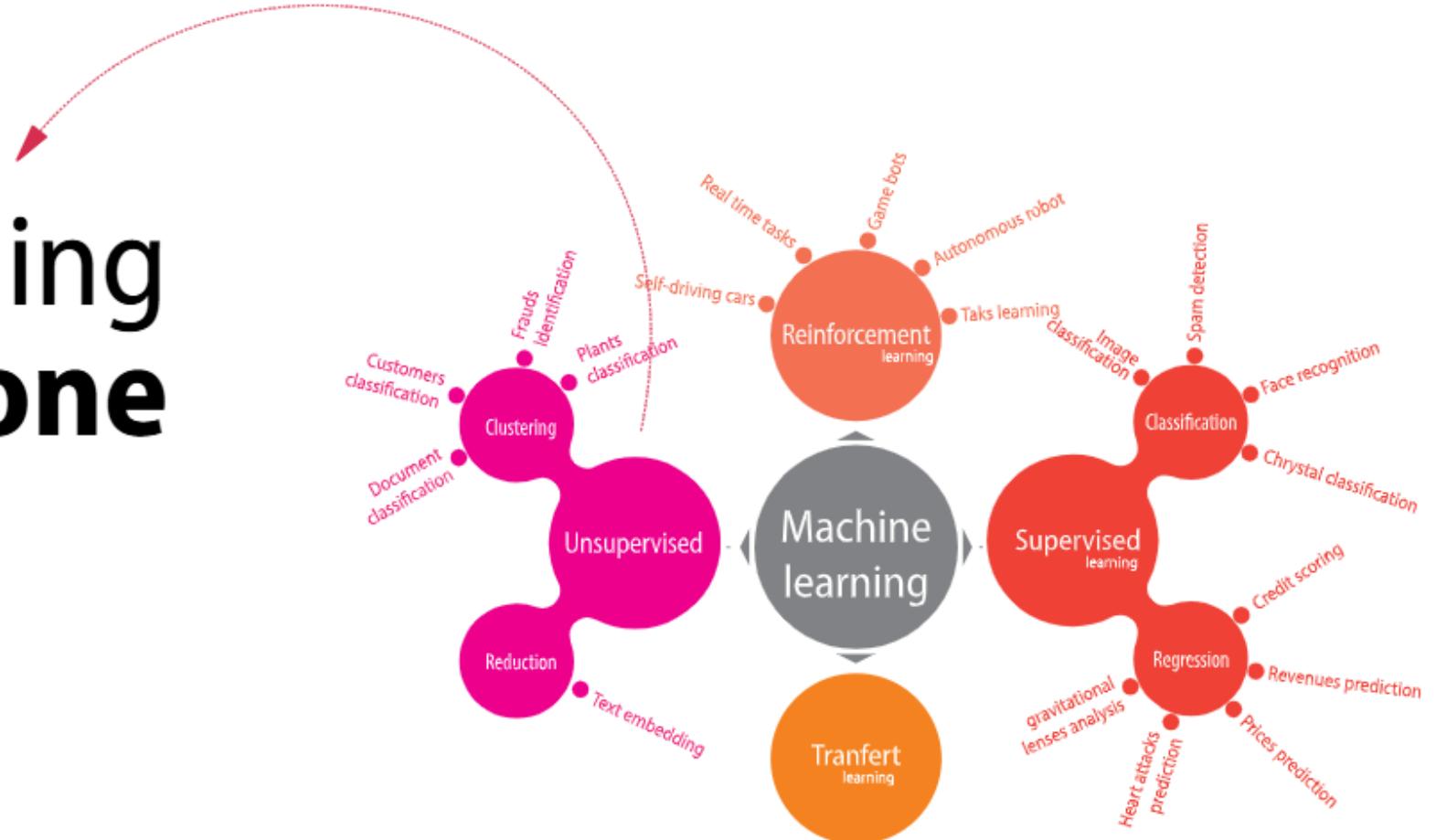
$$P_n(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_n \cdot x^n = \sum_{i=0}^n a_i \cdot x^i$$

# Underfitting and overfitting



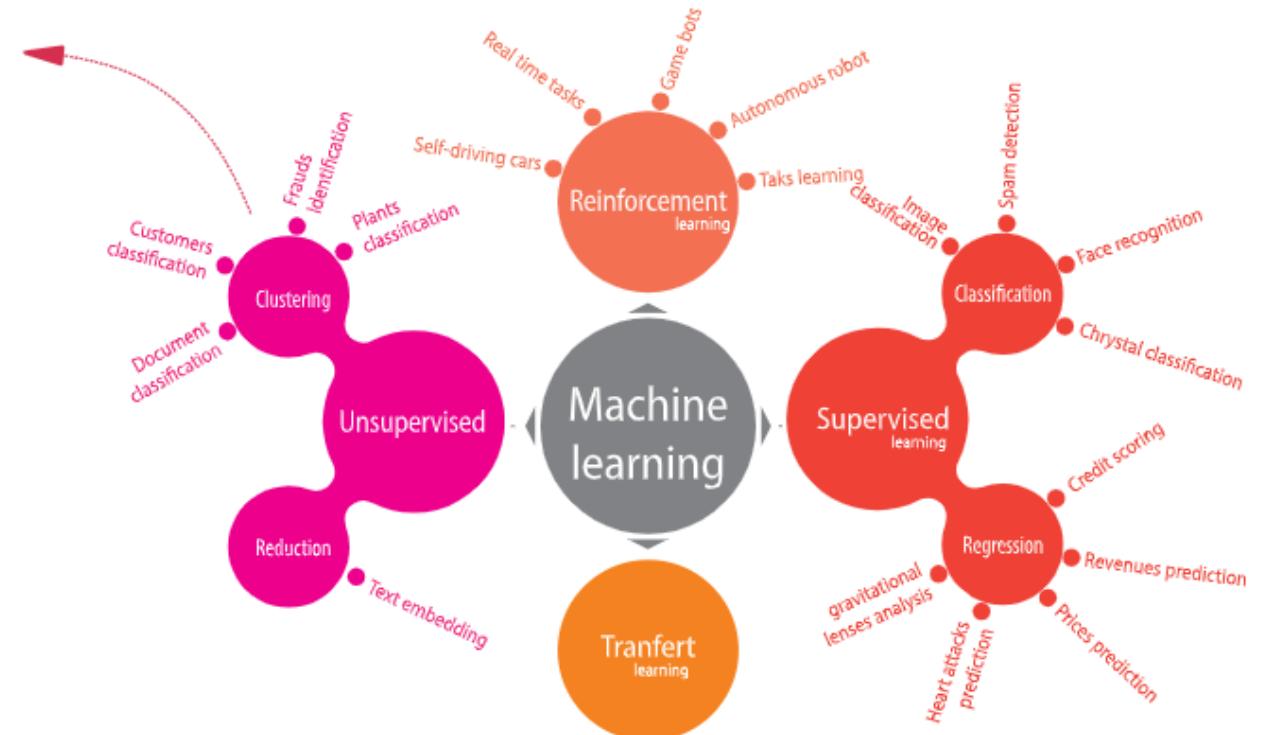
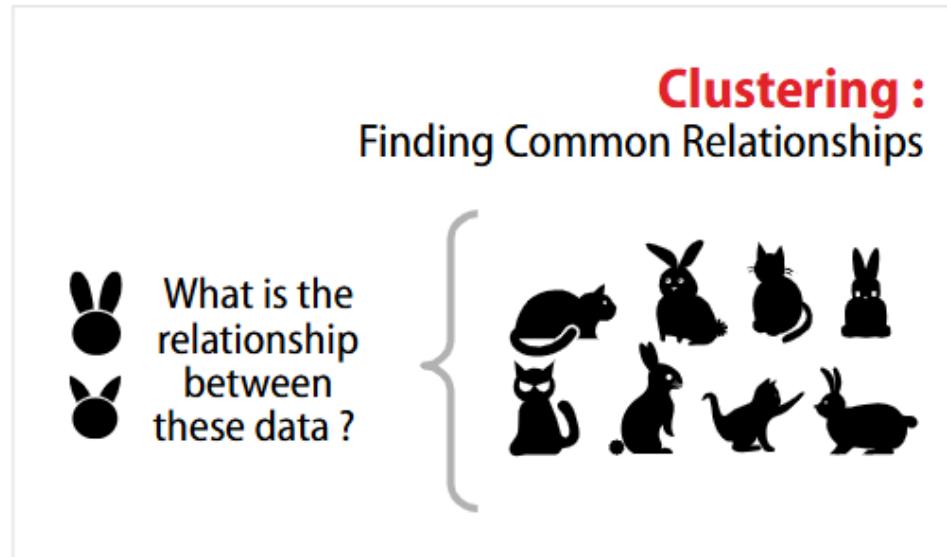
# Unsupervised learning

Learning  
from data **alone**



# Unsupervised learning

Learning from data alone



## Modéliser le cerveau :

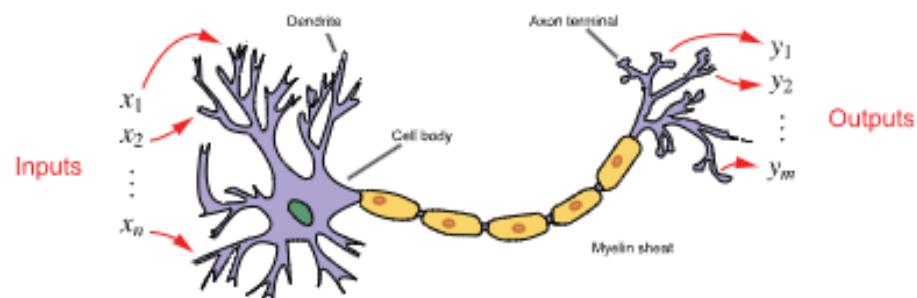
« Penser s'apparente à un calcul massivement parallèle de **fonctions élémentaires**.

L'information est un **signal** avant d'être un code »<sup>1</sup>



### Connexionniste

Modéliser le cerveau  
Modelling the brain



vs

### Symbolique

Forger une opinion  
Making a mind

Tout [homme] est [mortel]  
[Socrate] est un [homme]  
Donc [Socrate] est [mortel]

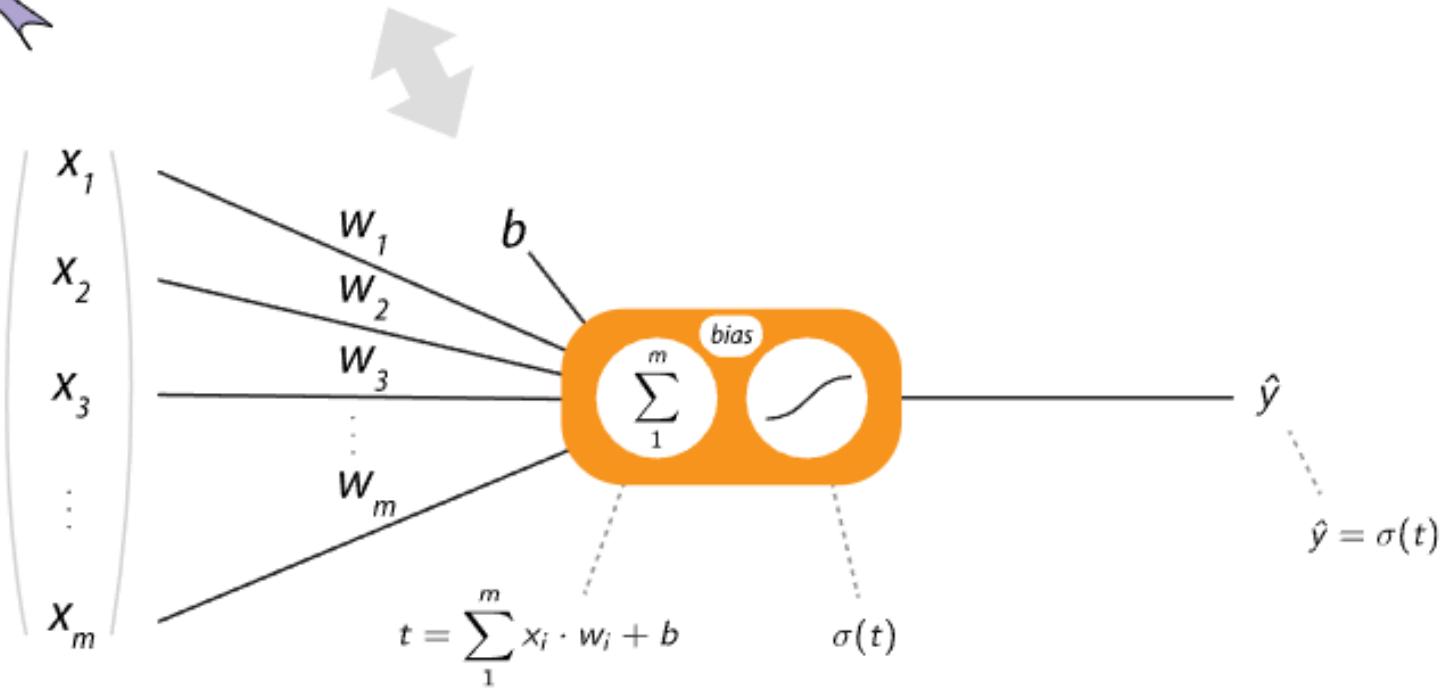
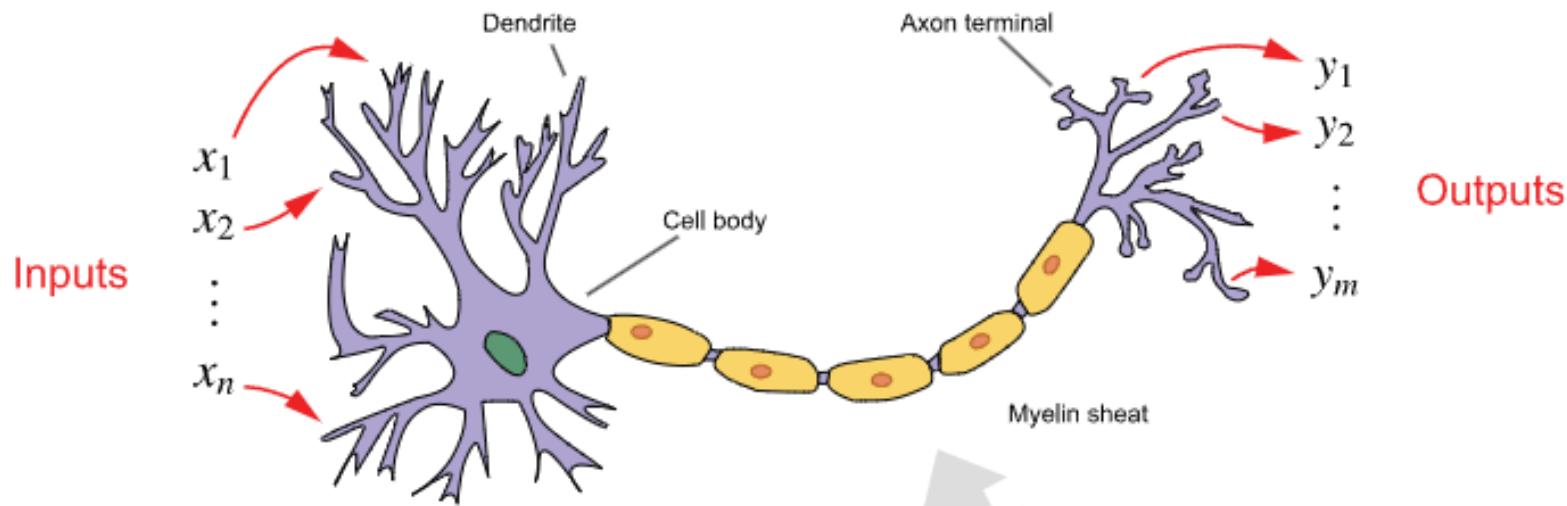


<sup>1</sup> D Cardon, JP Cointet, A Mazieres (2018), La revanche des neurones  
<https://dx.doi.org/10.3917/res.211.0173>

## Forger une opinion :

« Penser, c'est calculer des **symboles** qui ont à la fois une réalité matérielle et une valeur sémantique de représentation »<sup>1</sup>

L'information est une donnée symbolique de **haut niveau**.



## Approche **inductive**



### Connexionniste

Modéliser le cerveau  
Modelling the brain

## Approche **déductive**



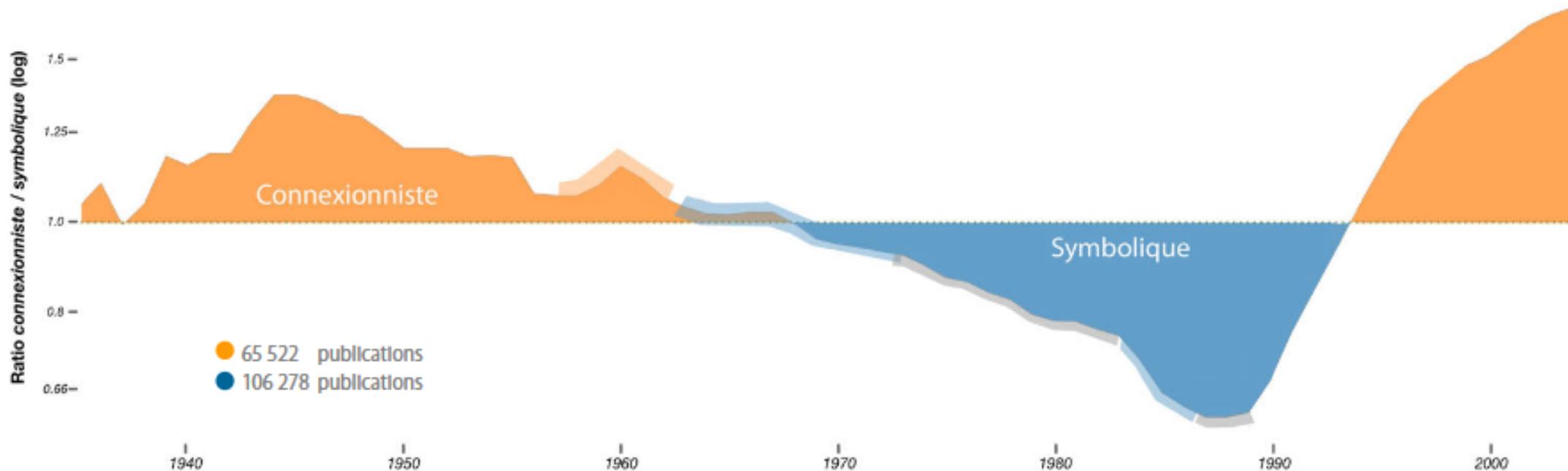
vs

### Symbolique

Forger une opinion  
Making a mind

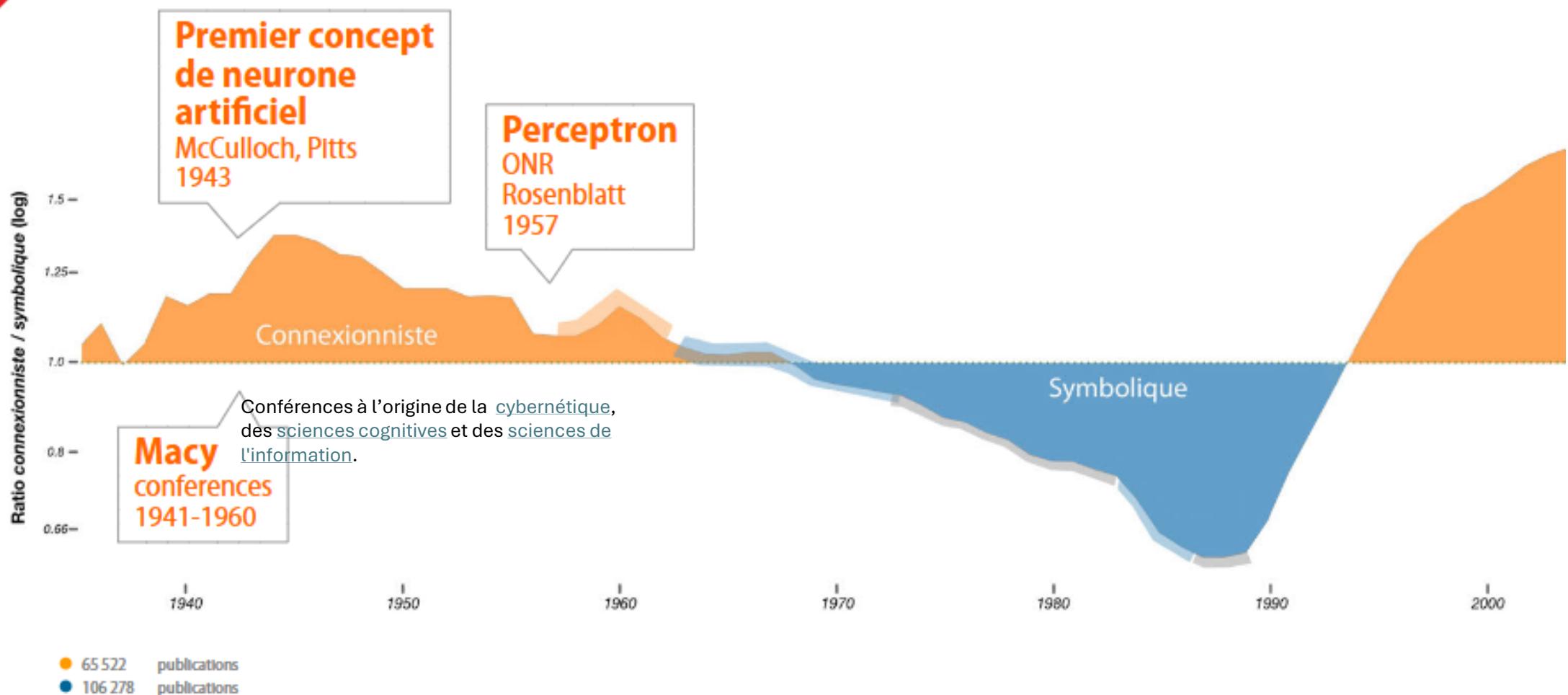
## Influence académique des approches connexioniste et symbolique<sup>1</sup>

Ratio des publications **connexionnistes / symbolistes**



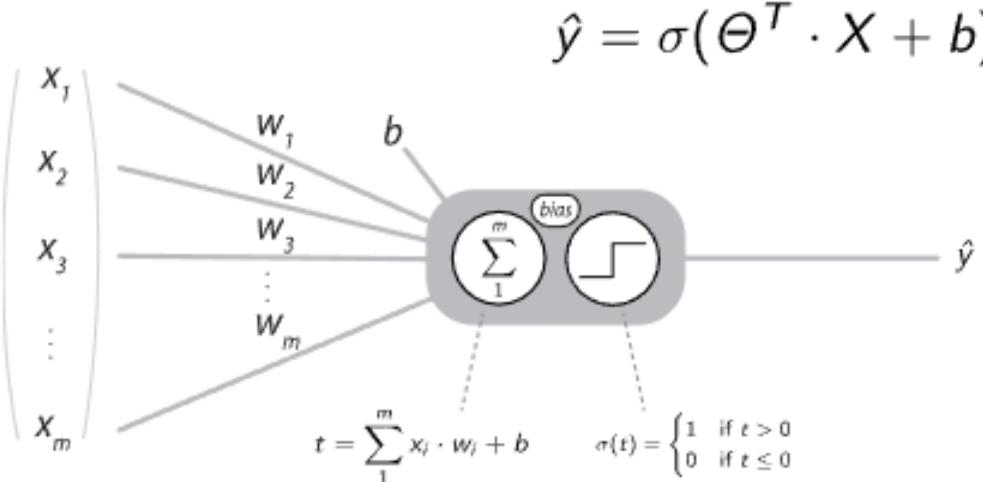
<sup>1</sup> D Cardon, JP Cointet, A Mazieres (2018), La revanche des neurones  
<https://dx.doi.org/10.3917/res.211.0173>

# Influence académique des approches connexionniste et symbolique<sup>1</sup>



<sup>1</sup> D Cardon, JP Cointet, A Mazieres (2018), La revanche des neurones  
<https://dx.doi.org/10.3917/res.211.0173>

# Perceptron



## THE PERCEPTRON

389

: sets of  
ich are  
tend to  
t sets of  
ve and/  
stimuli  
y facilita-  
tion of

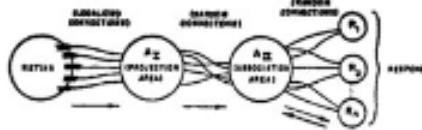
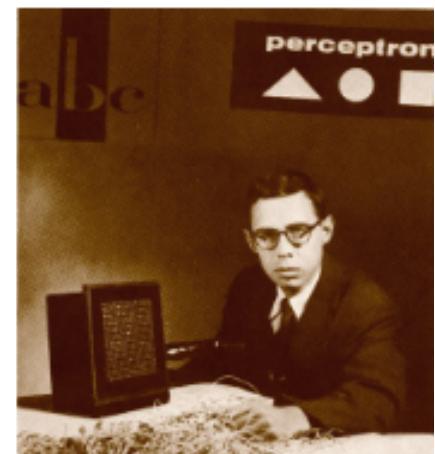


FIG. 1. Organization of a perceptron.

The cells in the projection area each receive a number of connections from

Perceptron  
Frank Rosenblatt  
1958



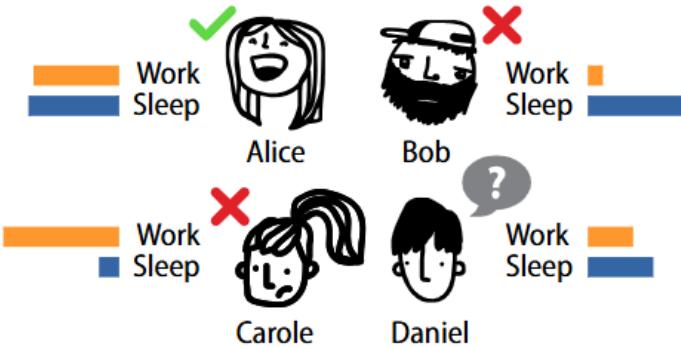
Linear and binary classifier

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain.  
<https://doi.org/10.1037/h0042519>

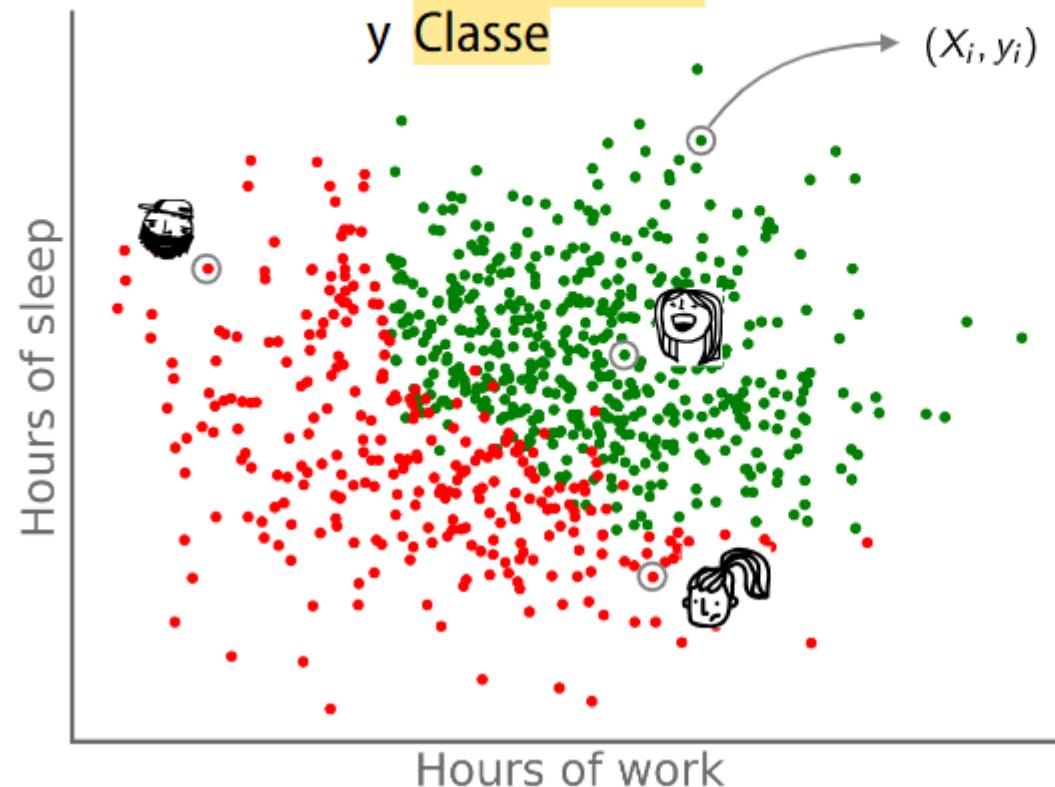
Okay, let's look at an example !

## Logistic regression

Let's try to predict our students' success in an exam, based on their work time and sleep time:



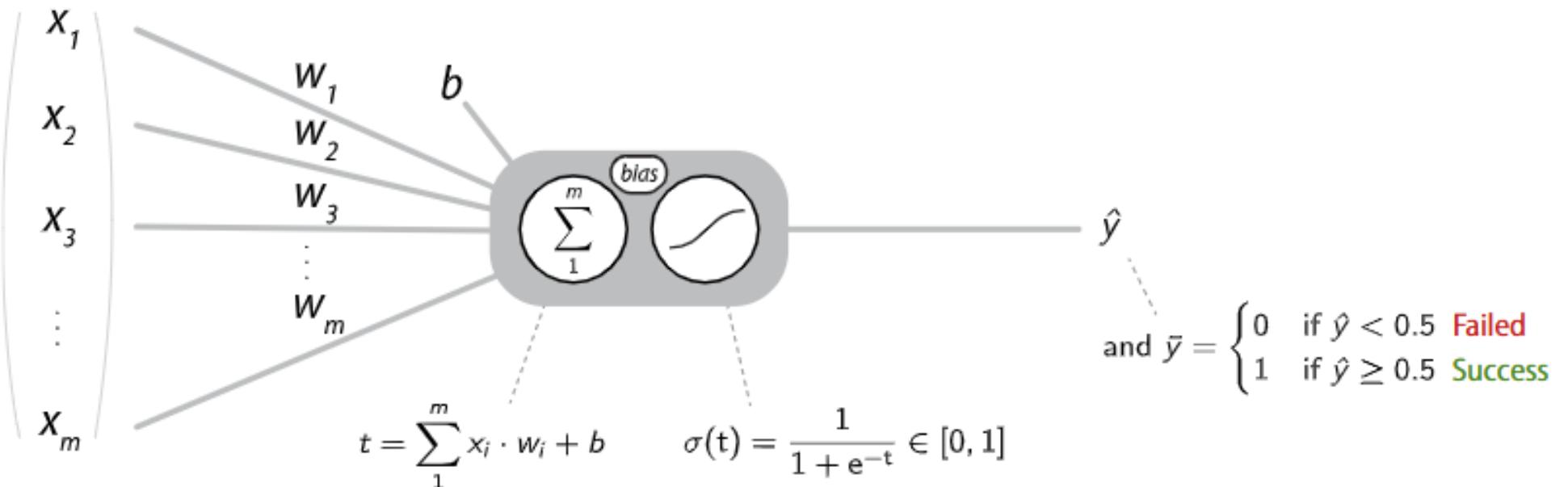
**Dataset :** X Observations  
y Classe



$$\begin{aligned} X_i &= \begin{pmatrix} x_{i1} = \text{Hours of work} \\ x_{i2} = \text{Hours of sleep} \end{pmatrix} \\ y_i &= \begin{cases} 1 & \text{Successful examination} \\ 0 & \text{Failed exam} \end{cases} \end{aligned}$$

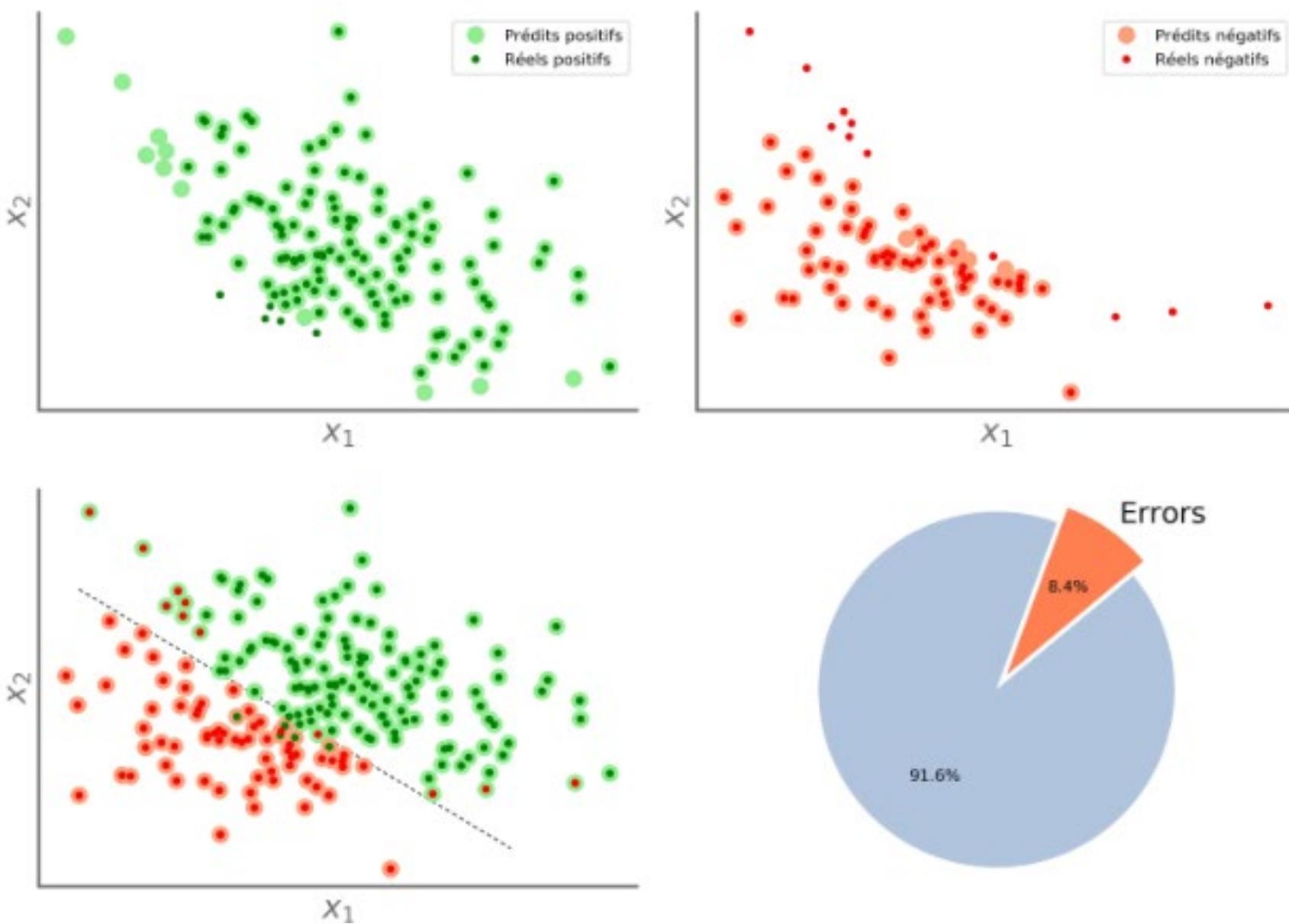
## Logistic regression

$$\hat{y} = \sigma(\Theta^T \cdot X + b)$$



Input	Bias / Weight	Activation function	Output
$X$	$\theta$	$\sigma(t)$	$\hat{y}$

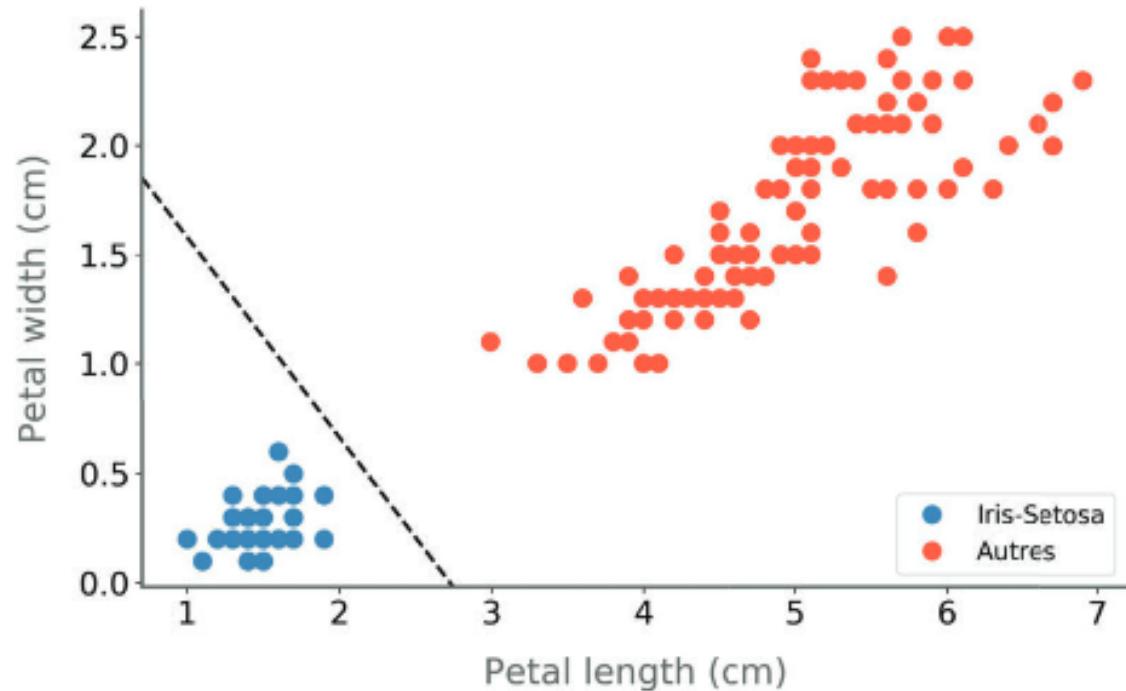
# Logistic regression



# Perceptron

## Iris plants dataset

Dataset from : Fisher, R.A. "The use of multiple measurements In taxonomic problems" Annual Eugenics, 7, Part II, 179-188 (1936)



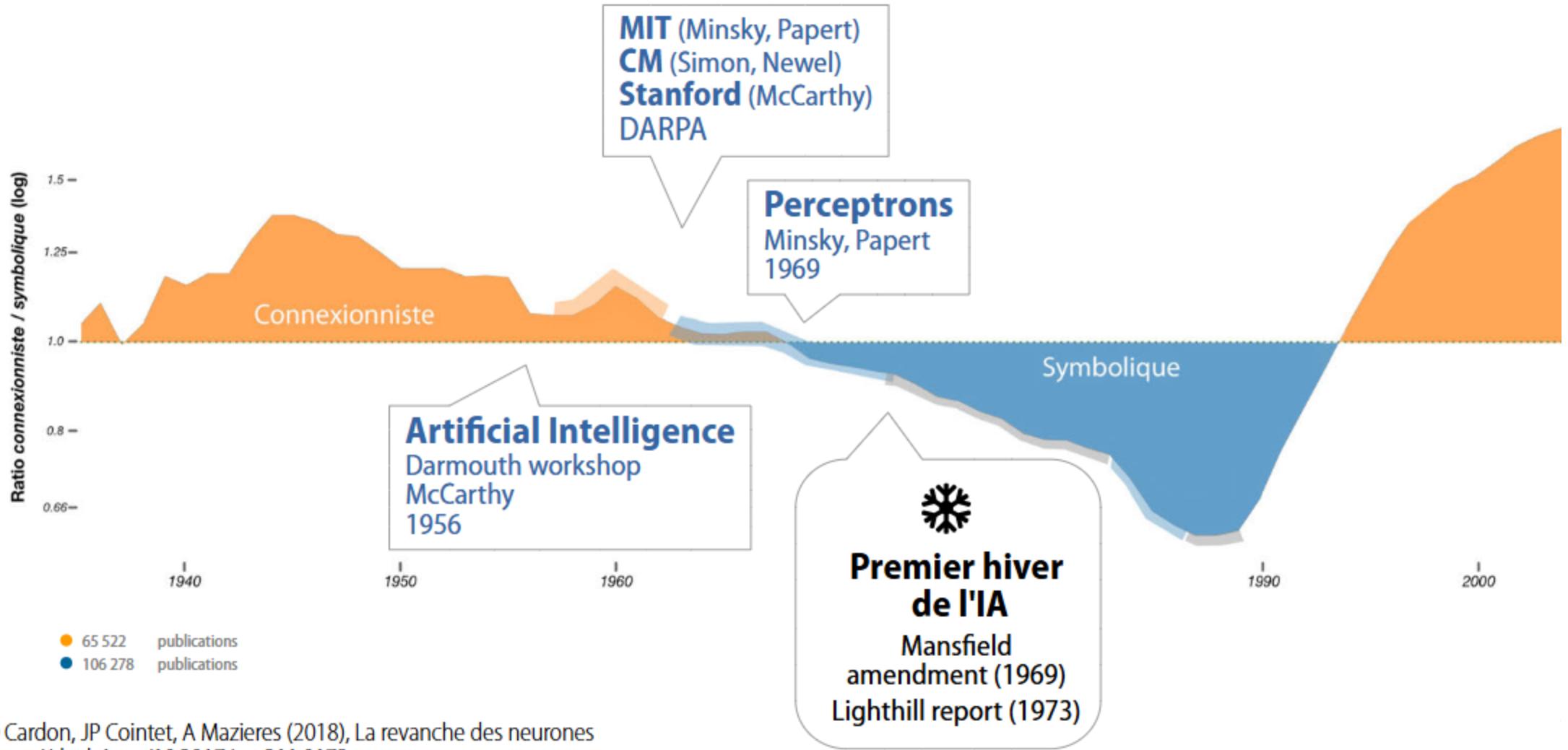
data(iris)

```
data <- iris[iris$Species %in% c("setosa", "versicolor"), ]
```

```
Y <- ifelse(data$Species == "setosa", 1, 0)
```

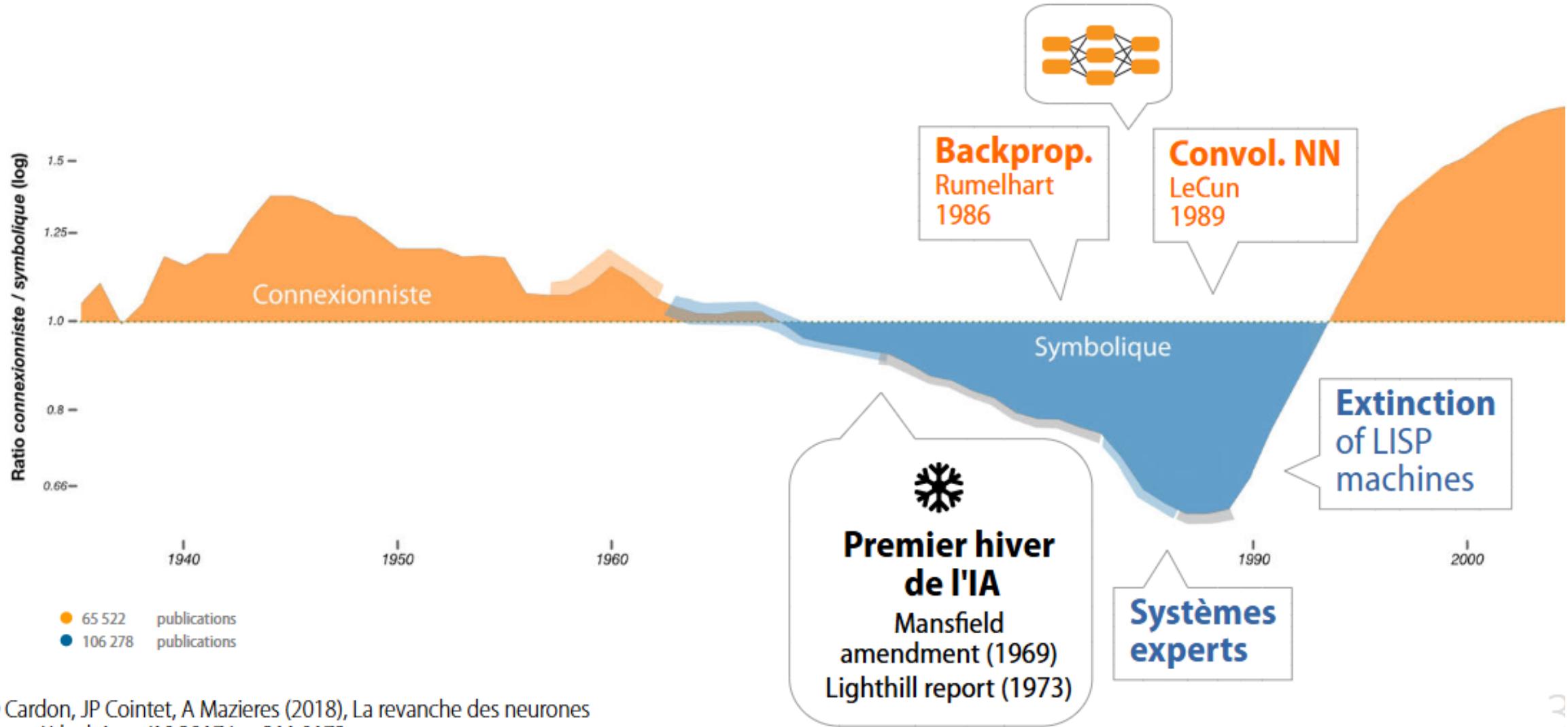
Length	Width	Iris Setosa (0/1)
$x_1$	$x_2$	y
1.4	1.4	1
1.6	1.6	1
1.4	1.4	1
1.5	1.5	1
1.4	1.4	1
4.7	4.7	0
4.5	4.5	0
4.9	4.9	0
4.0	4.0	0
4.6	4.6	0
(...)		

# Influence académique des approches connexionniste et symbolique<sup>1</sup>



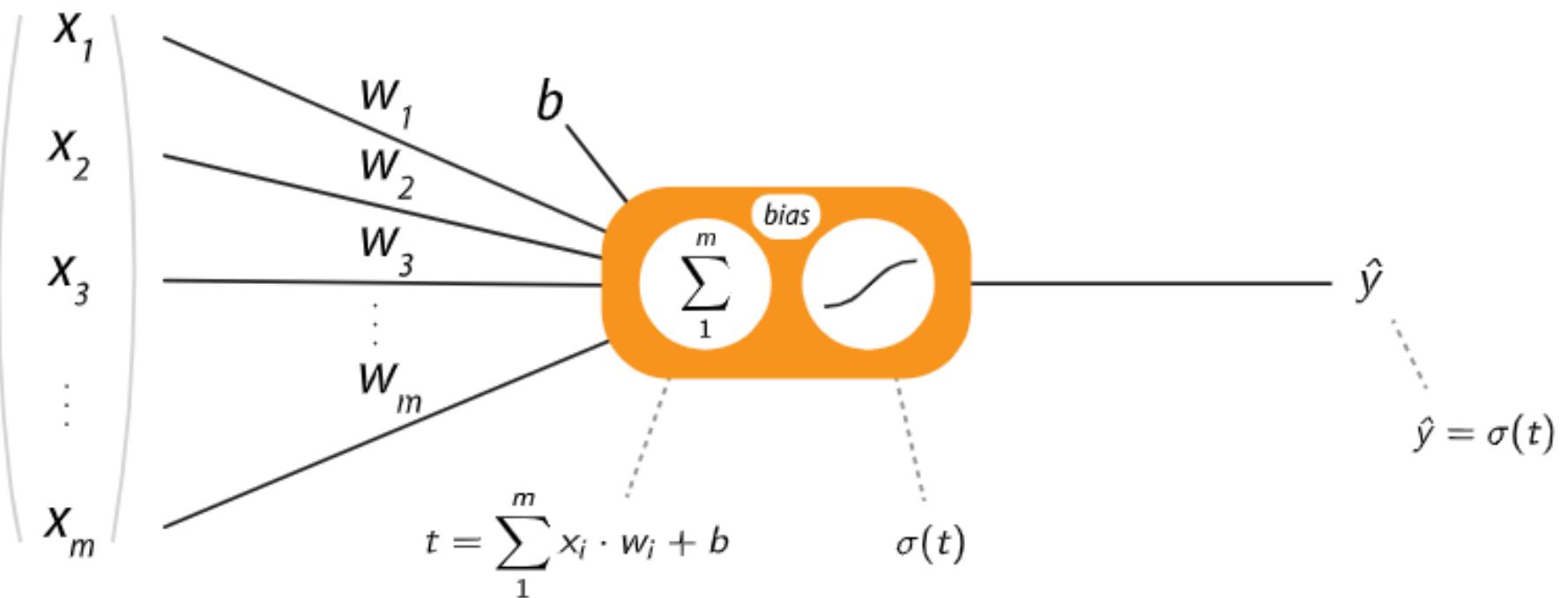
<sup>1</sup> D Cardon, JP Cointet, A Mazieres (2018), La revanche des neurones  
<https://dx.doi.org/10.3917/res.211.0173>

# Influence académique des approches connexionniste et symbolique<sup>1</sup>

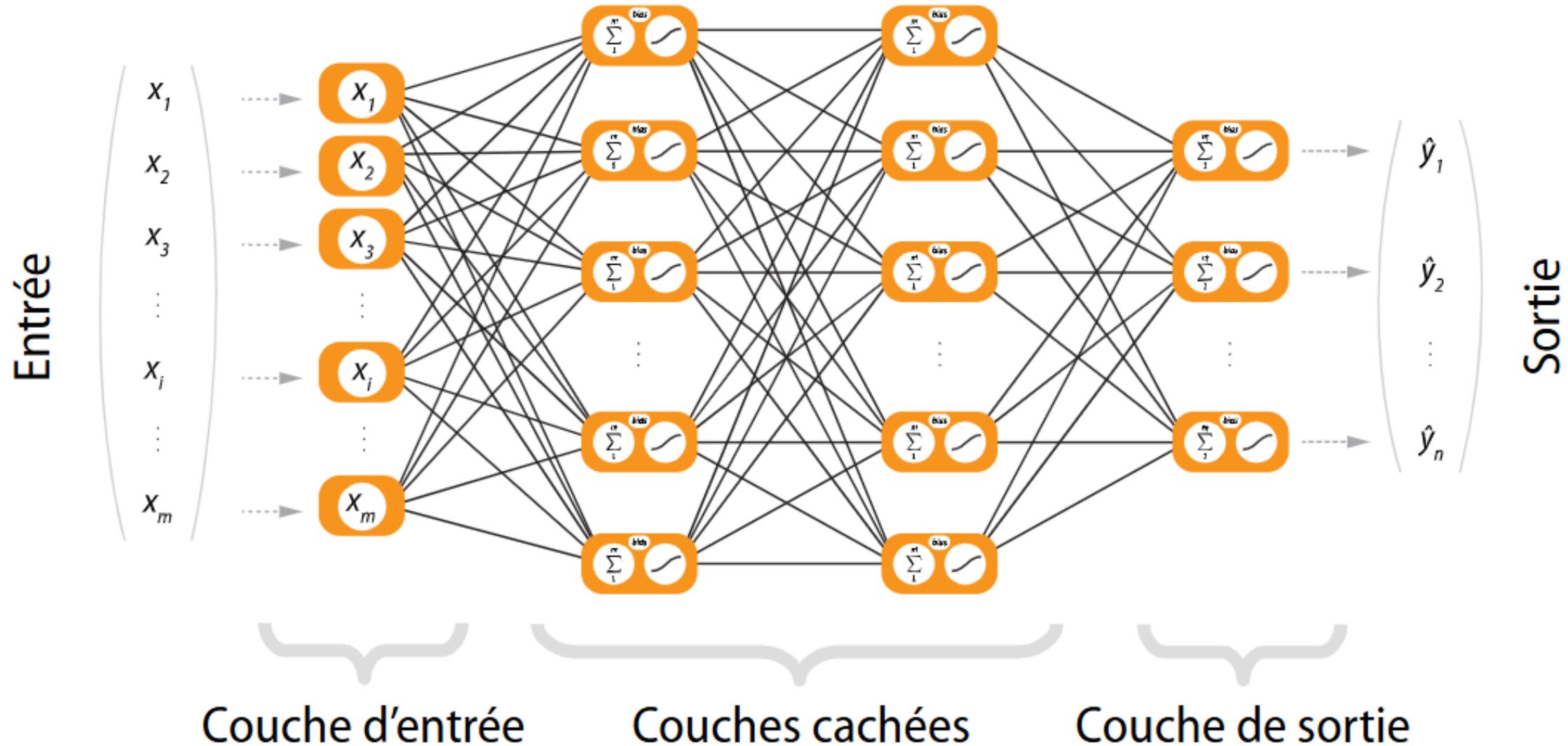


<sup>1</sup> D Cardon, JP Cointet, A Mazieres (2018), La revanche des neurones  
<https://dx.doi.org/10.3917/res.211.0173>

# Deep Neural Networks

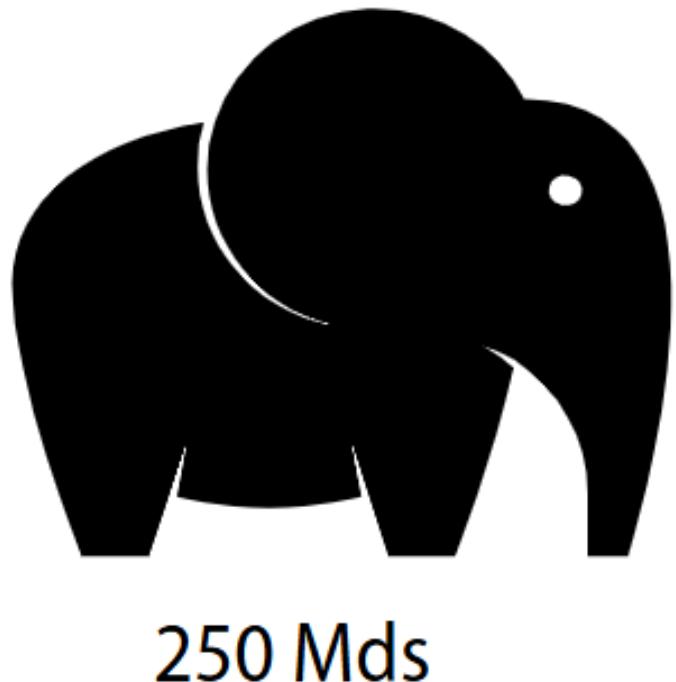
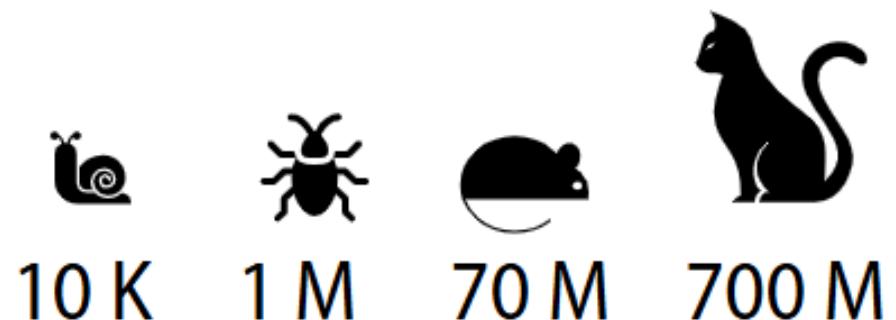
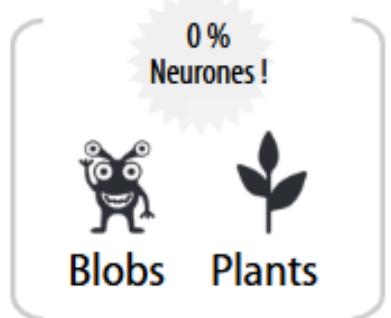


# Deep Neural Networks



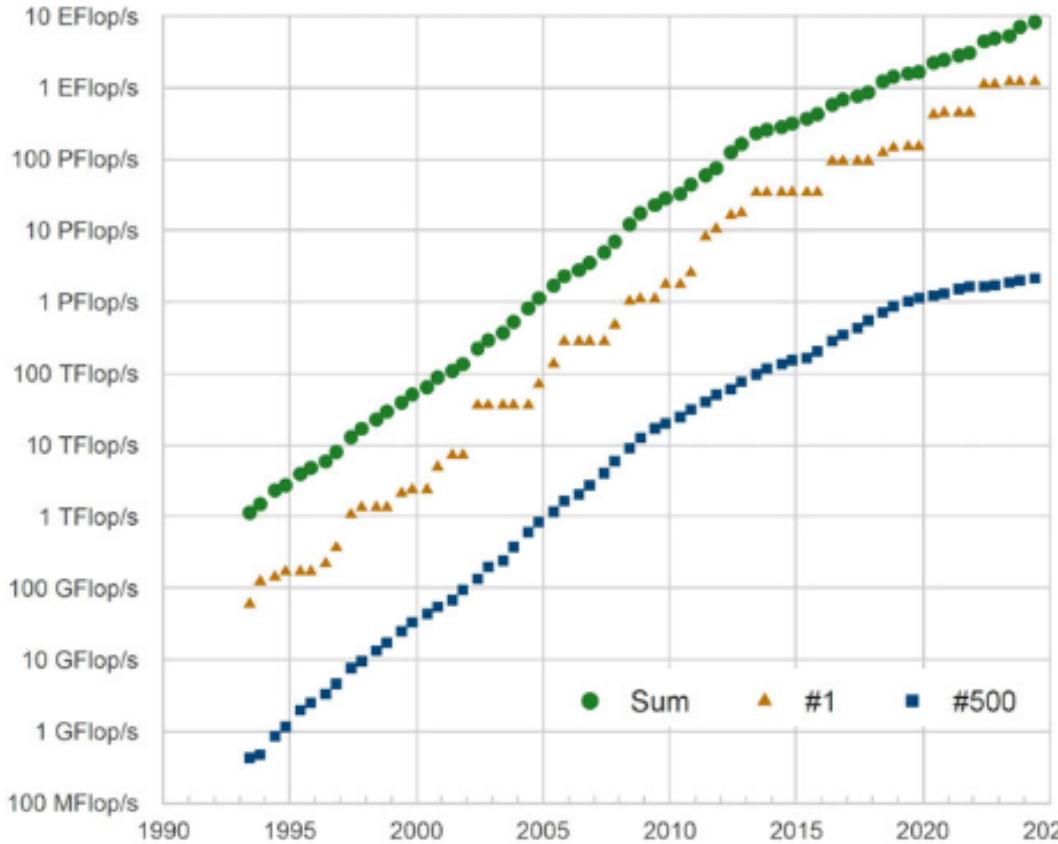
# Combien de neurones pour être intelligent ?

Un neurone, c'est **bien**... mais plusieurs, **c'est mieux** !





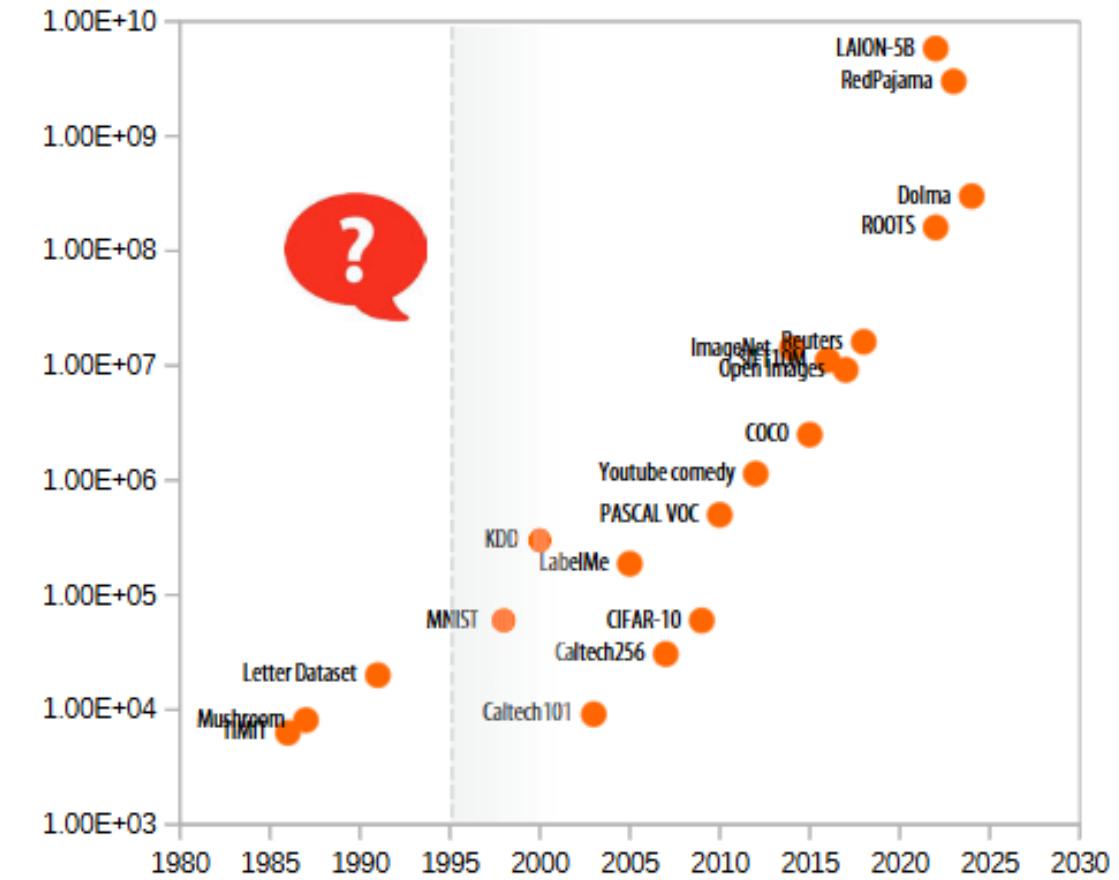
## Performances du Top 500<sup>1</sup>



TOP500 :

<https://www.top500.org/>

## Evolution des Datasets<sup>2</sup>



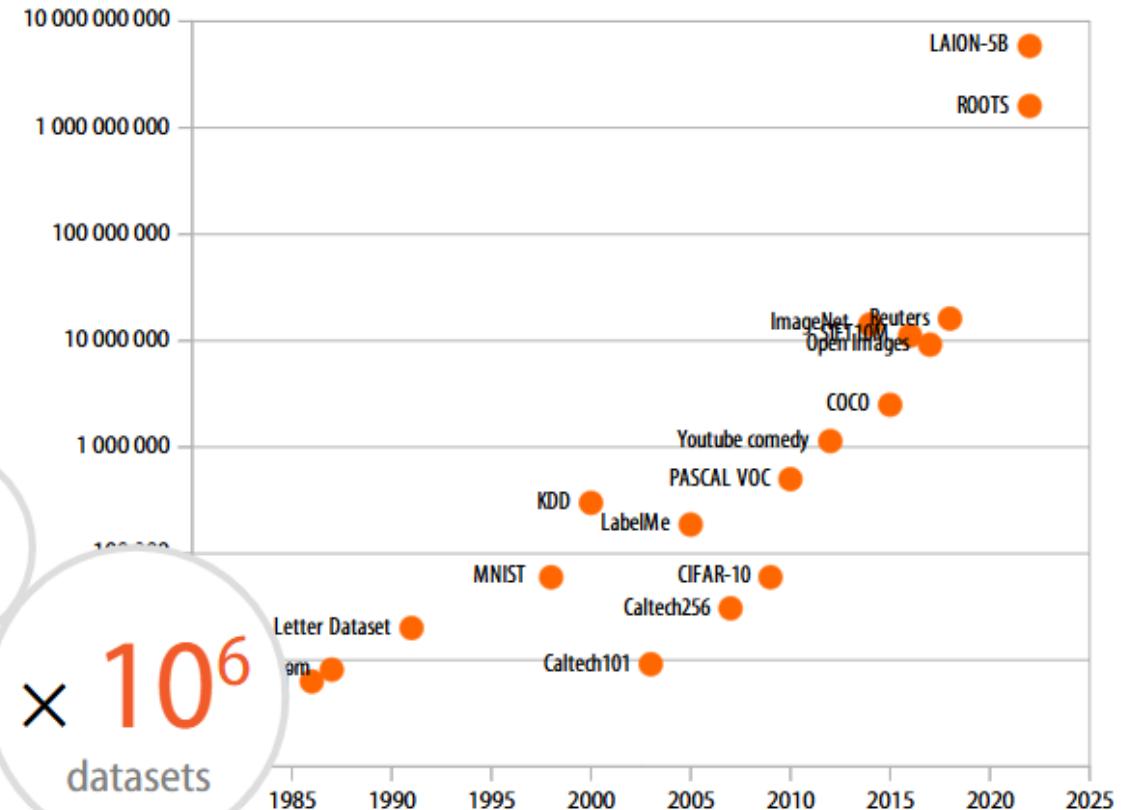
[https://en.wikipedia.org/wiki/List\\_of\\_datasets\\_for\\_machine-learning\\_research](https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research)

# The big Controversy



Laboratoire  
Cas particulier → Monde réel

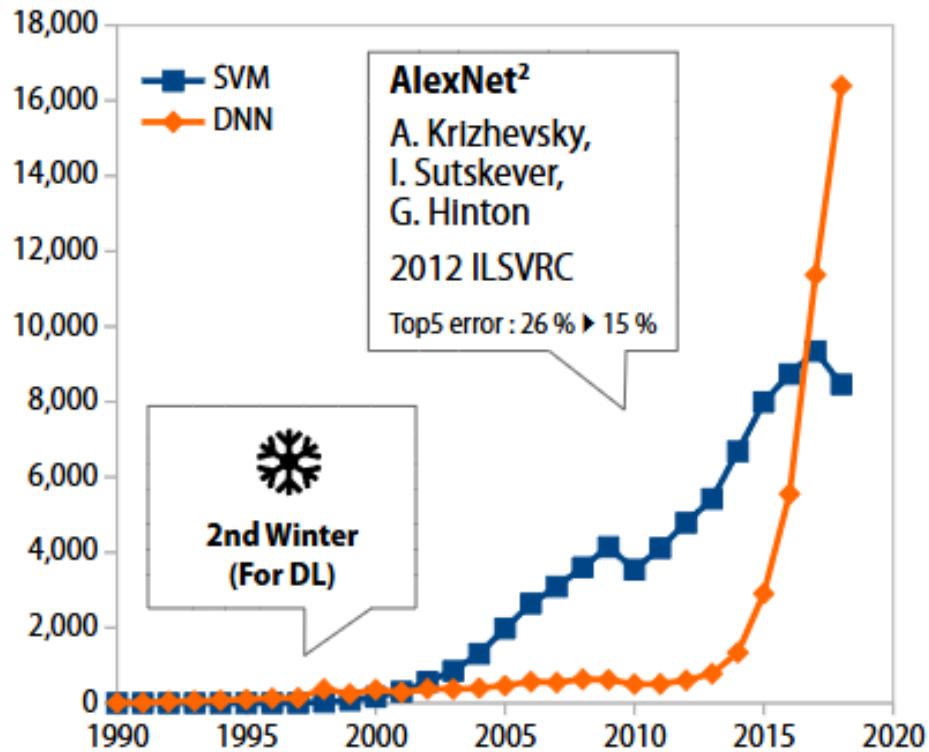
## Datasets for machine-learning<sup>2</sup>



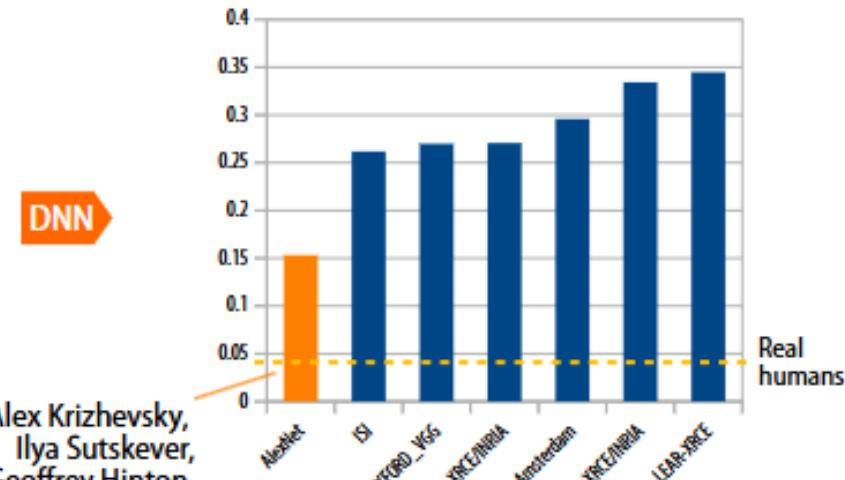
# Le tournant en 2012

The big  
Controversy

## Publications SVM vs DNN<sup>1</sup>



## Classification d'images Top 5 error at ILSVRC 2012<sup>3,4</sup>



Bien que sans garantie mathématique, les DNN se sont révélés bien plus efficaces face à la **complexité du monde réel !**

<sup>1</sup> Web of Science

<sup>2</sup> AlexNet at ILSVRC, <https://image-net.org/challenges/LSVRC/2012/results.html>

<sup>3</sup> ImageNet Large Scale Visual Recognition, <https://www.image-net.org/challenges/LSVRC/>

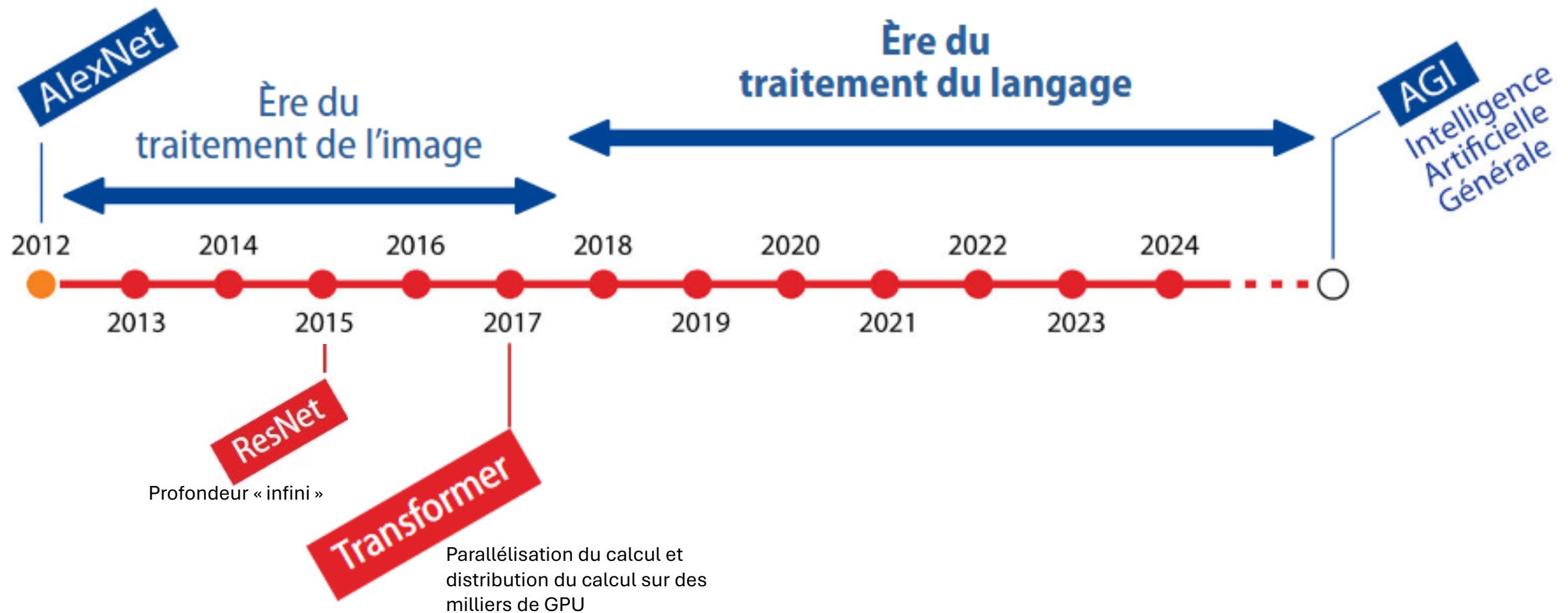
<sup>4</sup> Similar evolution in NLP, translation, board games, etc.  
See : DeepL.com, AlphaGo, AlphaZero, ...

# Histoire du Deep Learning

D'AlexNet vers l'AGI...



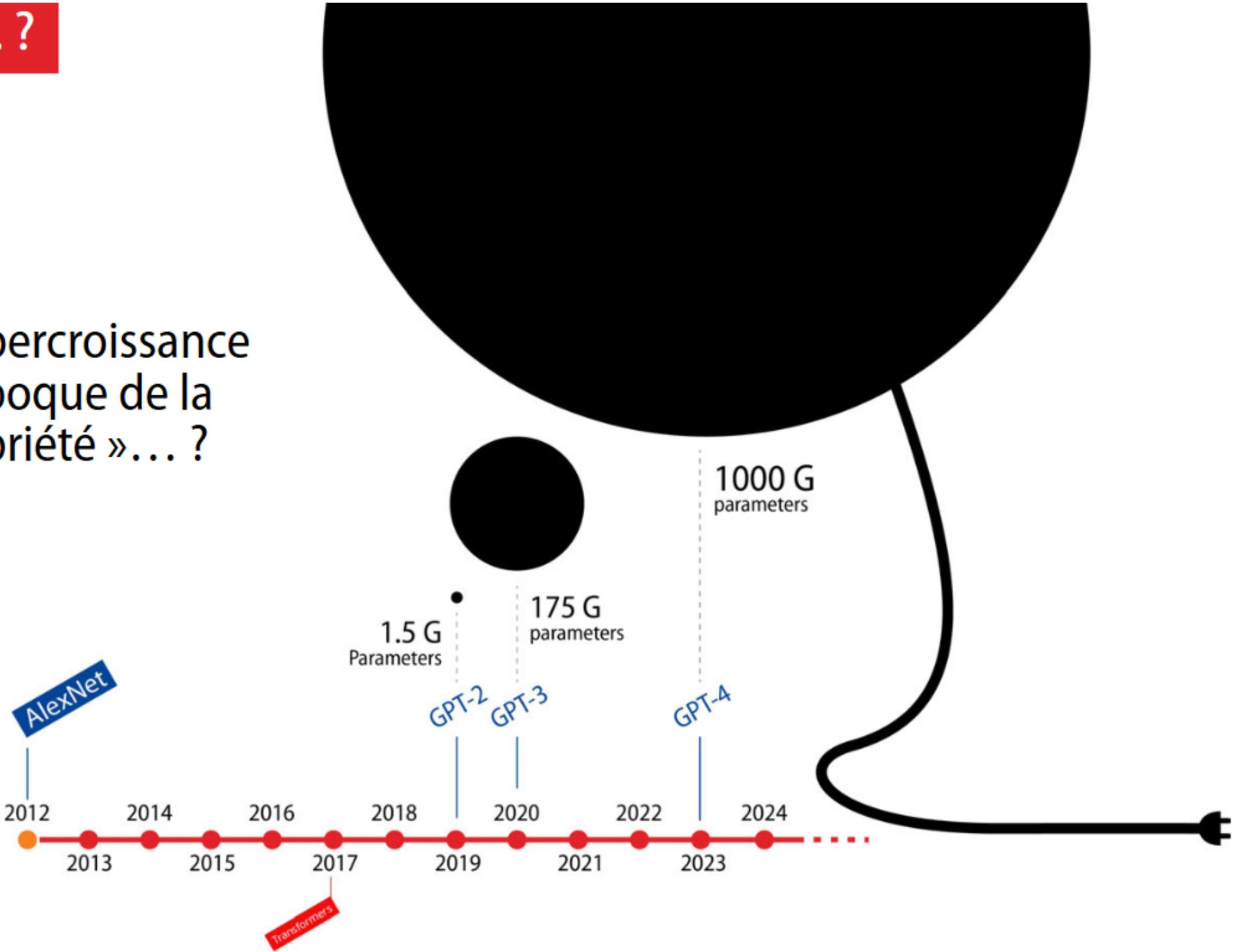
# D'AlexNet vers l'AGI...



# Sobriété... ?



L'hypercroissance  
à l'époque de la  
« sobriété »... ?



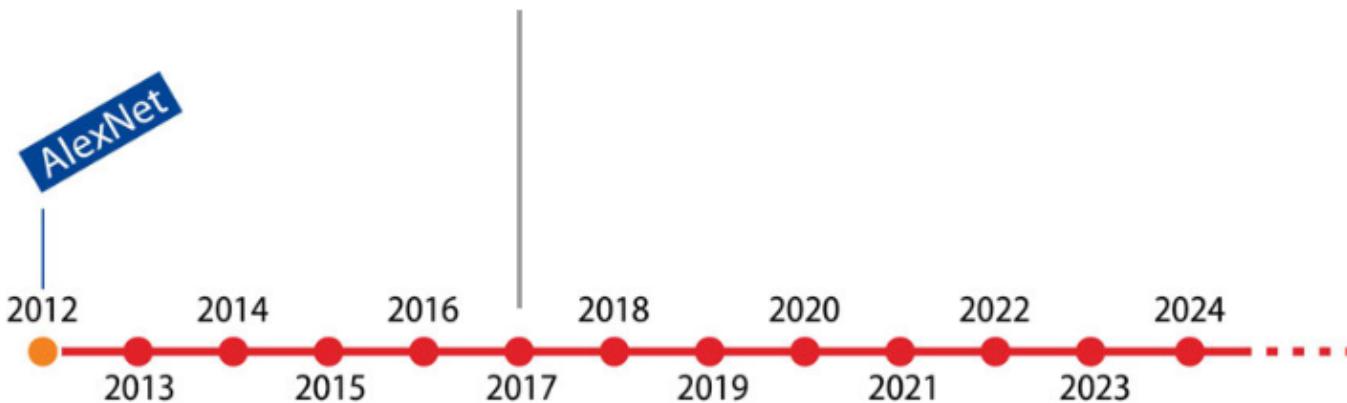


14 Jours  
1 NVIDIA M40 GPU

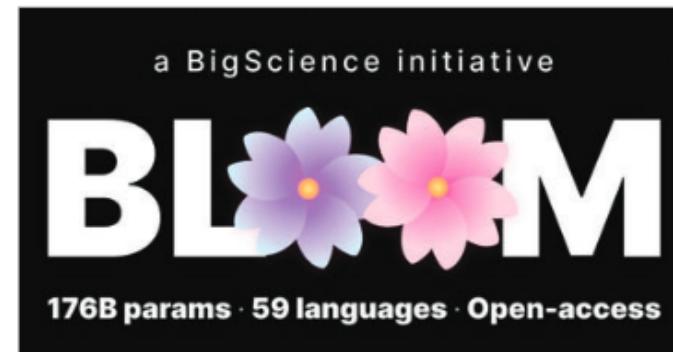
14 Jours GPU

**Classification** (1000 classes)

**25 M paramètres (Resnet-50)  
1,2 M labeled images**



Sobriété... ?



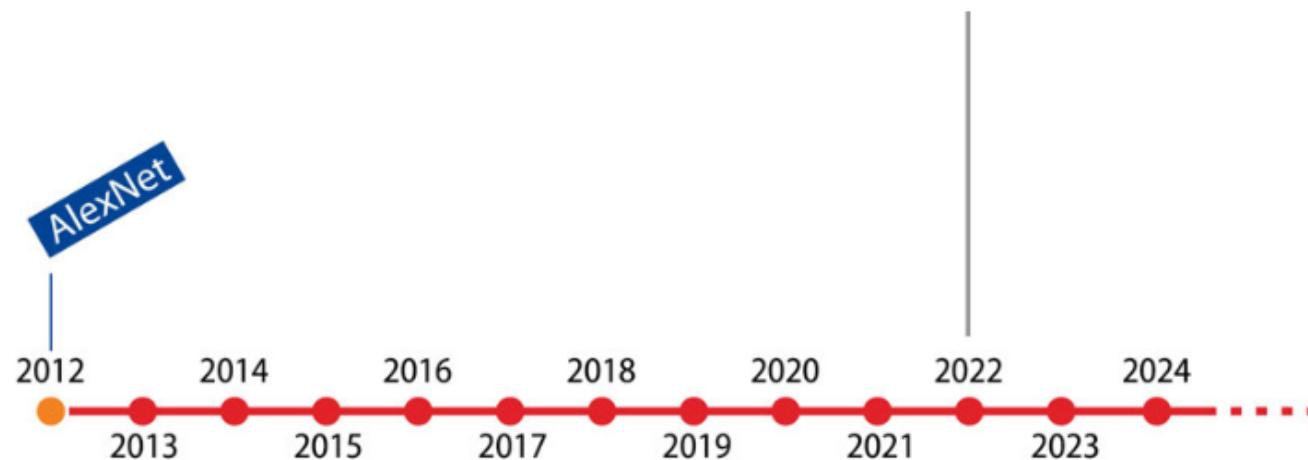
117 Jours  
384 A100 GPU

**44 928 Jours GPU**

## Génération de texte (LLM)

**176 B paramètres**  
**366 B tokens (~3 M livres)**

123 ans



Meta

## Llama 3

### Génération de texte (LLM)

**405 B** paramètres

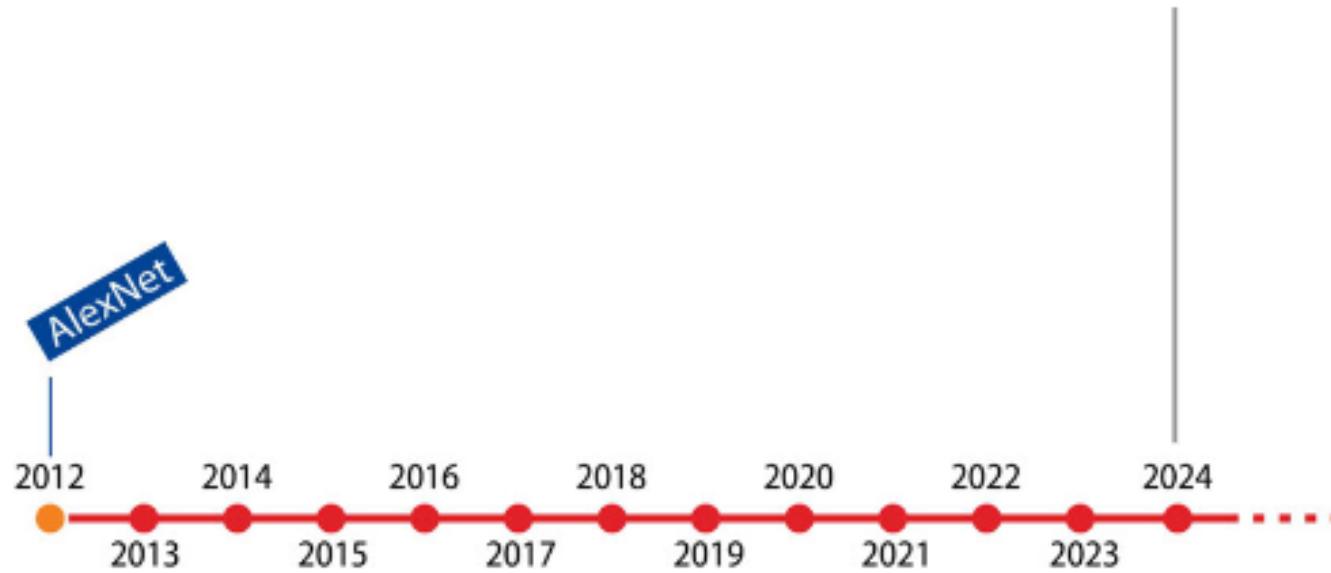
**15 T** tokens (~150 M livres)

54 Jours

16 384 H100 GPU

**884 736 Jours GPU**

2400 ans



# Quand l'IA ...est devenue Hype !!



# Des progrès fulgurants !



# Quand l'IA ...est devenue Hype !!

OpenAI

ChatGPT: Optimizing Language Models for Dialogue

Question ChatGPT va-t-il bouleverser notre monde

Depuis sa mise en ligne, le 30 Janvier 2022, ChatGPT a séduit plus de 100 millions d'utilisateurs dans le monde ! Ce "bot" capable de répondre à toutes les questions fascine et effraie.

1 ChatGPT, comment ça marche ? ChatGPT fonctionne comme un robot programmé pour dialoguer de la façon la plus humaine possible. C'est un modèle de langage entraîné par une intelligence artificielle et capable d'interagir avec les humains de manière à accomplir une mission scientifique complexe ou décrire une recette délicieuse.

2 Qui se cache derrière cette innovation ? Une startup californienne, OpenAI, fondée en 2015 par Sam Altman, et un certain Elon Musk ! Depuis, on connaît les succès de l'IA utilisée pour développer les voitures autonomes d'Elon et la cryptomonnaie Meme (SpaceX).

3 A-t-il à tout prix à battre ? ChatGPT est devenu rapidement l'un des meilleurs assistants virtuels au monde, grâce à une interface conviviale et intuitive qui permet de répondre à des questions complexes et variées.

4 Dites-nous ce que vous pensez de l'avenir de l'IA et de ses impacts sur la société. Nous sommes curieux de vos avis !



The New York Times

A New Chat Bot Is a 'Code Red' for Google's Search Business

A new wave of chat bots like ChatGPT use artificial intelligence that could reinvent or even replace the traditional internet search engine.

What is ChatGPT? The AI chatbot talked up as a potential Google killer

After all, the AI chatbot seems to be staying a great deal of search engine

BBC NEWS FOX BUSINESS

'Google killer' ChatGPT spars in AI chatbot race

POTENTIAL GOOGLE KILLER COULD CHANGE US WORKFORCE

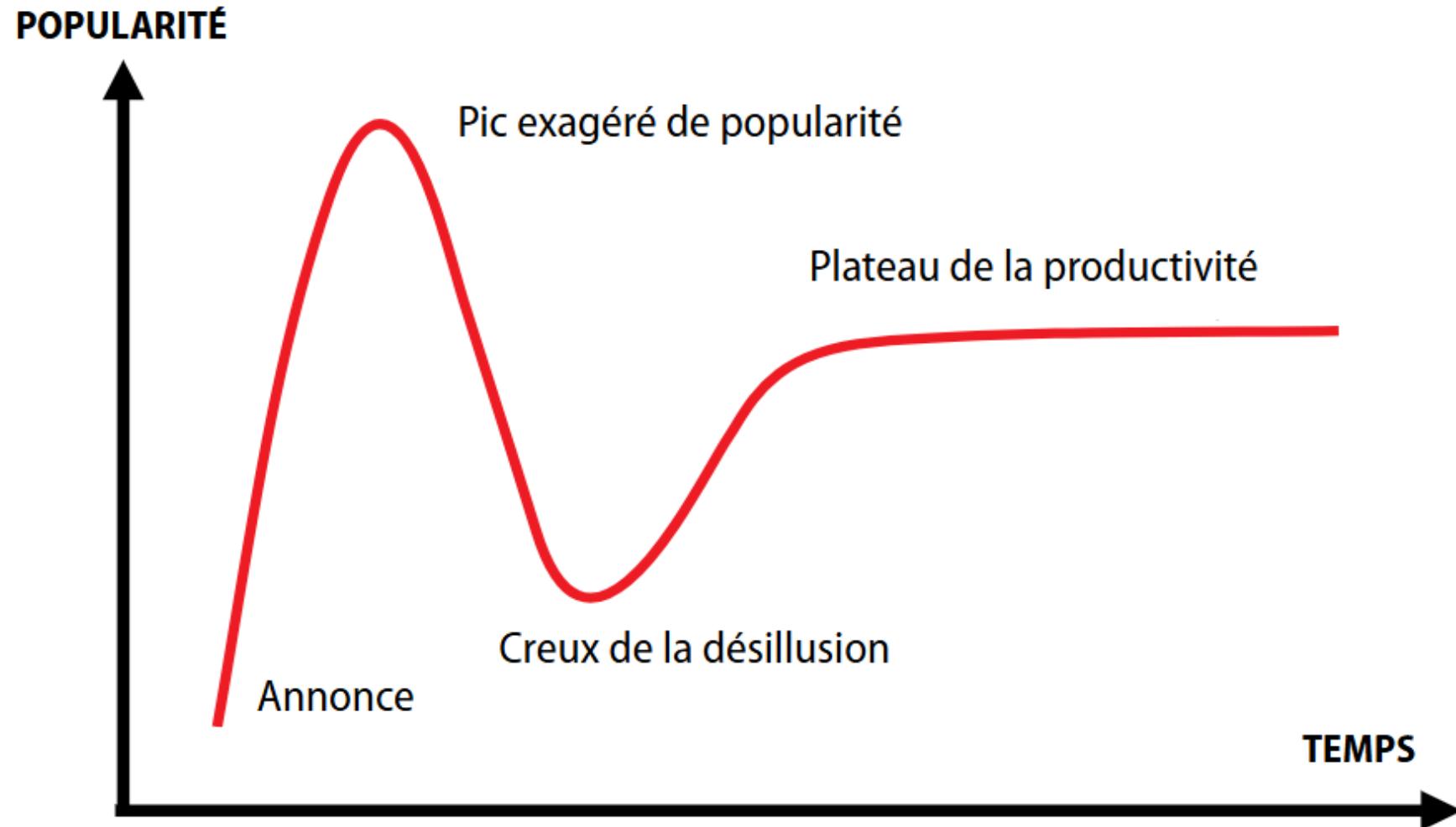
ARTIFICIAL INTELLIGENCE Published January 29, 2023 09:00 ET

ChatGPT fuels concern over AI's impact on jobs

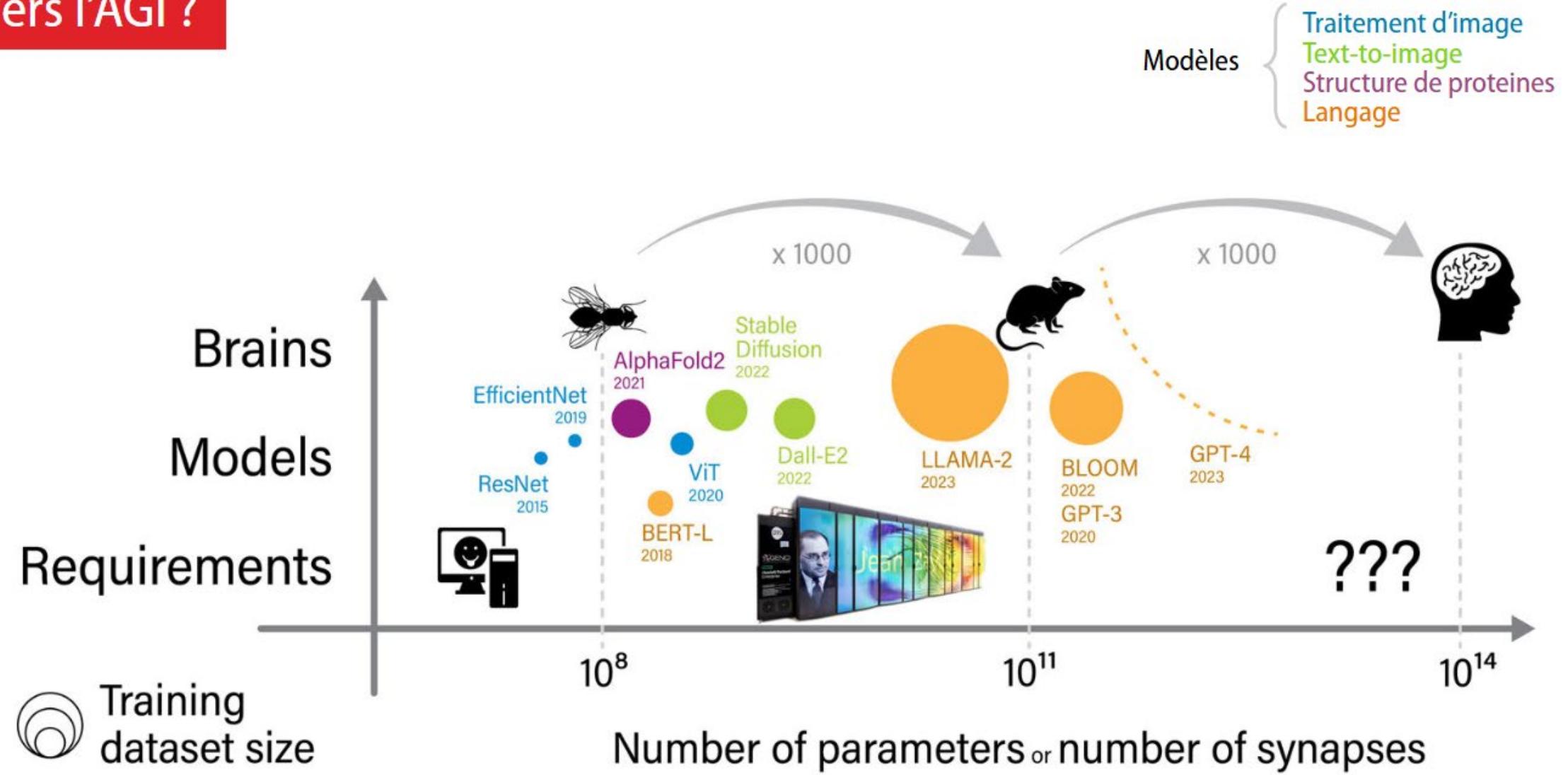
Source: Bloomberg

Bloomberg Opinion

# Quand l'IA ...est devenue Hype !!



# Vers l'AGI ?



Des datasets de plus en plus gros, atteignant les limites de l'univers numérique !!

## L'enjeu des données...

Les données sont essentielles...

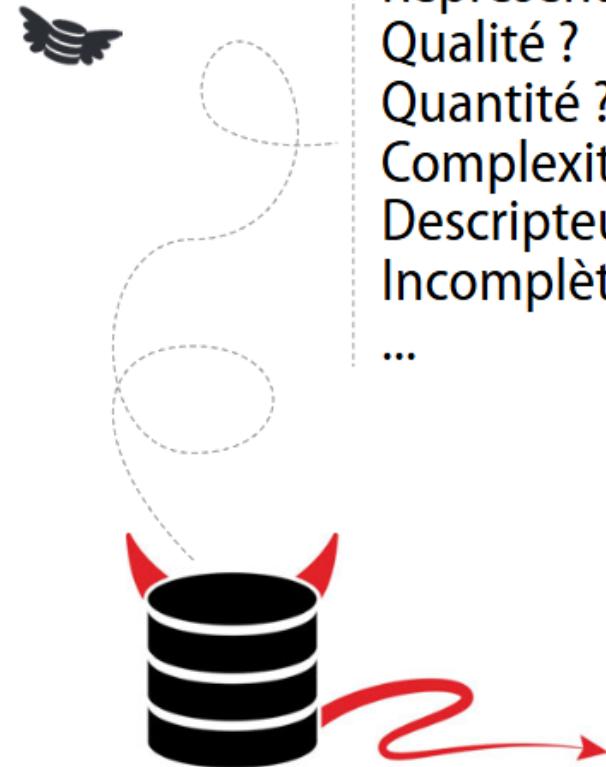
Elles peuvent être :



Fermée



Ouvertes



Trop souvent  
infernales !!

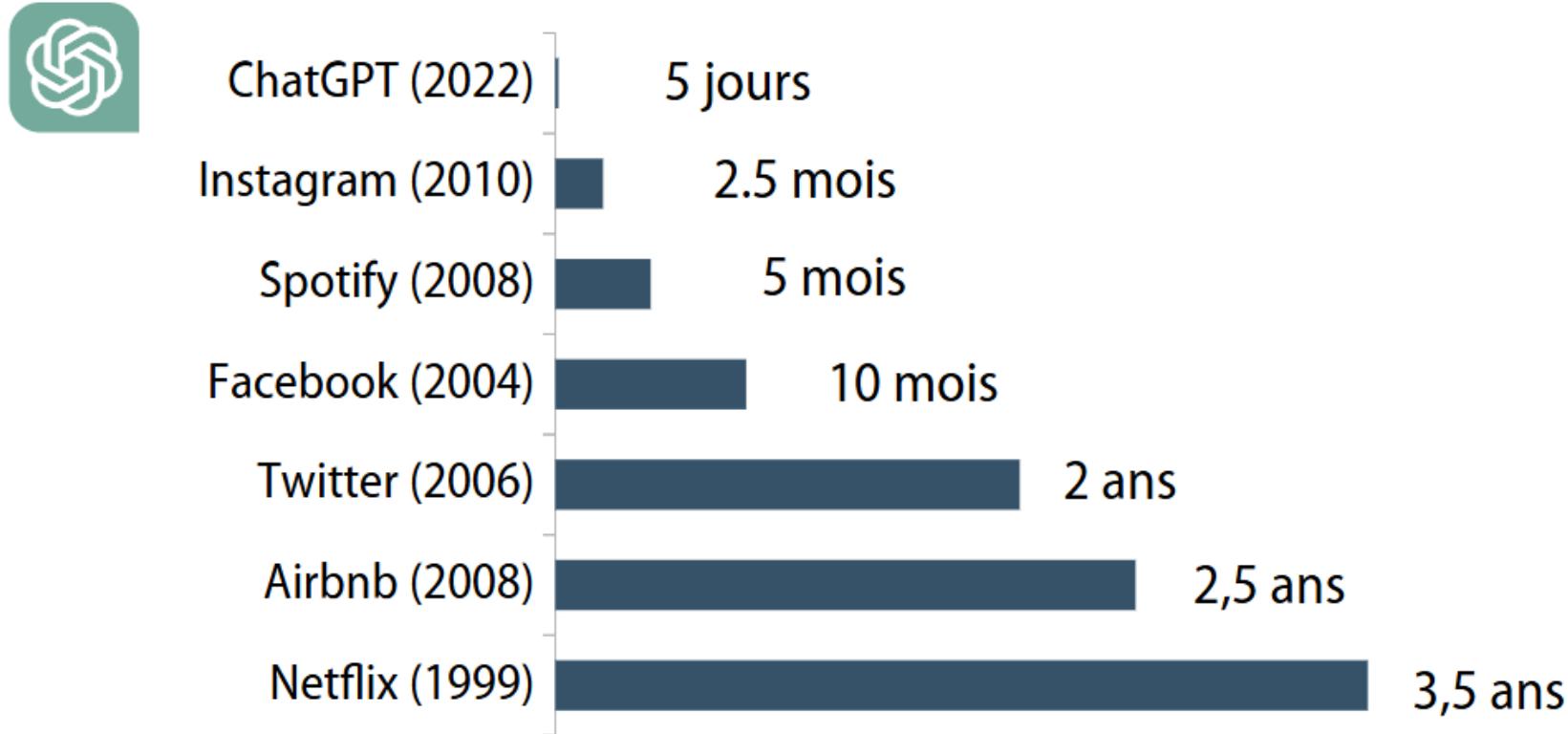
Représentativité ?  
Qualité ?  
Quantité ?  
Complexité ?  
Descripteur ?  
Incomplètes ?  
...

# Apropiation... ?



Une appropriation sociale sans précédent.

Temps nécessaire pour d'atteindre 1 million d'utilisateurs :





Par exemples :



Biais ?

Explicabilité ?

Impacts sociaux et environnementaux ?

Syndrome MAD\* ?

Transposition dans le monde réel ?

Capacité à suivre ?

Etc.



>9,000 publications par mois...

# Références :

- IA générative : usages et applications avancées pour la veille technologique – YouTube
- L'intelligence artificielle, objet philosophique
- Pour une introduction au Deep Learning FIDLE
- **L'IA et ses défis :** <https://www.college-de-france.fr/fr/agenda/colloque/ia-et-ses-defis>
- Implications philosophiques de l'IA
- Guide pratique de l'IA à l'Inspé
- **Guide pratique de l'intelligence artificielle dans l'entreprise (Stéphane RODER - Eyrolles)**
- guide d'accompagnement : Décryptage des enjeux et opportunités de l'IA pour les entreprises et les salariés
- Guide pratique de l'université de Genève
- Comment les machines pourraient-elles atteindre l'intelligence humaine? Conférence de Yann LeCun