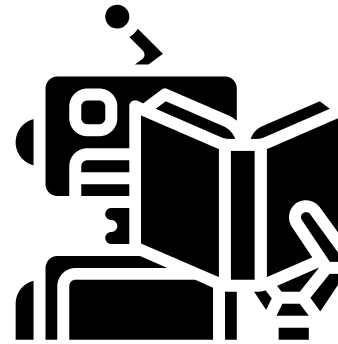
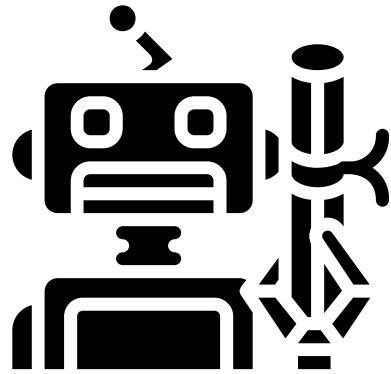


# Introduction rapide au *Machine Learning*

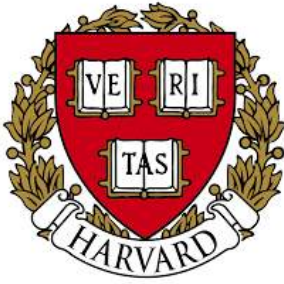
Démystification & exemples



# Teaser




<http://enqueteur.cgdd.developpement-durable.gouv.fr/index.php?sid=49623&lang=fr>




**nature**  
International journal of science

Letter | Published: 29 August 2018

# Deep learning of aftershock patterns following large earthquakes

Phoebe M. R. DeVries , Fernanda Viégas, Martin Wattenberg & Brendan J. Meade

*Nature* **560**, 632–634 (2018) | [Download Citation](#) 

Une intelligence artificielle pour prédire les répliques de séismes, développée sans sismologue, et sur le papier plus efficace que les modèles métiers.

« Les grands acteurs du numérique [...] collectent des données de manière **non scientifiquement contrôlée**, grâce à des capteurs, au recueil de traces d'utilisation, ou encore à la contribution des internautes. **Elles n'ont ni la robustesse, ni la complétude des données scientifiques.** Elles créent cependant progressivement, de par leur seul volume, une forme d'empreinte du réel qui peut à son tour être interprétée et être utilisée produire un **savoir activable.** »



# Au-delà du buzzword, le *machine learning*, c'est quoi ?

« Contrairement à la programmation symbolique, le *machine learning* nourrit le système de données, qui apprend seul de ses actions, sans lui avoir imposé des règles sur un mode itératif »



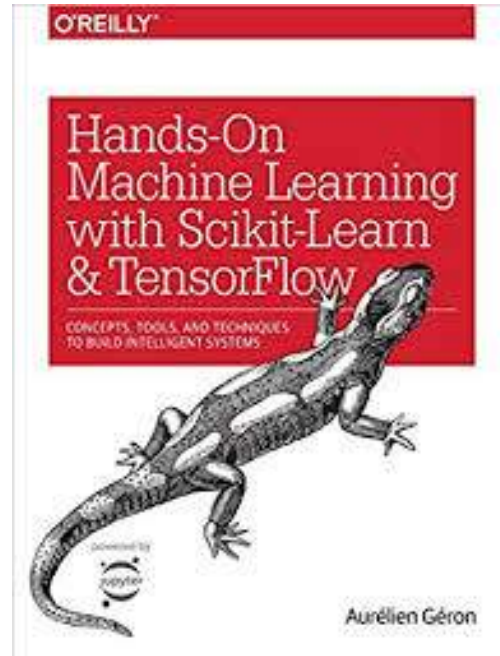
"Le *machine learning* est un ensemble d'outils statistiques ou géométriques et d'algorithmes informatiques qui permettent d'automatiser la construction d'une fonction de prédiction  $f$  à partir d'un ensemble d'observations que l'on appelle l'ensemble d'apprentissage." Big Data et Machine Learning, Lemberger et al. , Dunod

« Algorithm that teach a computer to perform a task from experience » Igor Halperin (Université de New York)



# Au-delà du buzzword, le *machine learning*, c'est quoi ?

« *Machine learning is the science (and art) of programming computers so they can **learn from data*** »

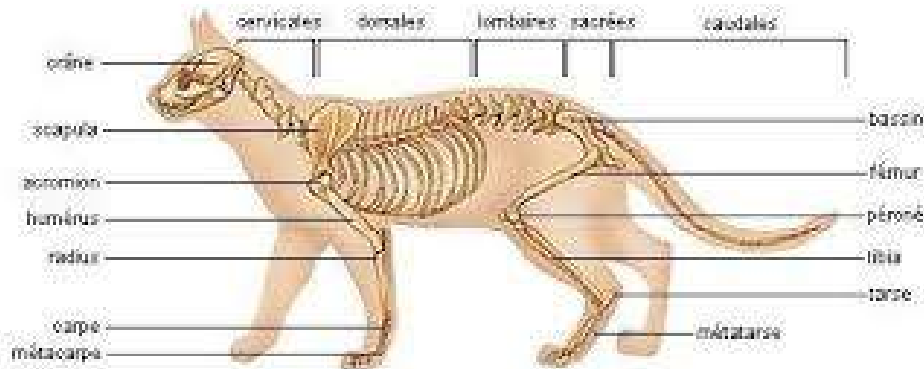


« ML is the field of study that gives computers the ability to learn **without being explicitly programmed** » (Arthur Samuel, 1959)

# Au-delà du buzzword, le *machine learning*, c'est quoi ?

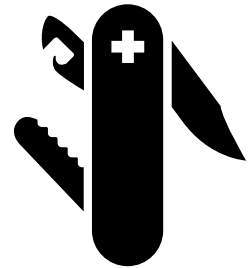
- Apprendre à partir de données puis généraliser !
- Apprendre sans programmation explicite !
- Pour estimer des fonctions trop complexes à écrire de manière explicite.

Exemple dans 20 minutes : détection d'images de chat sans n'avoir jamais écrit ce qu'est un chat, juste en montrant des images de chats à la machine.



# Le machine learning c'est aussi et surtout

- Un peu de stats
- Un peu de mathématiques appliquées (et un max d'algèbre linéaire)
- Un peu de proba
- Un peu de théorie de l'information
- Un peu d'informatique
- Un peu de neurosciences, de biologie, de physique ...
- Un peu de bidouille ...
- Des essais/erreurs : il faut pratiquer !





*"Dans sa forme la plus utilisée, l'apprentissage machine est supervisé : on montre en entrée de la machine une photo d'un objet, par exemple une voiture, et on lui donne la sortie désirée pour une voiture. Puis on lui montre la photo d'un chien avec la sortie désirée pour un chien. Après chaque exemple, la machine ajuste ses paramètres internes de manière à rapprocher sa sortie de la sortie désirée. Après avoir montré à la machine des milliers ou des millions d'exemples étiquetés avec leur catégorie, la machine devient capable de classifier correctement la plupart d'entre eux. Mais ce qui est plus intéressant, c'est qu'elle peut aussi classifier correctement des images de voiture ou de chien qu'elle n'a jamais vues durant la phase l'apprentissage. C'est ce qu'on appelle la capacité de généralisation."*

UDACITY TALKS

Yann LeCun

Director of AI Research

facebook

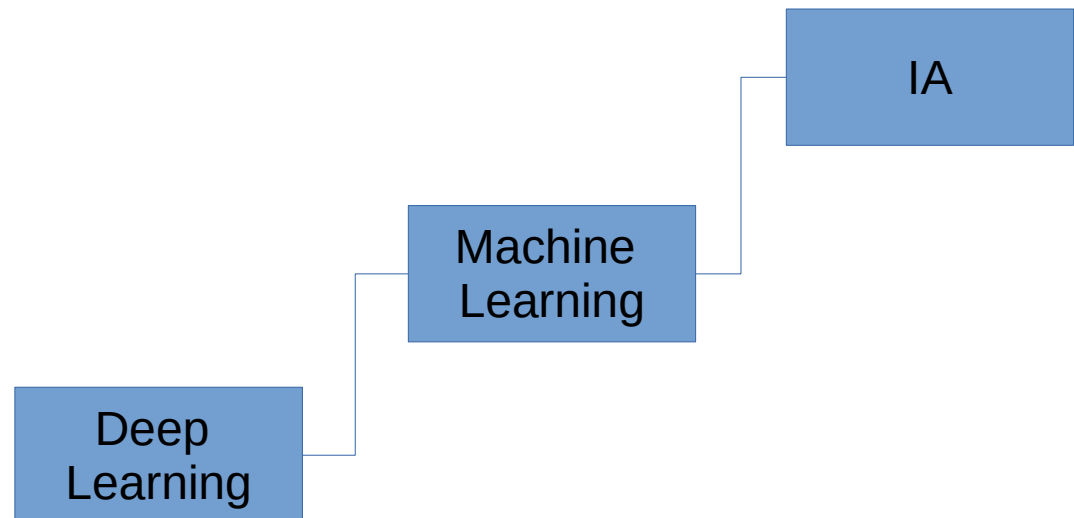


# Et l'IA dans tout ça ?

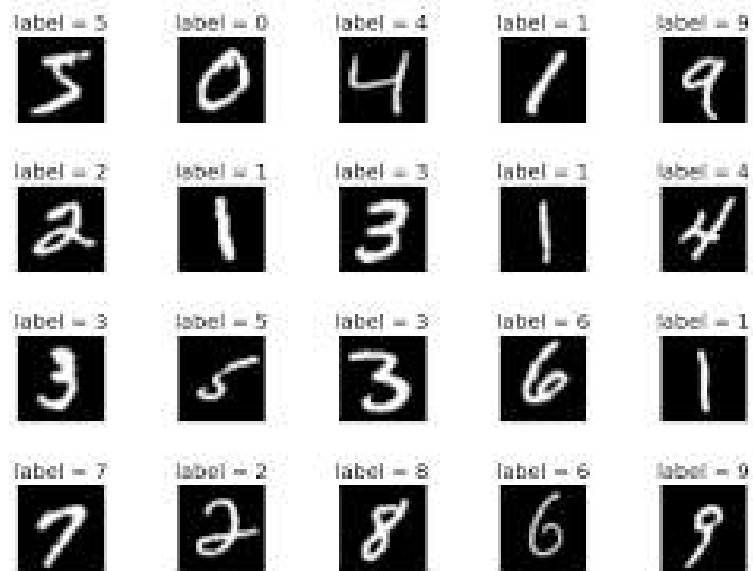
Le *machine learning* est une branche de l'IA

# Et le deep learning dans tout ça ?

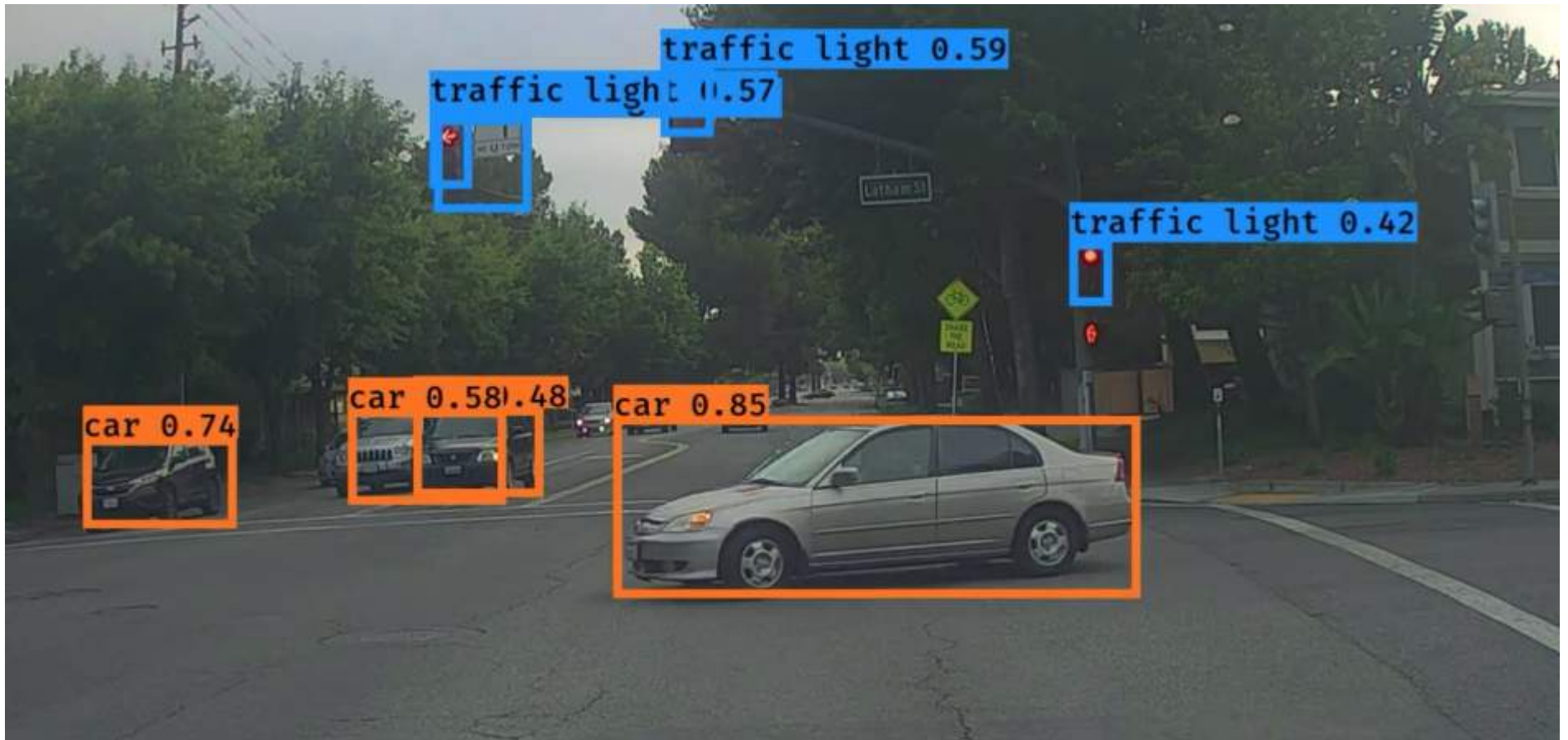
Le *deep learning* (réseau de neurones) est un branche du *machine learning*



# Et ça marche ?



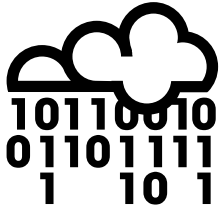
# Et ça marche ?



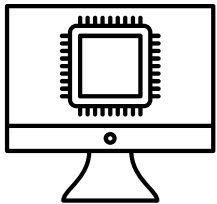
Source : [deeplearning.ai](https://deeplearning.ai) / [drive.ai](https://drive.ai)

# Le *machine learning* n'est pas forcément nouveau, pourquoi ce buzz ?

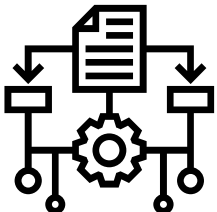
*Il y a 20 ans, on parlait déjà des réseaux de neurones ...*



Production et disponibilité exponentielles de données (IoT, open data, smartphone ...) et de plus en plus de données « étiquetées »



Augmentation significative des puissances de calcul

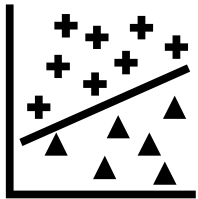


Des progrès algorithmiques récents

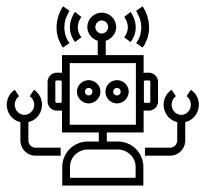
# Schématiquement (et pour simplifier), 3 types de machine learning



Apprentissage supervisé



Apprentissage non supervisé

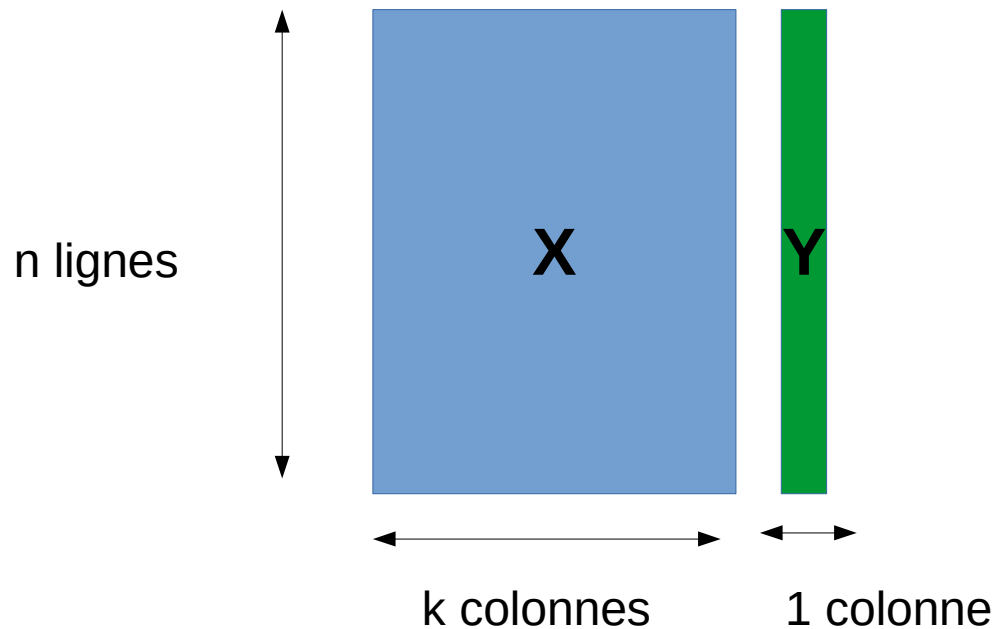


Apprentissage par renforcement



# Apprentissage supervisé

Les données (X) sont accompagnées des « solutions », appelées *labels* (Y)

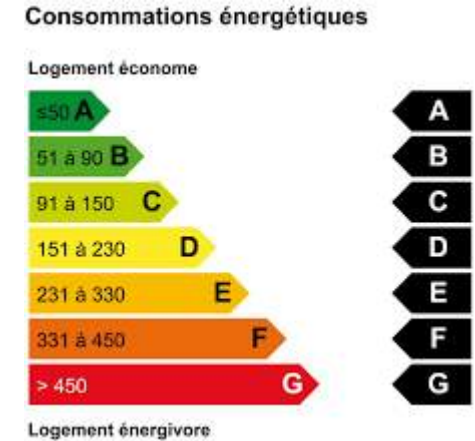


« L'erreur commise par le modèle est connue *a priori* par l'algorithme. » Pierre-Cyril Aubin

**Machine learning :**

Trouver une fonction  $f$  telle  
que  $f(X) \approx Y$

# Exemple, les DPE



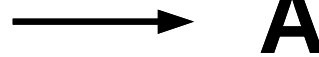
1940  
Fuel  
3 étages



**E**



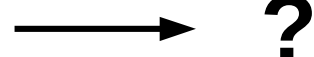
2018  
Elec. solaire  
2 étages



**A**



1970  
Bois  
4 étages



**?**



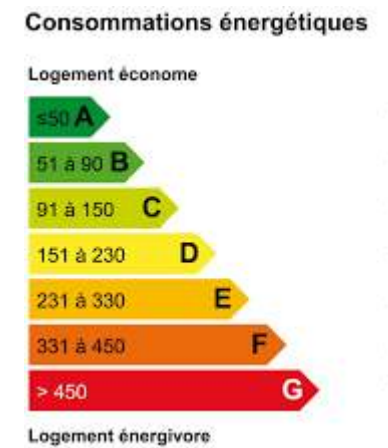
2005  
Gaz  
3 étages



**C**

## Classification

# Exemple, les DPE



1940  
Fuel  
3 étages

→ 320 kWh<sub>ep</sub>/m<sup>2</sup>.an



2018  
Elec. solaire  
2 étages

→ 35



1970  
Bois  
4 étages

→ ?



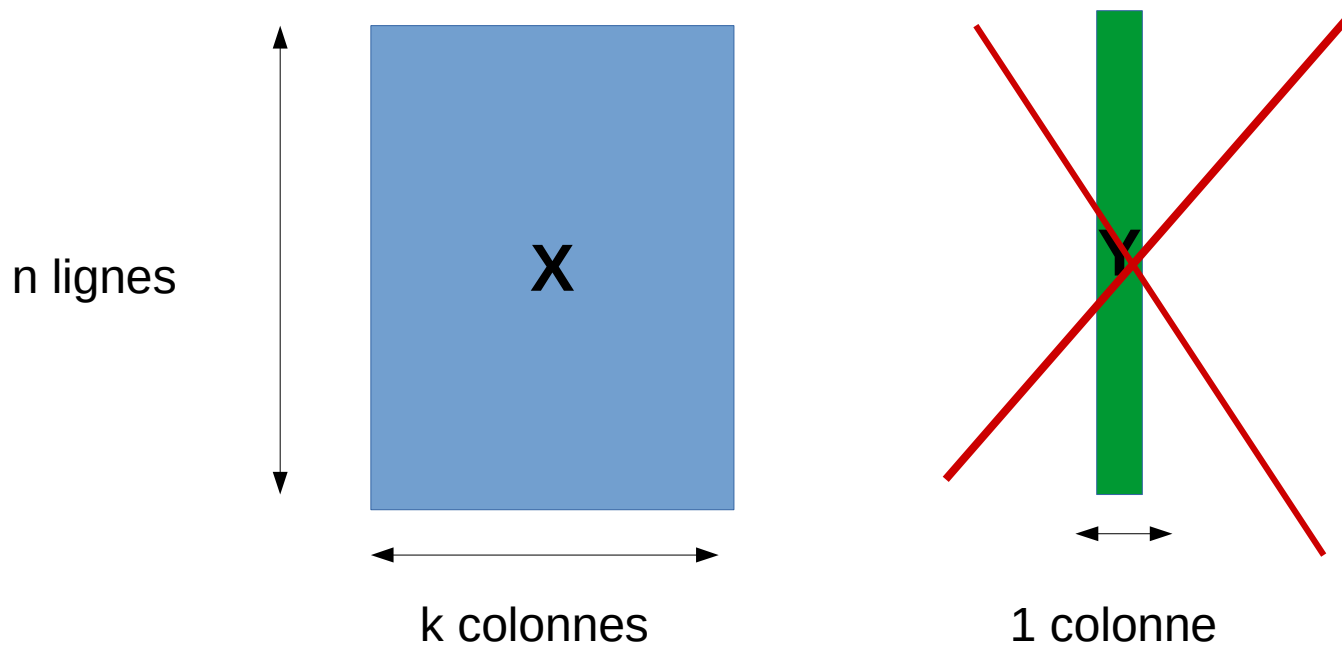
2005  
Gaz  
3 étages

→ 149

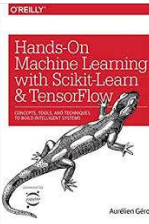
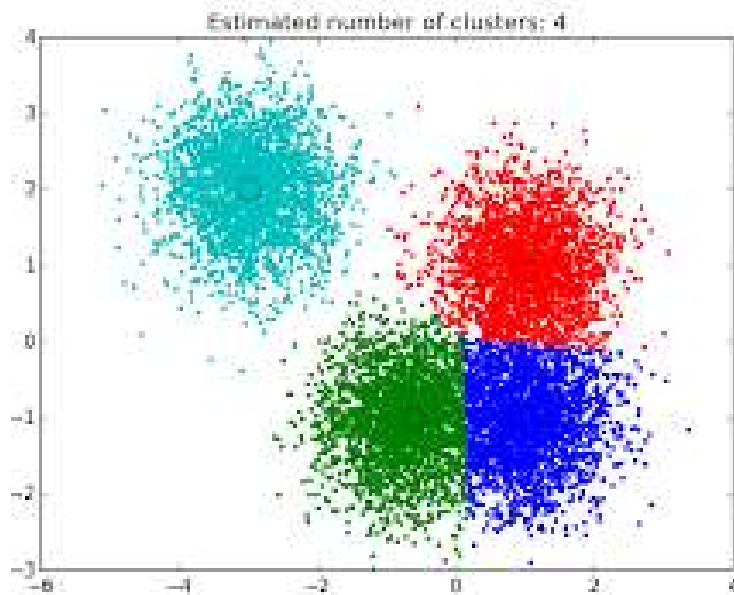
## Régression

# Apprentissage non supervisé

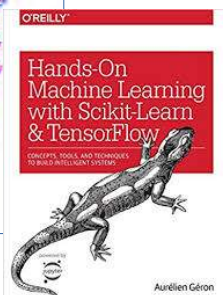
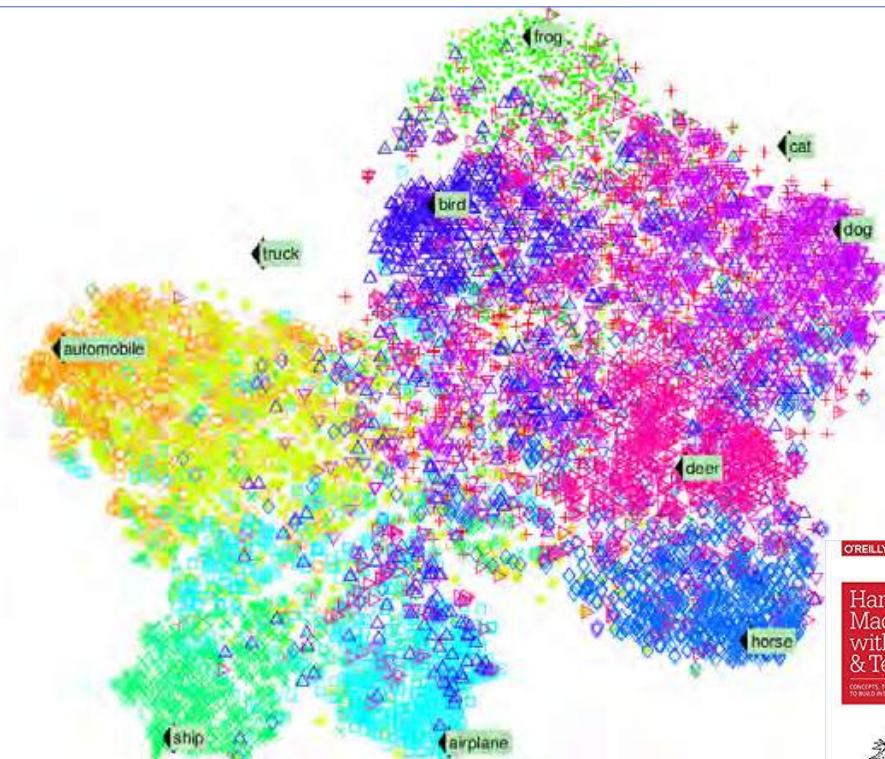
Les données ne sont pas accompagnées des « solutions » ...



Clustering, réduction de dimensions, représentation de données à grande dimension, détection d'anomalies, découverte d'associations (ex: achat combiné de couches et de bière etc.). ...

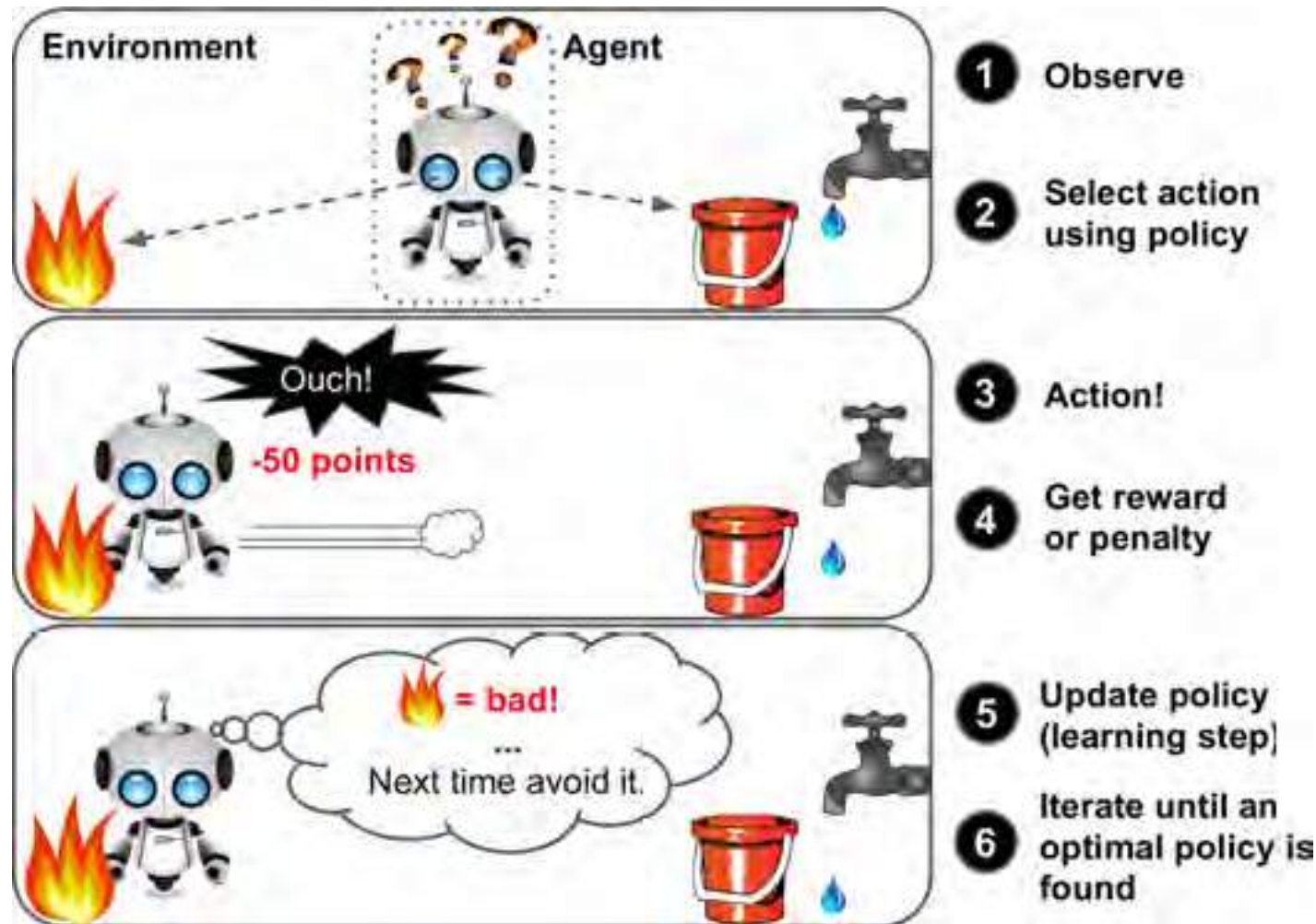


- + cat
- o automobile
- truck
- frog
- ship
- airplane
- ◊ horse
- △ bird
- ▽ dog
- ◀ deer

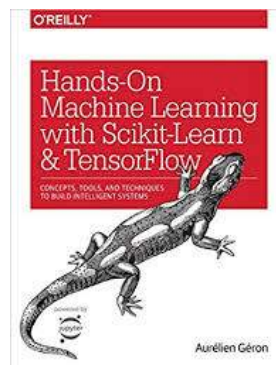


# Apprentissage par renforcement

Il y a un « superviseur » mais il ne donne qu'un feedback partiel (récompenses selon les actions)



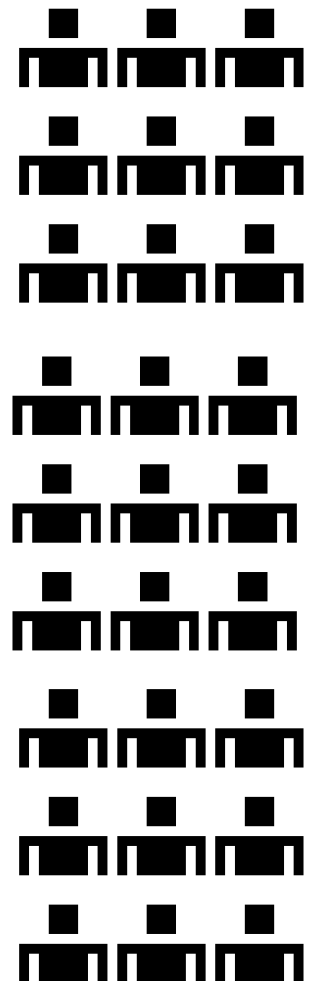
« L'erreur commise par le modèle est connue *a posteriori* par l'algorithme. »





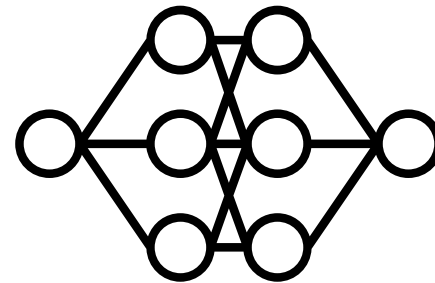
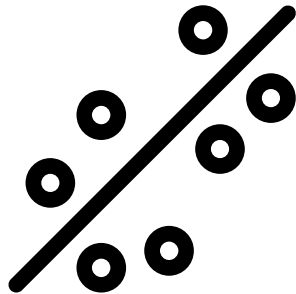
# Des dizaines d'algorithmes ...

- Régression linéaire ou logistique (et si !)
- Arbres de décisions (et si !)
- Forêts aléatoires, méthodes dites « Ensemble »
- Support Vector Machine (SVM)
- Réseau de neurones et son bestiaire (CNN, RNN, LSTM ...)
- K-Means
- Décomposition en composantes principales
- T-distributed stochastic neighbor embedding (t-SNE)
- Q-Learning ...
- ....

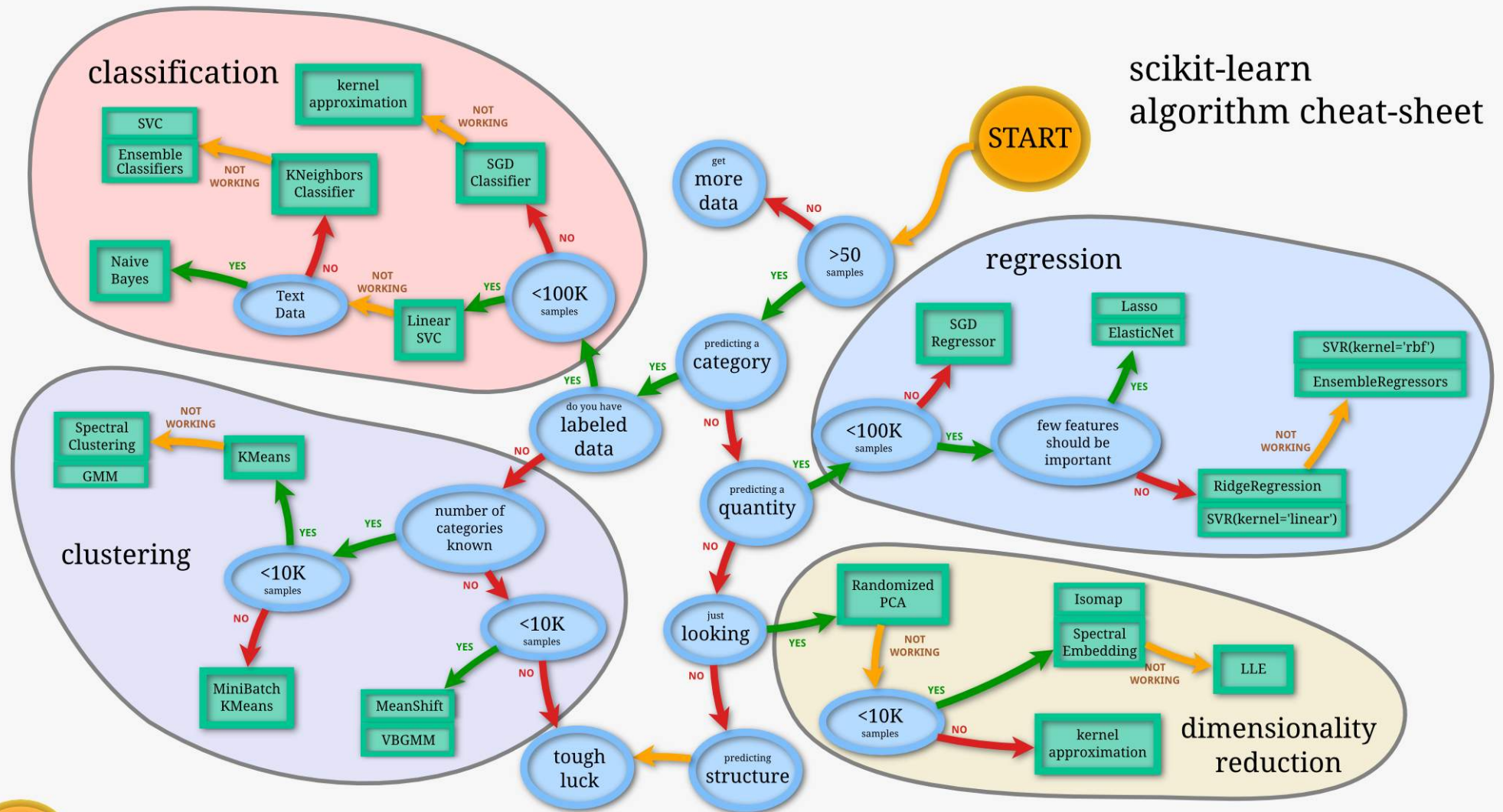


# ... mais le *free lunch theorem* (*Wolpert, 1996*)

- Il n'y a aucun algorithme universellement meilleur que les autres ...
- Dans certains cas, une bonne vieille régression linéaire est plus efficace qu'un gros réseau de neurones ...



# scikit-learn algorithm cheat-sheet



# Machine Learning

Apprentissage supervisé

Apprentissage non supervisé

Apprentissage par  
renforcement

Classif.

Régression

Clustering

Représen-  
tation

Optimisation d'une  
stratégie pour une  
tâche

Apprendre la  
fonction  $f$

$$f : \mathbb{R}^n \rightarrow \{1, 2, \dots, k\}$$

en connaissant  
les paires

$$(X_i, C_i)$$

Apprendre la  
fonction  $f$

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

en connaissant  
les paires

$$(X_i, Y_i)$$

Apprendre la  
fonction  $f$

$$f : \mathbb{R}^n \rightarrow \{1, 2, \dots, k\}$$

$K$  le nombre  
de clusters

en connaissant  
seulement

$$(X_i)$$

Apprendre la  
fonction  $f$

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^k$$

Réduction de  
dimensions  
...

en connaissant  
seulement

$$(X_i)$$

Apprendre la  
fonction  $f$

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^k$$

Trouver l'action optimale pour  
garantir la plus grande  
récompense finale

en connaissant

$$(X_i, a_i, X_{i+1}, r_i)$$

# Machine Learning

Apprentissage supervisé

Apprentissage non supervisé

Apprentissage par  
renforcement

Classif.

Régression

Clustering

Représen-  
tation

Optimisation d'une  
stratégie pour une  
tache

Reconnaissance  
d'image

Classification  
de documents ...

Détection d'  
anomalies

Prédiction

Trading

Segmentation de  
clients

Détection d'  
anomalies

Reconnaissance  
de textes,  
traduction ...

Prédiction  
d'éléments  
suivants dans  
des séquences

Débruitage

Robotique  
Stratégie de trading  
Stratégie de gestion de patrimoine ...  
Jeux ...

# Machine Learning

Apprentissage supervisé

Apprentissage non supervisé

Apprentissage par  
renforcement

Classif.

Régression

Clustering

Représen-  
tation

Optimisation d'une  
stratégie pour une  
tache

Régression  
logistique

SVM

Arbres

Boost

Deep Learning

...

Arbres

Boost

Deep Learning

K-Means

Hiéarch. Clust.

Deep Learning

...

PCA

ICA

t-SNE

Deep Learning

...

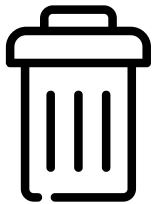
Algo RL

Deep Learning

...



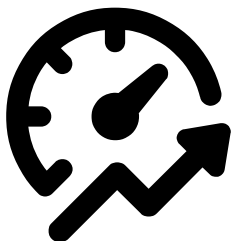
# Quelques principes ...



*Garbage in / Garbage out* : le *machine learning* nourri avec des données de mauvaise qualité donnera un résultat mauvais ...

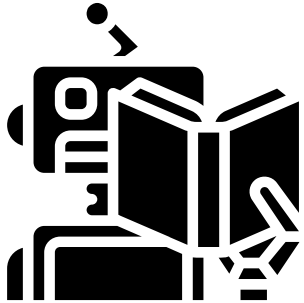


Si on entraîne l'algorithme à reconnaître des chats et des chiens sans lui monter des canards, il ne saura pas reconnaître des canards ...

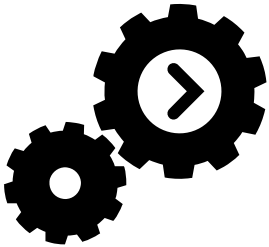


Reconnaître les chats avec 80 % de réussite c'est facile. 95 % c'est extrêmement difficile.

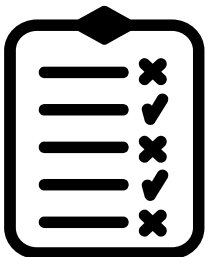
# Revenons sur le cas le simple : Apprentissage supervisé



La machine apprend à partir de données d'apprentissage pour bâtir le modèle

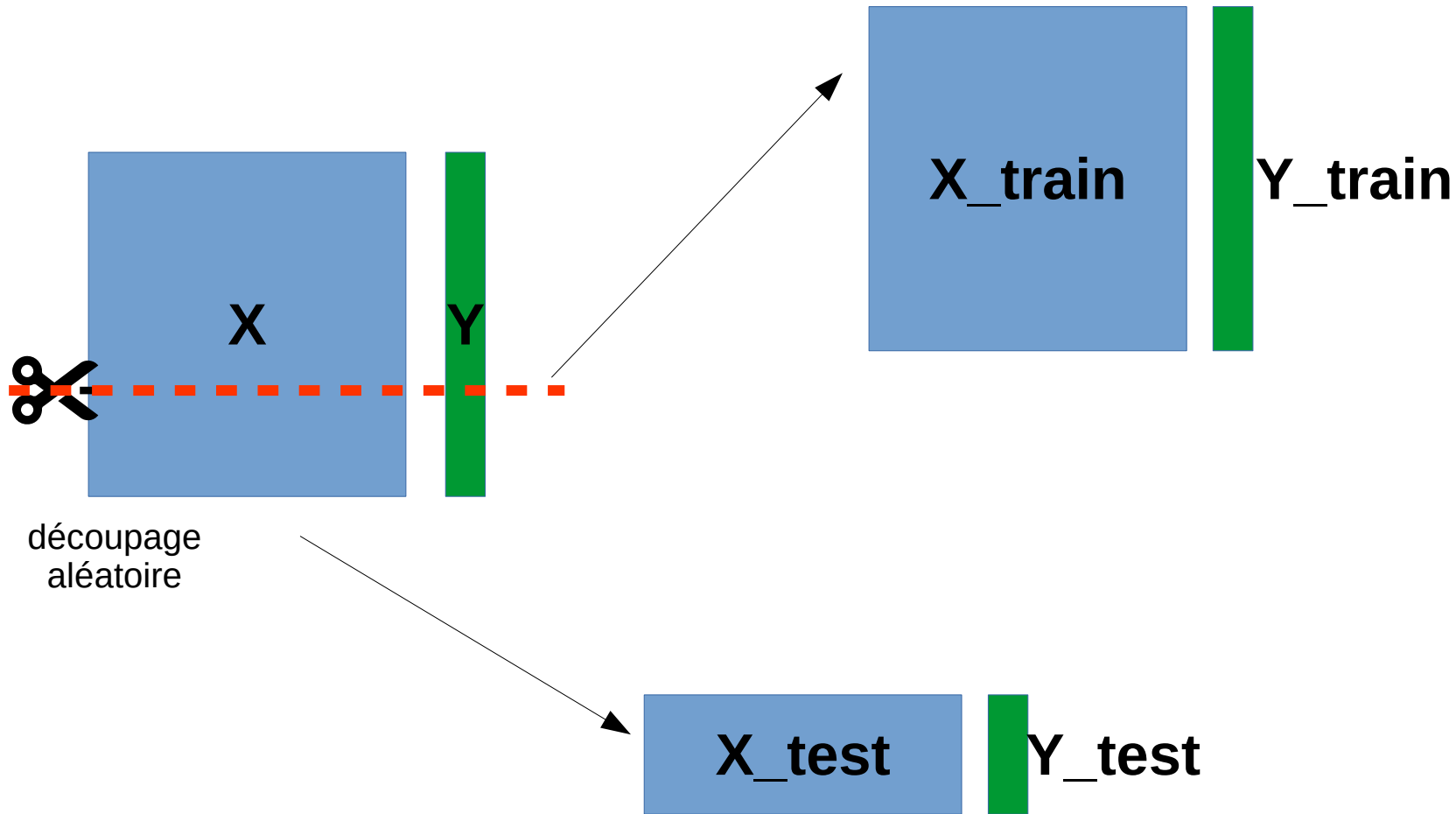


On teste en exécutant le modèle sur des données que la machine n'a « jamais vues »

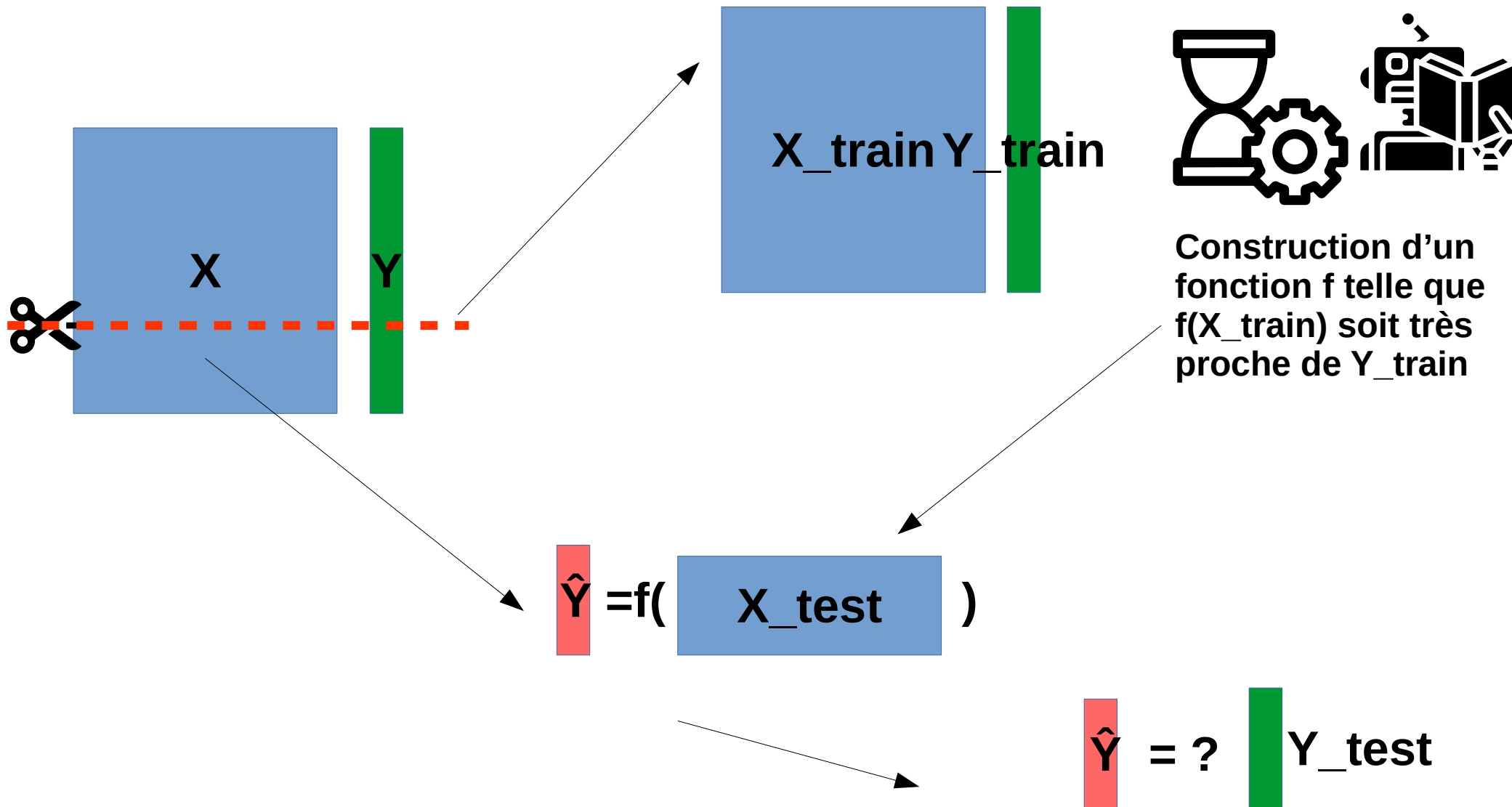


On évalue le modèle en fonction des résultats

Schéma typique (simplifié) :  
il existe d'autres approches plus sophistiquées



# Schéma typique (simplifié) : il existe d'autres approches



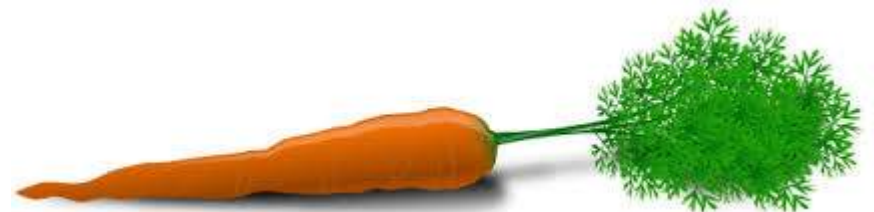
# Machine learning / Comment commencer ?

## Quels outils ?

- Sous python :



- Sous R : caret notamment



# Machine learning / Comment commencer ?

## Quels outils ?

- Commencer par utiliser les NoteBook Jupyter



en utilisant au début les outils gratuits disponibles (MS Azure, Google Colab ...) puis en les installant en local

- Se faire la main avec les exemples de scikit Learn sur les jeux de données classiques (Iris)


<http://scikit-learn.org/stable/tutorial/basic/tutorial.html>





# Machine learning / Comment commencer ?

## Quels outils ?

Se faire la main avec des challenges sur 



### Digit Recognizer

Learn computer vision fundamentals with the famous MNIST data

[Getting Started](#) - Ongoing - 📌 multiclass classification, object identification, image data, tabular data



### Titanic: Machine Learning from Disaster

Start here! Predict survival on the Titanic and get familiar with ML basics

[Getting Started](#) - Ongoing - 📌 tabular data, tutorial, binary classification

# Machine learning / Comment commencer ?

## Quels outils ?

D'excellents MOOC dont ceux d'Andrew NG sur



## Deep Learning Specialization

Deep Learning is transforming multiple industries. This five-course specialization will help you understand Deep Learning fundamentals, apply them, and build a career in AI.



Course 1  
Neural Networks and  
Deep Learning



Course 2  
Improving Deep Neural  
Networks



Course 3  
Structuring Machine  
Learning Projects



Course 4  
Convolutional Neural  
Networks

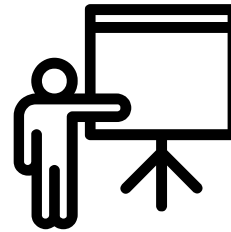


Course 5  
Sequence Models

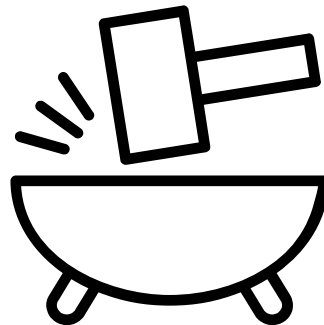
# Encore beaucoup beaucoup de choses à dire



Mais assez de slides



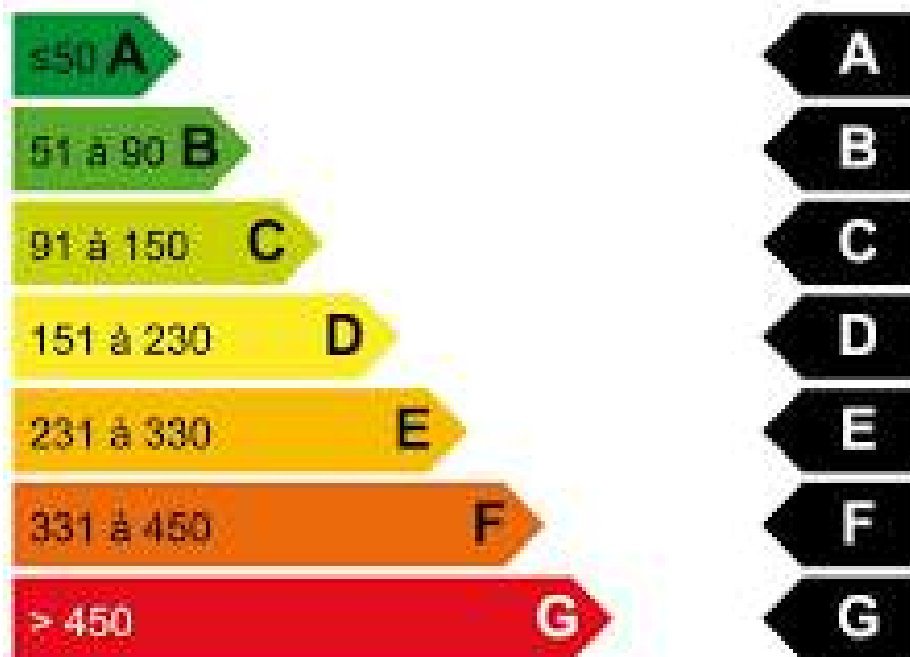
passons à la pratique sur un *notebook*



# 1<sup>er</sup> exemple

## Consommations énergétiques

Logement économe



Logement énergivore

## 2 ième exemple (CIFAR 10)

airplane



automobile



bird



cat



deer



dog



frog



horse



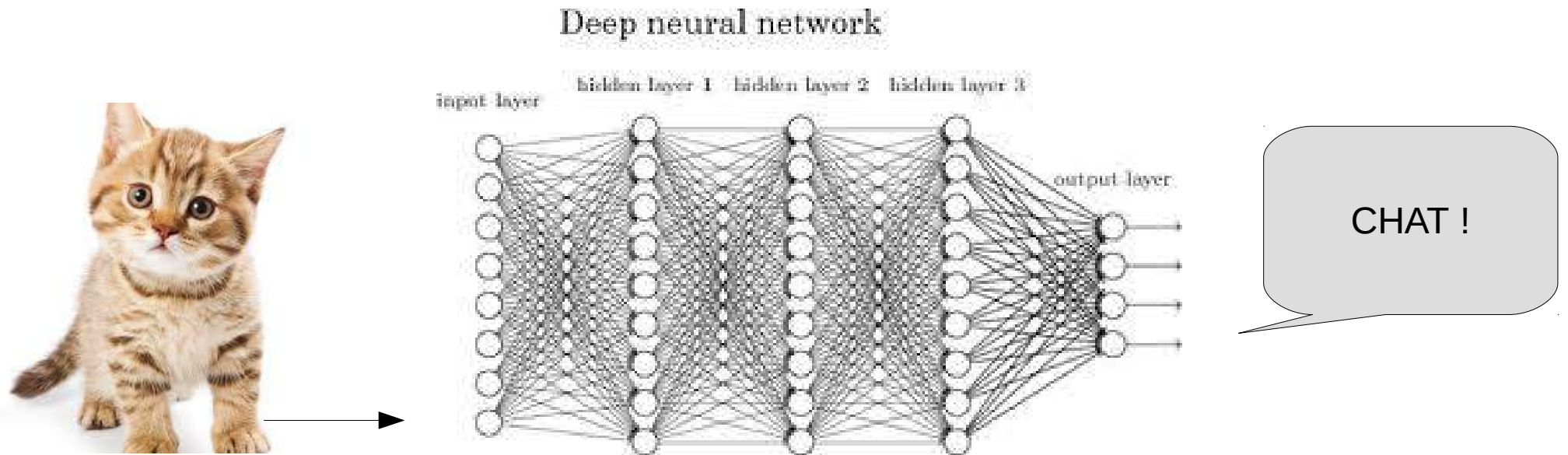
ship



truck



# Et le Deep Learning ?



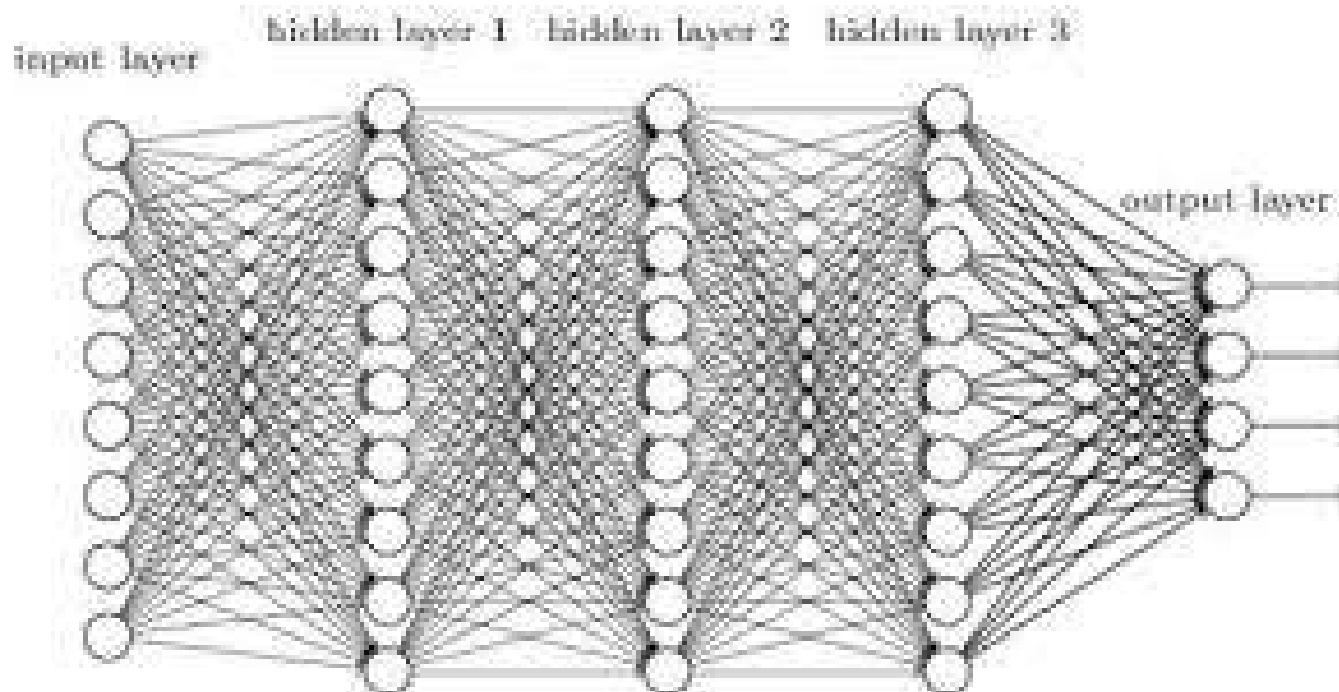
Etiquette : Chat

Réseau de neurone = un graphe de neurones  
Neurone = un traitement simple avec des paramètres.

Ces derniers doivent être ajustés pendant  
l'apprentissage pour que les sorties correspondent  
bien aux étiquettes

# Micro focus sur le deep learning

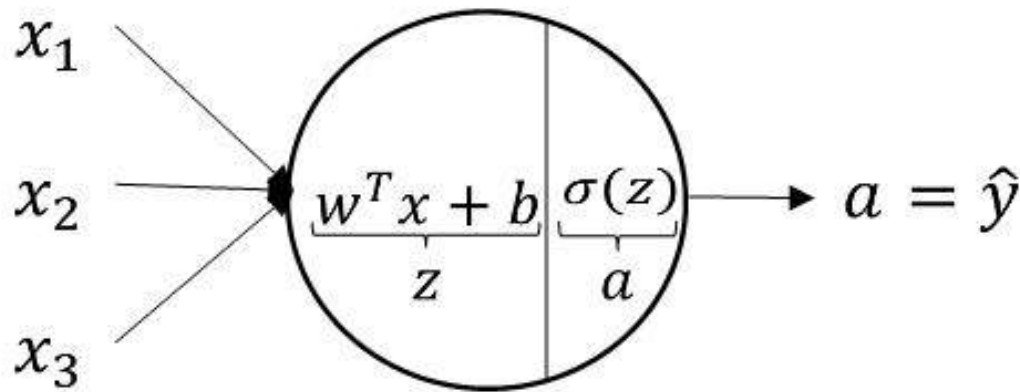
## Deep neural network



Le réseau de neurones constitue un graphe de calcul à opérer sur la donnée en entrée (ex : une image). Il en sort un résultat (ex : la classification de l'image permettant d'affirmer que c'est un chat). Chaque "neurone" constitue un calcul intermédiaire du traitement global. Chaque neurones de la couche  $n$  prend en entrée les sorties des neurones de la couche  $n-1$ .



# Et les neurones, ils font quoi ?



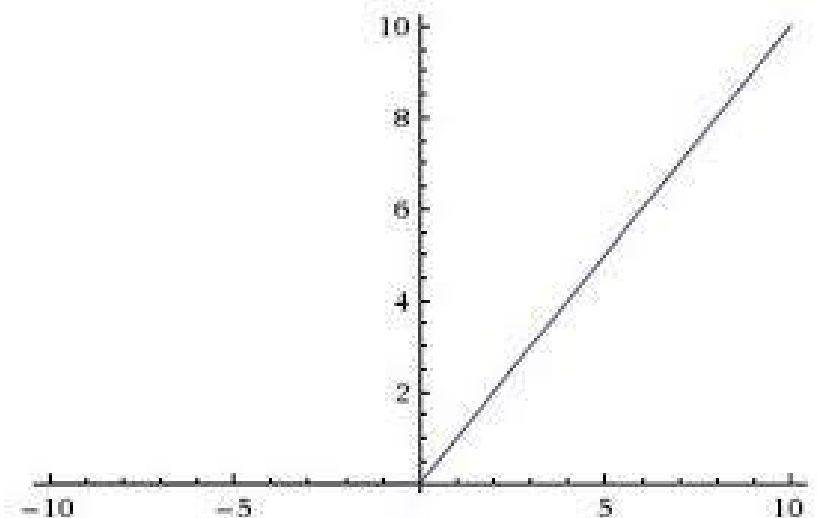
Source : deeplearning.ai

$$z = w^T x + b$$

$$a = \sigma(z)$$

**Apprentissage :**

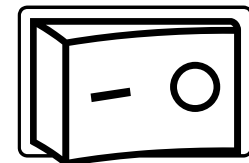
Trouver les meilleurs  
**w** et **b**



Fonction ReLu :  $f(x) = \max(0, x)$

# Fonction d'activation

- Les neurones calculent une sorte de somme pondérée à partir de paramètres du neurone (paramètres = les poids  $w_i$  + un biais  $b$ ). La sortie passe ensuite par une fonction non linéaire (pour casser la linéarité du modèle ... sinon on aurait l'équivalent d'un réseau à une couche ; car la juxtaposition de couches linéaires donne un résultat linéaire).
- C'est la fonction d'activation qui joue aussi le rôle de fonction de seuil : en-dessous d'un certain seuil, le neurone est inactif. Au-dessus, il est activé. Il existe plusieurs fonctions d'activation : ***sigmoid*** (utile pour les régressions logistiques), ***tanh***, mais la plus connue et la plus utilisée (car elle se calcule très vite) est la fonction ***ReLu*** (*rectified linear unit*).



# Apprentissage

**L'apprentissage consiste à définir les meilleurs poids possibles pour chaque neurone.** Plus concrètement, il s'agit de minimiser une fonction de coût, sorte de distance entre les valeurs calculées par le modèle et les vraies valeurs (celles des étiquettes qui accompagnent chaque donnée). La fonction de coût la plus utilisée est la fonction de l'entropie croisée (*cross entropy*).

$$L(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N H(p_n, q_n) = -\frac{1}{N} \sum_{n=1}^N \left[ y_n \log \hat{y}_n + (1 - y_n) \log(1 - \hat{y}_n) \right]$$

Source: Wikipedia,

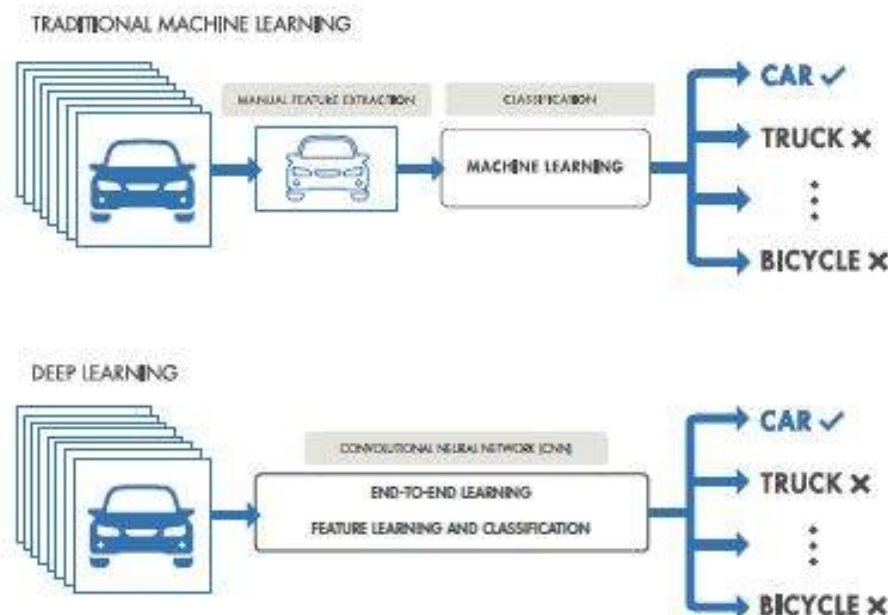
Les  $y$  sont les valeurs étiquettes ; les  $y$  *chapeau* sont les valeurs calculées par le réseau.

La minimisation de la fonction de coûts utilise la technique classique dite de la **descente de gradient** (et la mise à jour des poids par backpropagation).

# What is the Difference Between Deep Learning and Machine Learning?

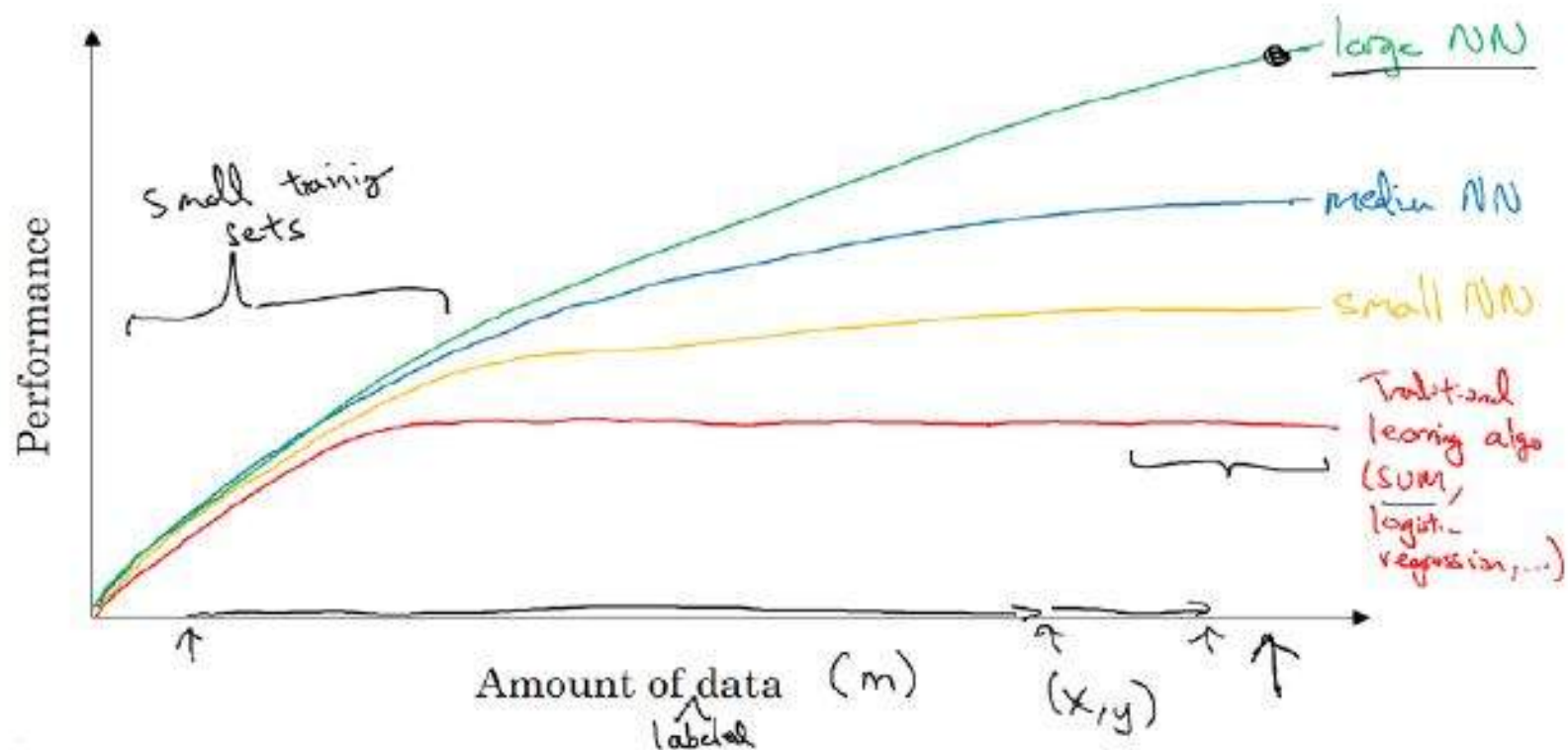
Deep learning is a subtype of machine learning. With machine learning, you manually extract the relevant features of an image. With deep learning, you feed the raw images directly into a deep neural network that learns the features automatically.

Deep learning often requires hundreds of thousands or millions of images for the best results. It's also computationally intensive and requires a high-performance GPU.



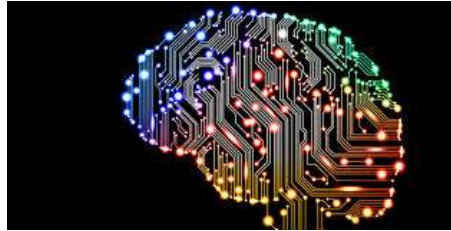
| Machine Learning   | Deep Learning                                   |
|--|---|
| + Good results with small data sets                                      | – Requires very large data sets                 |
| + Quick to train a model   | – Computationally intensive                     |
| – Need to try different features and classifiers to achieve best results | + Learns features and classifiers automatically |
| – Accuracy plateaus  | + Accuracy is unlimited                         |

## Scale drives deep learning progress



Source DeepLearning.ai

# Ce que nous allons faire !



ça, c'est un chat !!!

ça, c'est un chien !!!

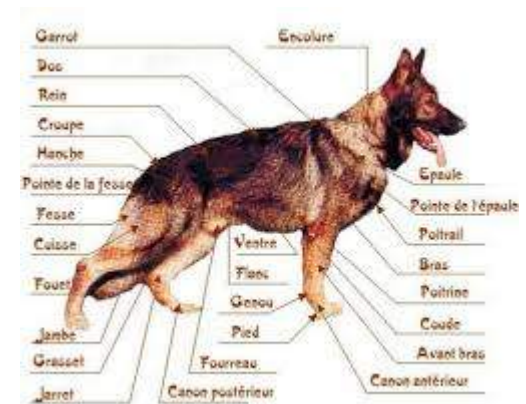
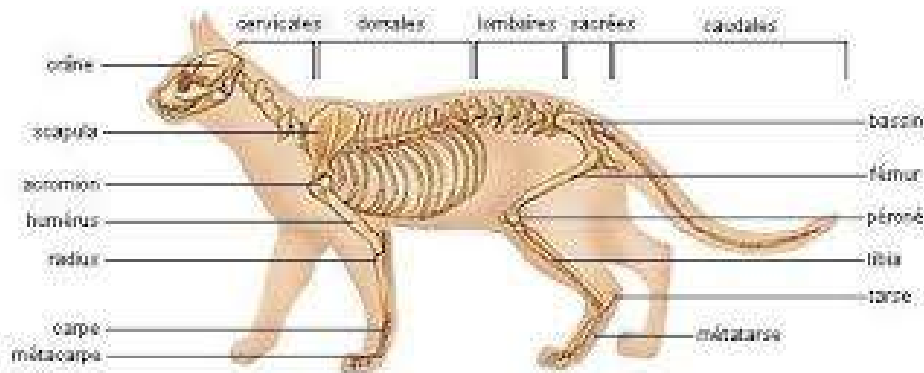
ça, c'est une voiture !!!

ça, c'est un chat !!!  
Euh, non, une voiture !!



# C'est pas gagné ...

- Comment expliquer à l'ordinateur ce qu'est un chat, un chien, une voiture, un camion etc... ?
- Un enfant de 4 ans sait reconnaître un chat ou un chien mais comment le formaliser ?
- Avec 100 000 lignes de code, peut-être, mais on a 30 minutes !





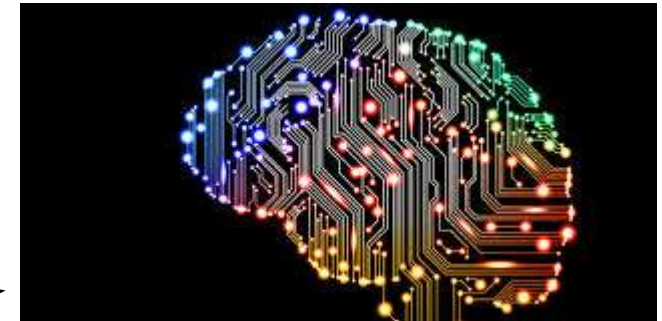
# Trichons un peu ...

- On va présenter 5000 images de chats à l'ordinateur en lui disant « ça c'est un chat », 5000 images de chiens en lui disant « ça, c'est un chien », 5000 images de voitures, 5000 de camions ...
- Et puis maintenant, toi l'ordinateur, apprends tout seul à faire la différence entre un chat et un chien ... zut, c'est ça aussi l'intelligence artificielle, débrouille-toi !!

# Apprentissage avec $10 \times 5000$ images !!

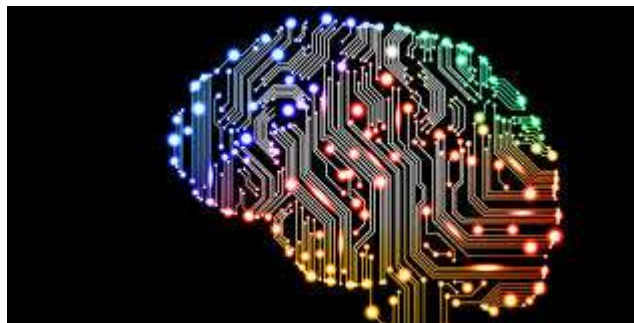
50 000 images  
à apprendre !!!

C'est long !



Verdict : Le modèle est-il capable  
de classer des images qu'il n'a  
jamais vues ?

10\*1000 images de test !



Zut, moi  
j'ai bachoté !!

# Bon maintenant à vous de jouer !

- Ce n'est pas réservé à des hyper-spécialistes !
- Les applications sont nombreuses !
- Imaginez le potentiel avec un vrai projet, une vraie équipe, avec des vraies compétences !
- Et en mixant les compétences métier (les vôtres) avec des compétences de ML !! *L'expertise augmentée ?*

# NoteBook Jupyter

- Les notebook :

<https://github.com/rst-analytics/atelier-ml/>

- Version ancienne de l'atelier à cloner :

<https://notebooks.azure.com/jetzgetzlos-POC/libraries/jouons-avec-le-ML/html/atelier-ML.ipynb>

- Version ancienne plus complète pour aller plus

**loin:** <https://notebooks.azure.com/jetzgetzlos-POC/libraries/POC-cat-classif/html/POC-cat.ipynb>