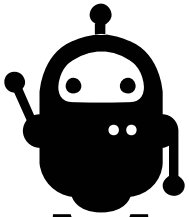
Version : 1.40

AI Machine Learning & deep learning



T8 : « Programmer et déployer votre IA »

Jean-Luc PAROUTY – CNRS/SIMaP

7 juillet 2020

cnrs SIMaP JDEV

Many thanks to :

Comité d'organisation
Comité de programme

Formation Introduction au Deep Learning

Soraya ARIAS – INRIA
Eric MALDONADO – INRAE
Jean-Luc PAROUTY – SIMaP

www.economie.gouv.fr

Lancement du partenariat mondial pour l'intelligence artificielle

www.economie.gouv.fr
16/06/2020

Intelligence artificielle : administrations, proposez vos projets d'expérimentation

www.etalab.gouv.fr
02/04/2019

www.defense.gouv.fr

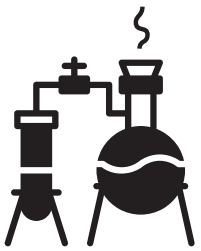
Vocabulaire de l'intelligence artificielle (liste de termes, expressions et définitions adoptés)

JORF n°0285 du 9 décembre 2018, texte n° 58

Ok, C’est à la mode… Mais de quoi parle t-on ?

[Méthode scientifique]

1st paradigm



Experimental science

2nd paradigm

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle$$

$$\nabla \times H = J + \frac{\partial D}{\partial t}$$

$$F = G \cdot \frac{m_1 \cdot m_2}{r^2}$$

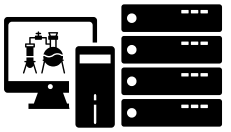
Theoretical science

3rd paradigm

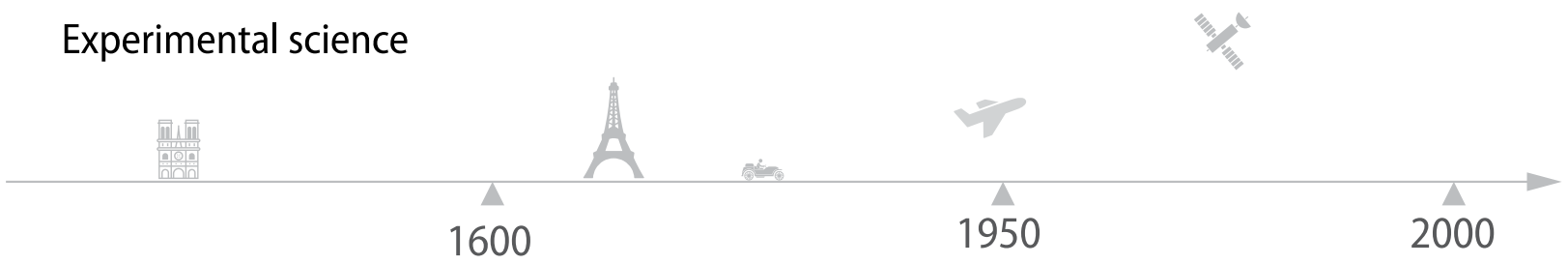
$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle$$

$$\nabla \times H = J + \frac{\partial D}{\partial t}$$

$$F = G \cdot \frac{m_1 \cdot m_2}{r^2}$$

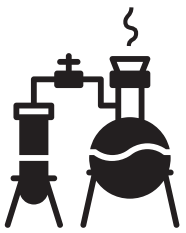


Computational science



[Méthode scientifique]

1st paradigm



Experimental science

2nd paradigm

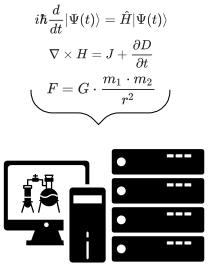
$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle$$

$$\nabla \times H = J + \frac{\partial D}{\partial t}$$

$$F = G \cdot \frac{m_1 \cdot m_2}{r^2}$$

Theoretical science

3rd paradigm



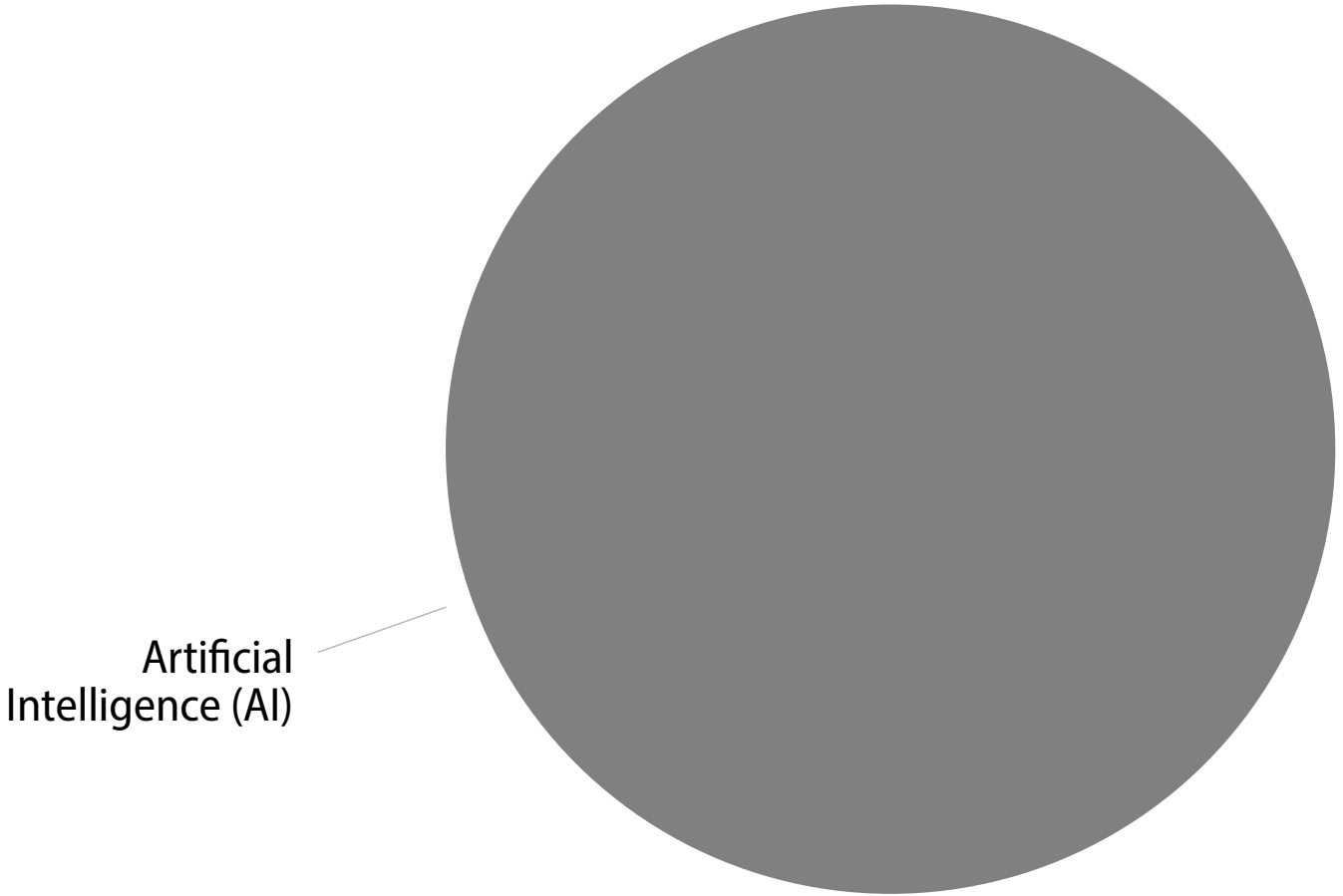
Computational science

4th paradigm[1]

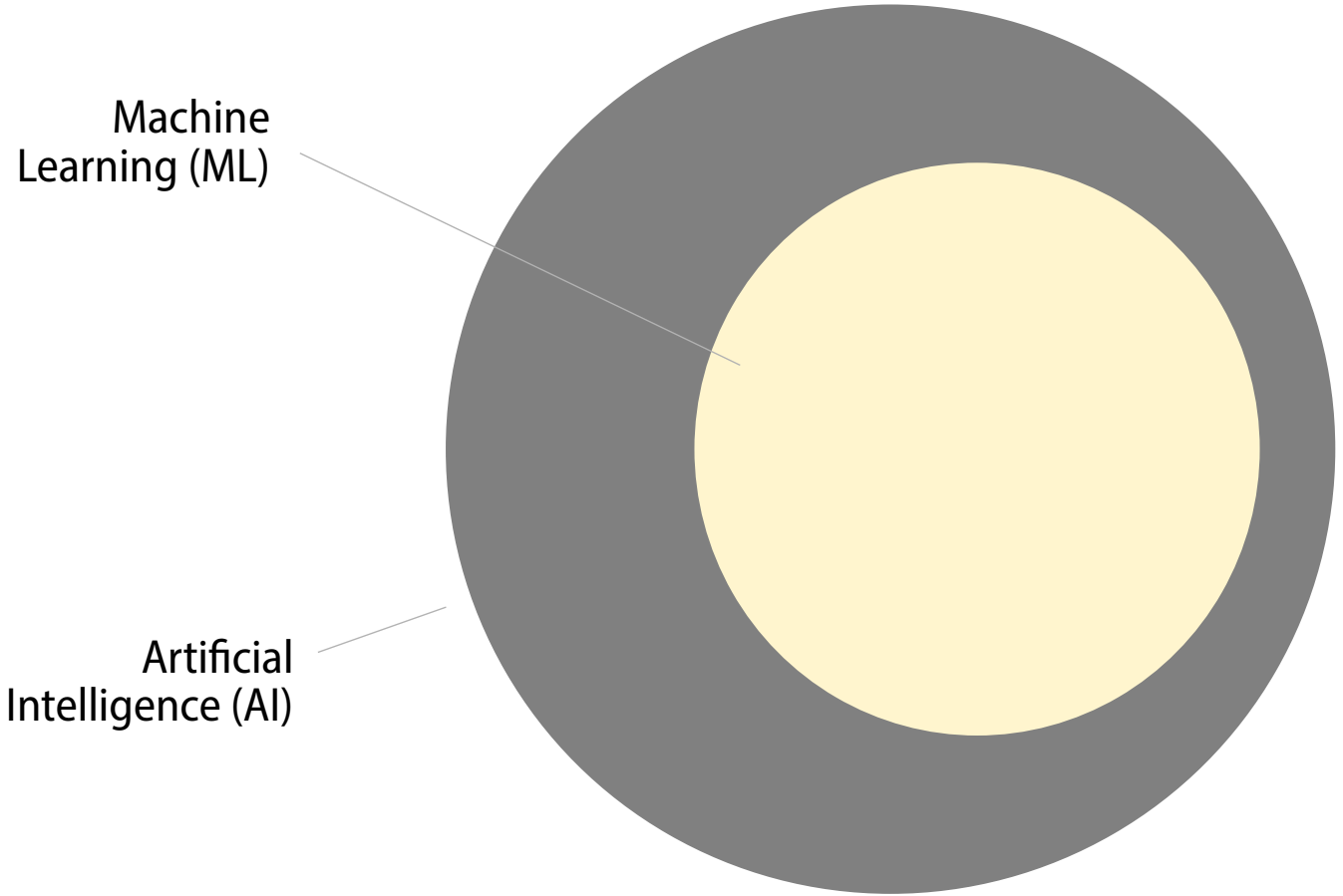
Data-driven science

[1] Jim Gray, 2007 [GRAY]

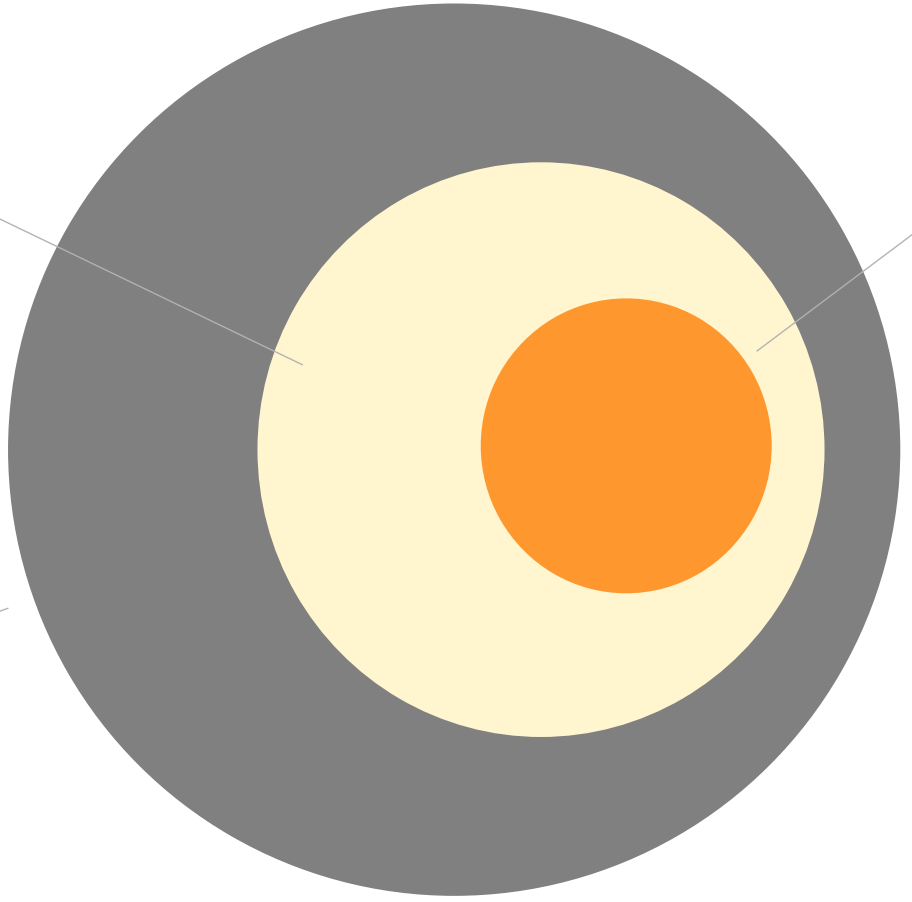
[*-learning]



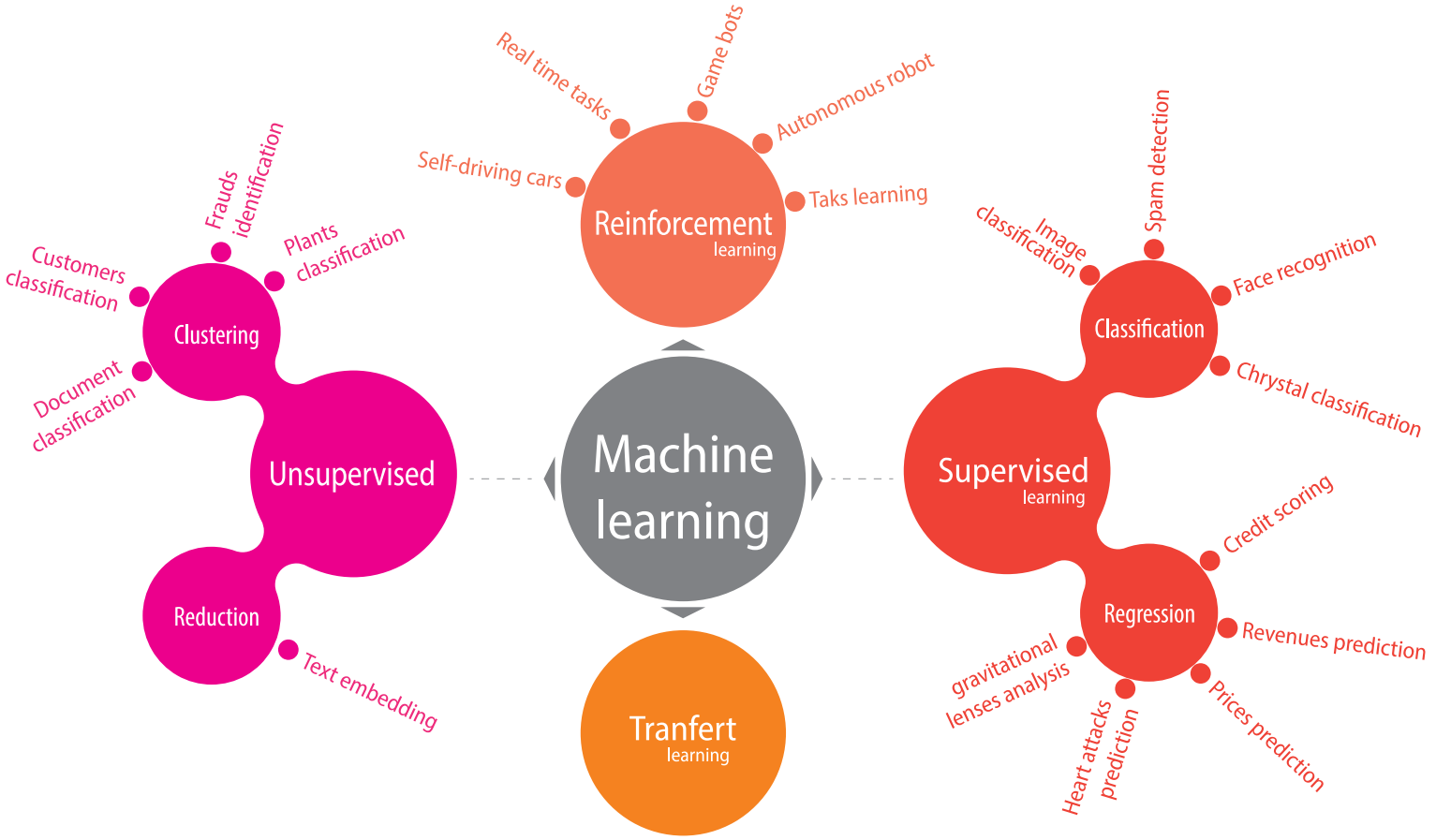
[*-learning]



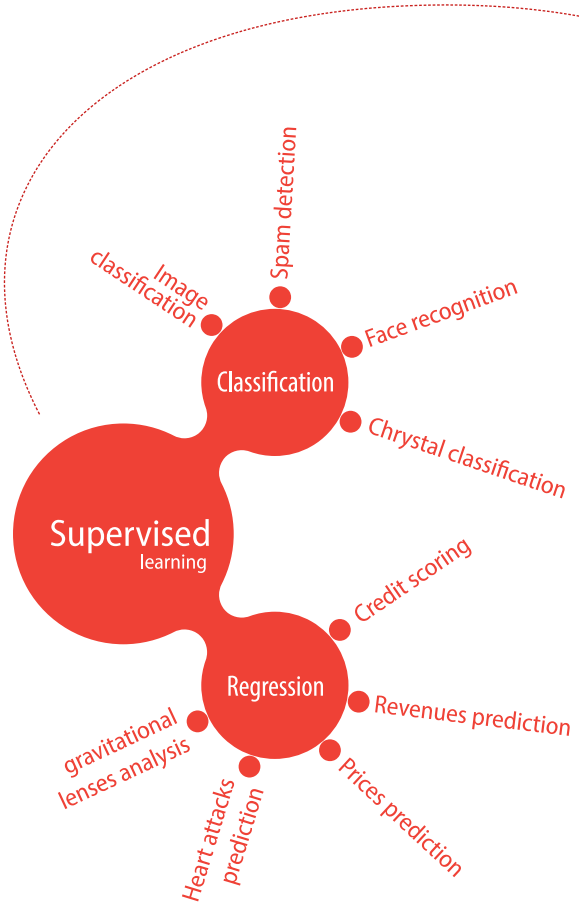
[*-learning]


Machine
Learning (ML)
Artificial
Intelligence (AI)
Deep
Learning (DL)
Based on artificial neural networks

[*-learning]

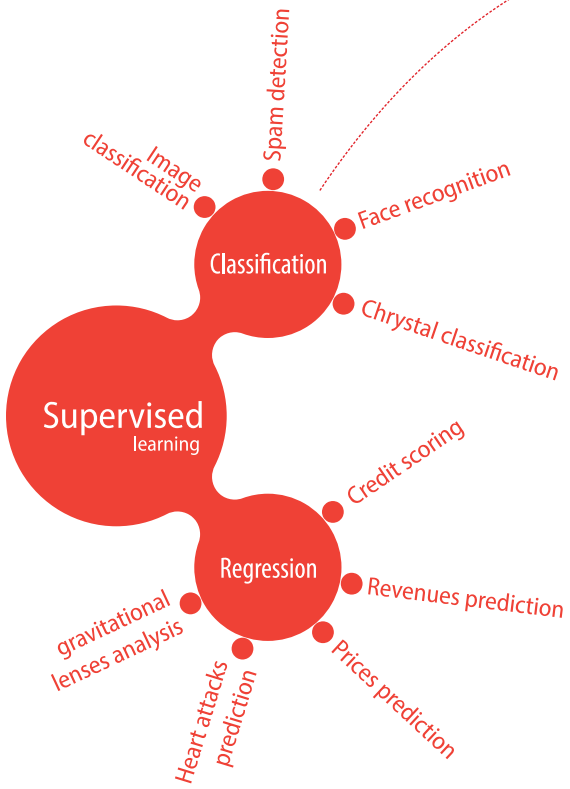

Machine learning
Unsupervised
Clustering
Frauds identification
Plants classification
Customers classification
Document classification
Reduction
Text embedding
Reinforcement learning
Real time tasks
Game bots
Autonomous robot
Self-driving cars
Taks learning
Tranfert learning
Supervised learning
Classification
Spam detection
Image classification
Face recognition
Chrystal classification
Regression
Credit scoring
Revenues prediction
Prices prediction
Heart attacks prediction
gravitational lenses analysis

Supervised learning

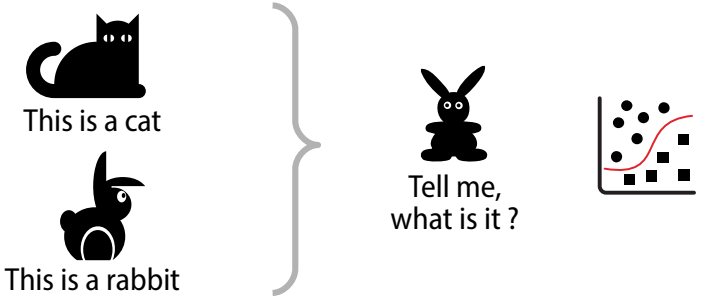


Learning from examples

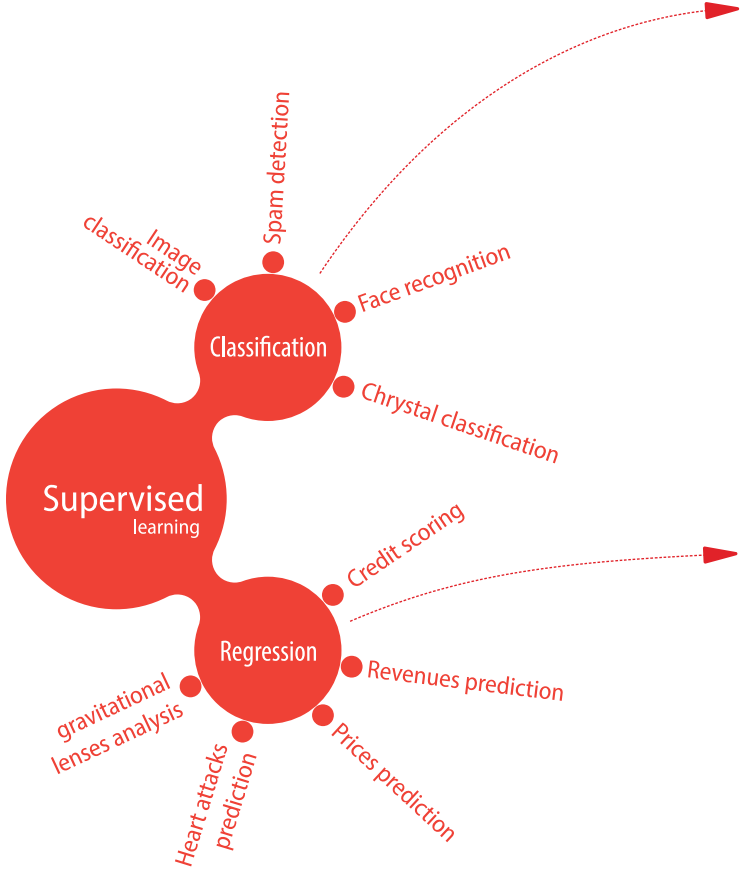
Supervised learning



Classification :
Predict qualitative informations

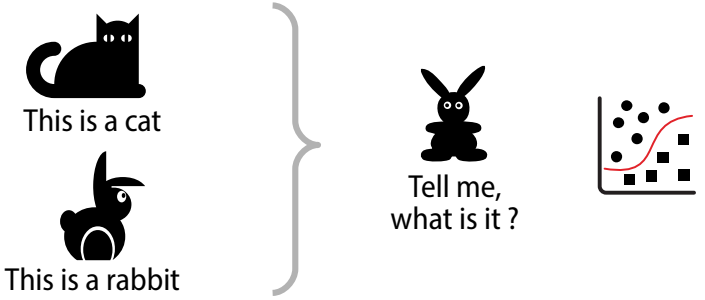


Supervised learning



Classification :

Predict qualitative informations

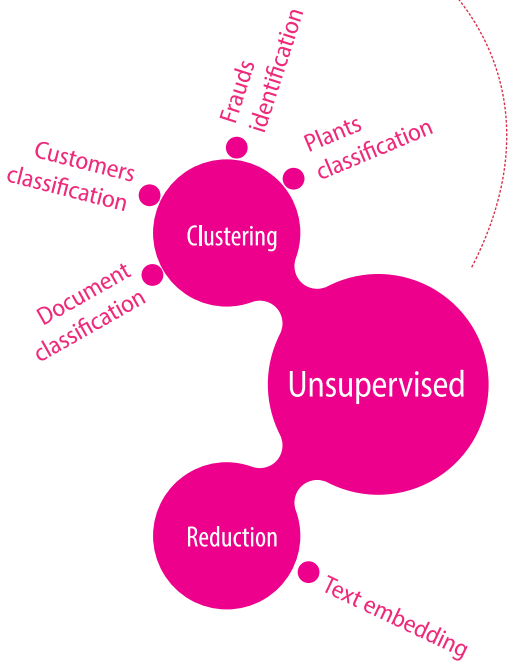


Régression :

Predict quantitative informations

Unsupervised learning

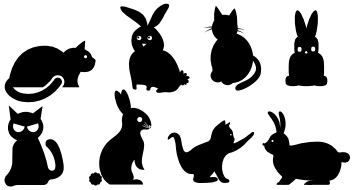
Learning from data alone

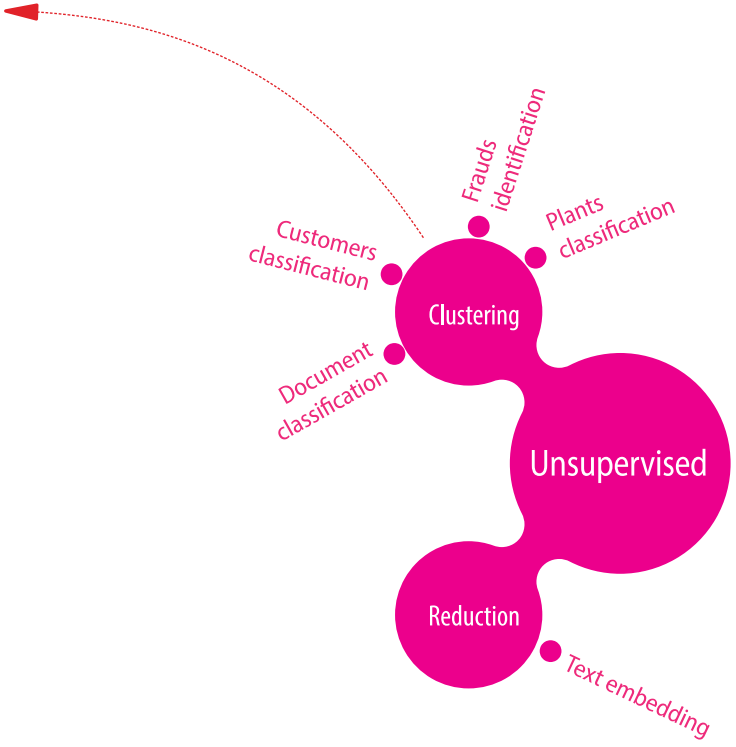


Unsupervised learning

Clustering :
Finding Common Relationships

What is the relationship between these data ?

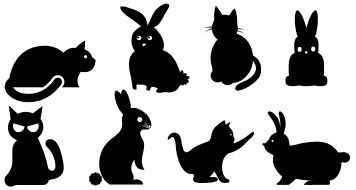




Unsupervised learning

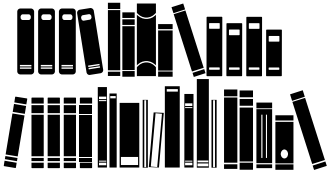
Clustering :
Finding Common Relationships

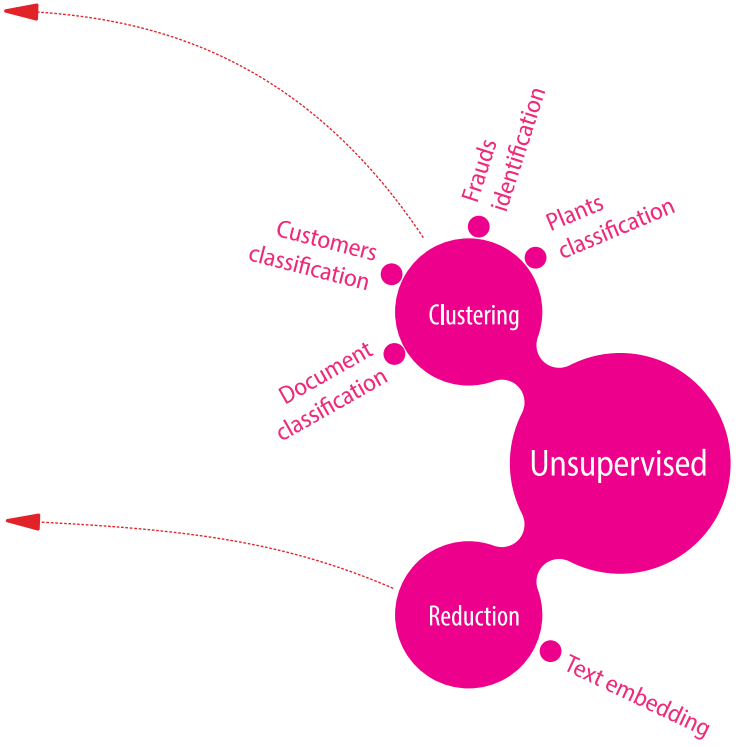
What is the relationship between these data ?



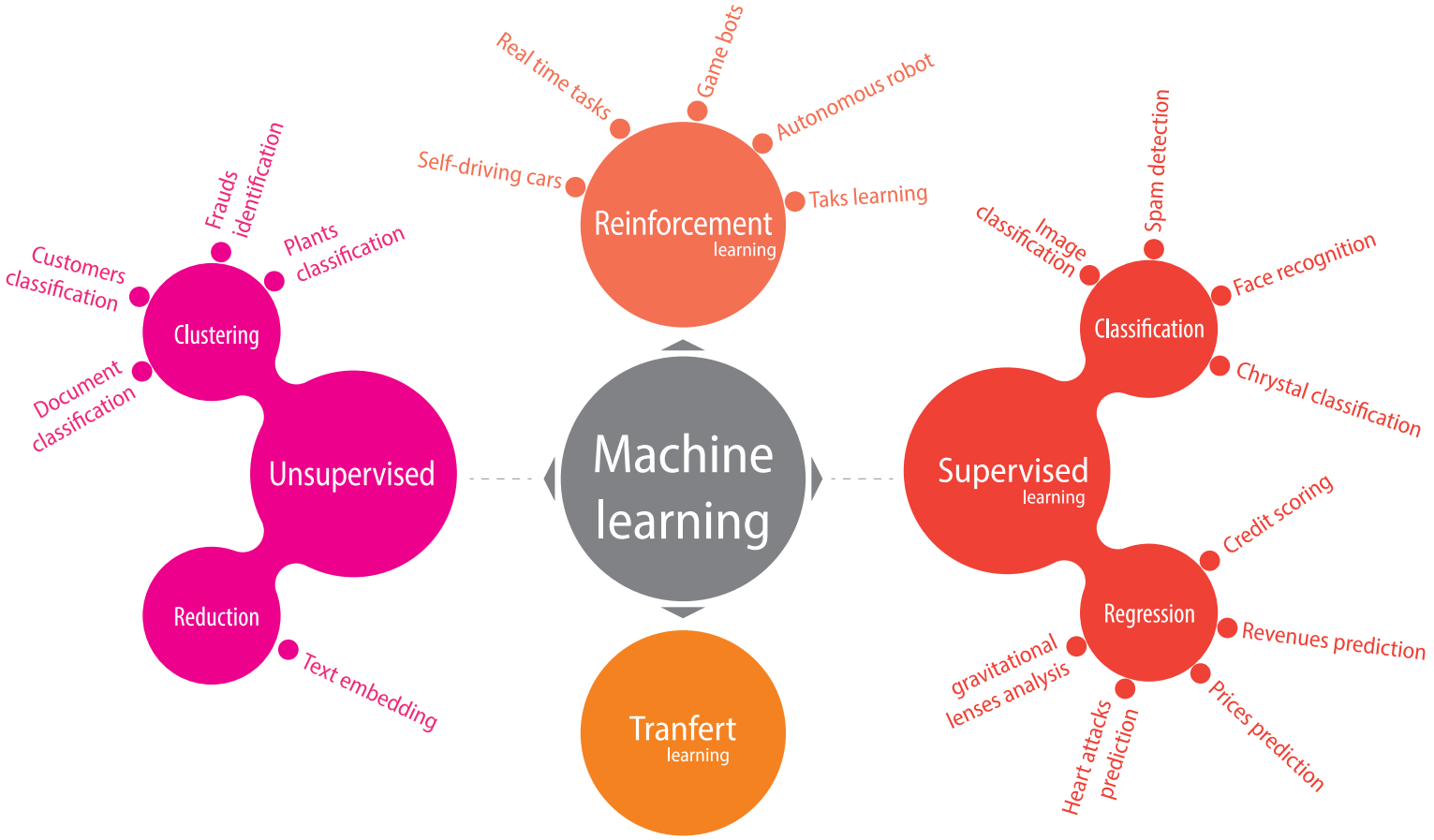
Reduction :
Reduce the number of dimensions

Simplify while keeping meaning

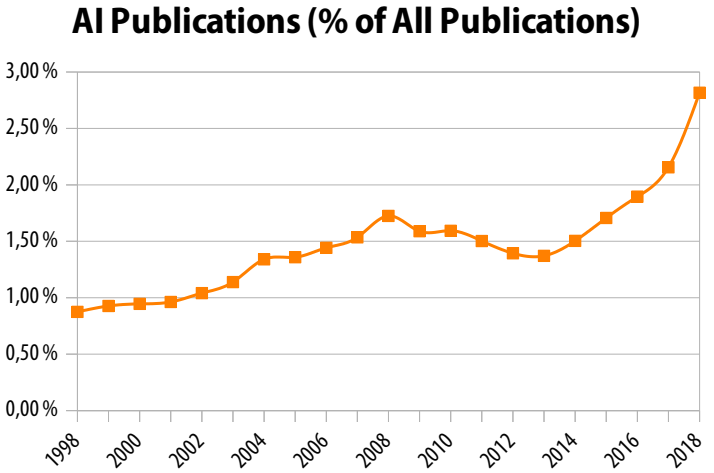


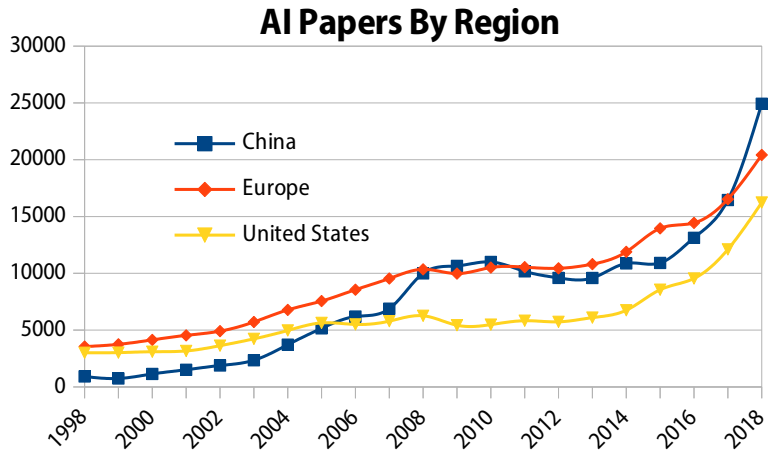


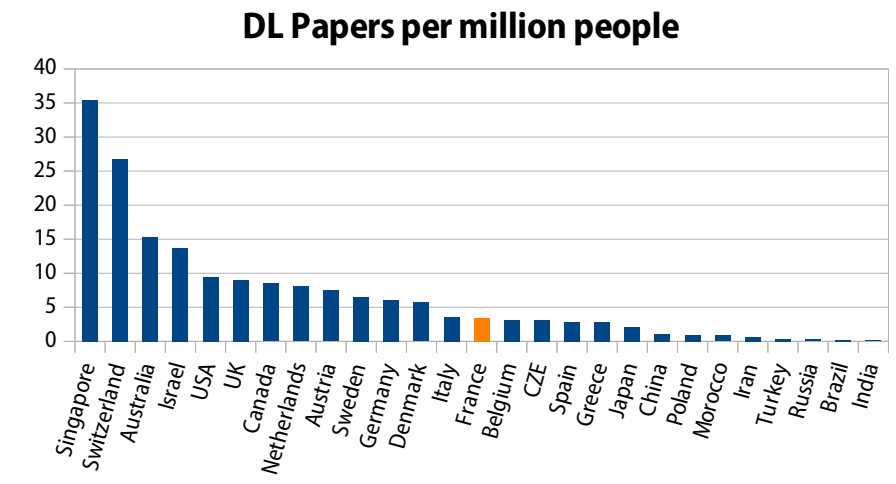
[*-learning]

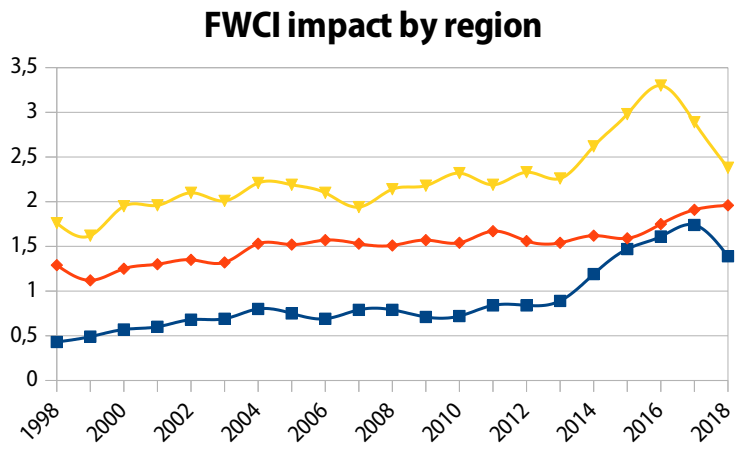

Machine learning
Unsupervised
Clustering
Frauds identification
Plants classification
Customers classification
Document classification
Reduction
Text embedding
Reinforcement learning
Real time tasks
Game bots
Autonomous robot
Self-driving cars
Taks learning
Tranfert learning
Supervised learning
Classification
Spam detection
Image classification
Face recognition
Chrystal classification
Regression
Credit scoring
Revenues prediction
Prices prediction
Heart attacks prediction
gravitational lenses analysis

A strong competition





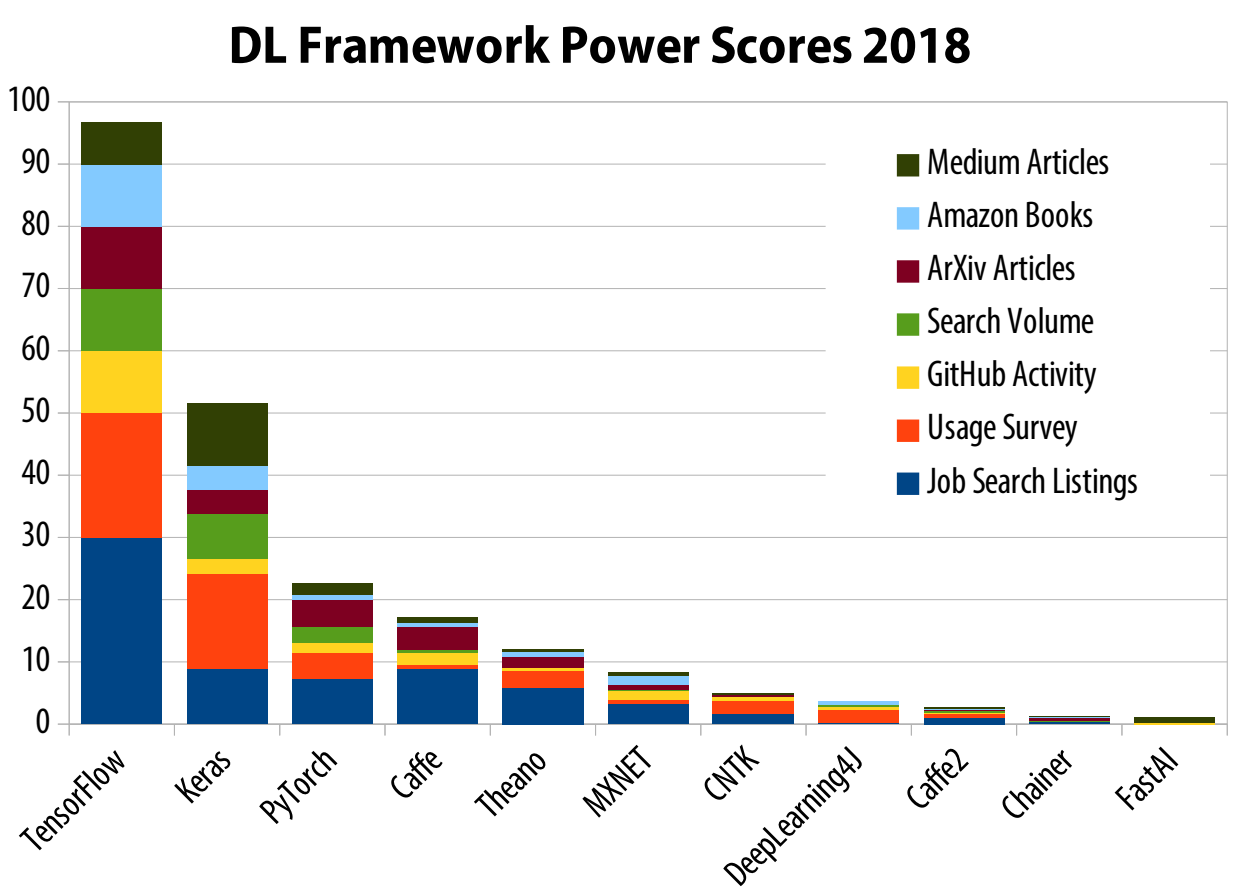




Source : AI Index [AIDEX]

FWCI : Field-Weighted Citation Impact

A Python centered world



Source : Jeff Hale, « Deep Learning Framework Power Scores 2018 » [DLPW]

TensorFlow

Most used DL framework
Supported by Google
Low level API – an hard way
Apache licence

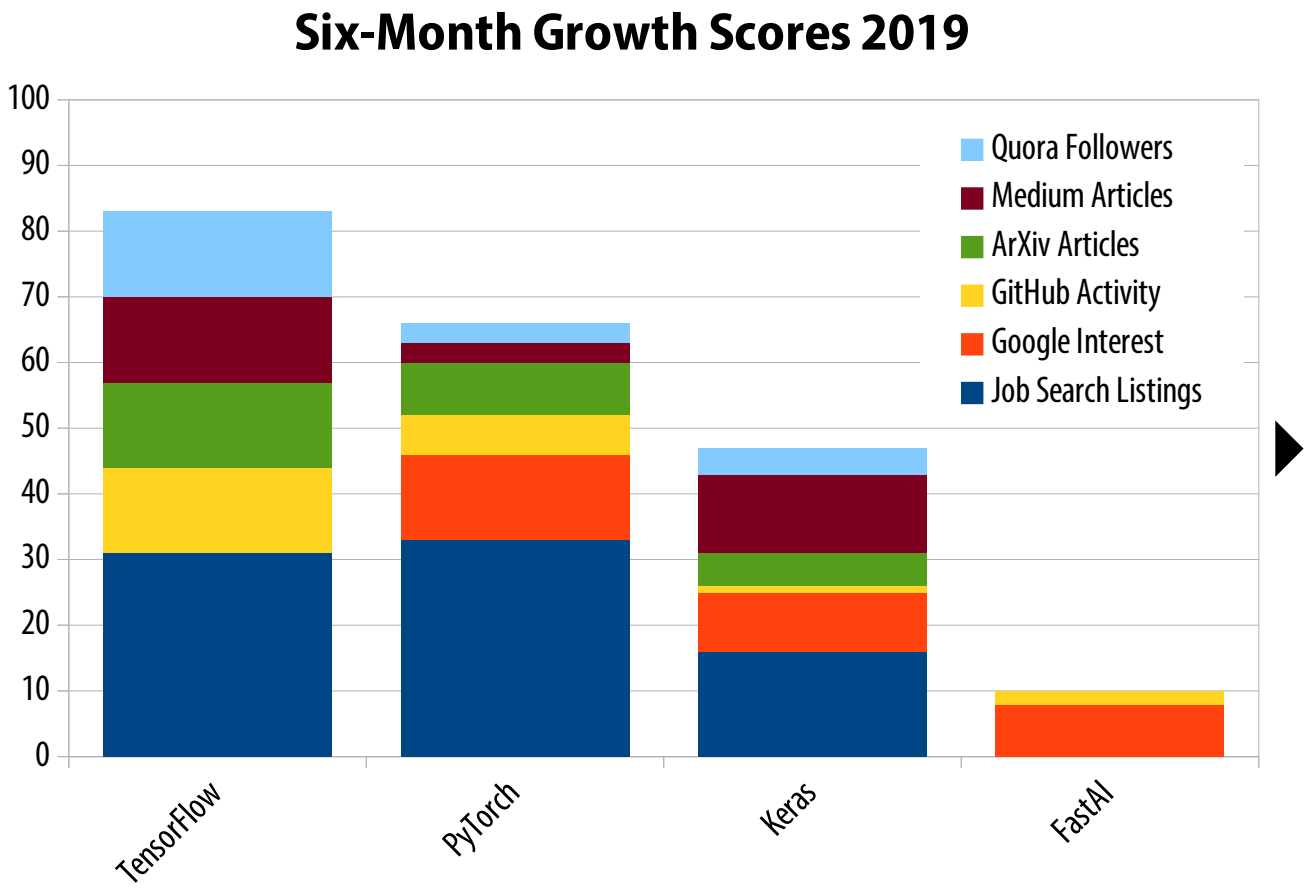
Keras

By François Cholet (Google)
Hight level API
Part on TensorFlow since 2017
MIT licence

PYTORCH

From Torch library
Supported by Facebook
BSD licence

A Python centered world



TensorFlow

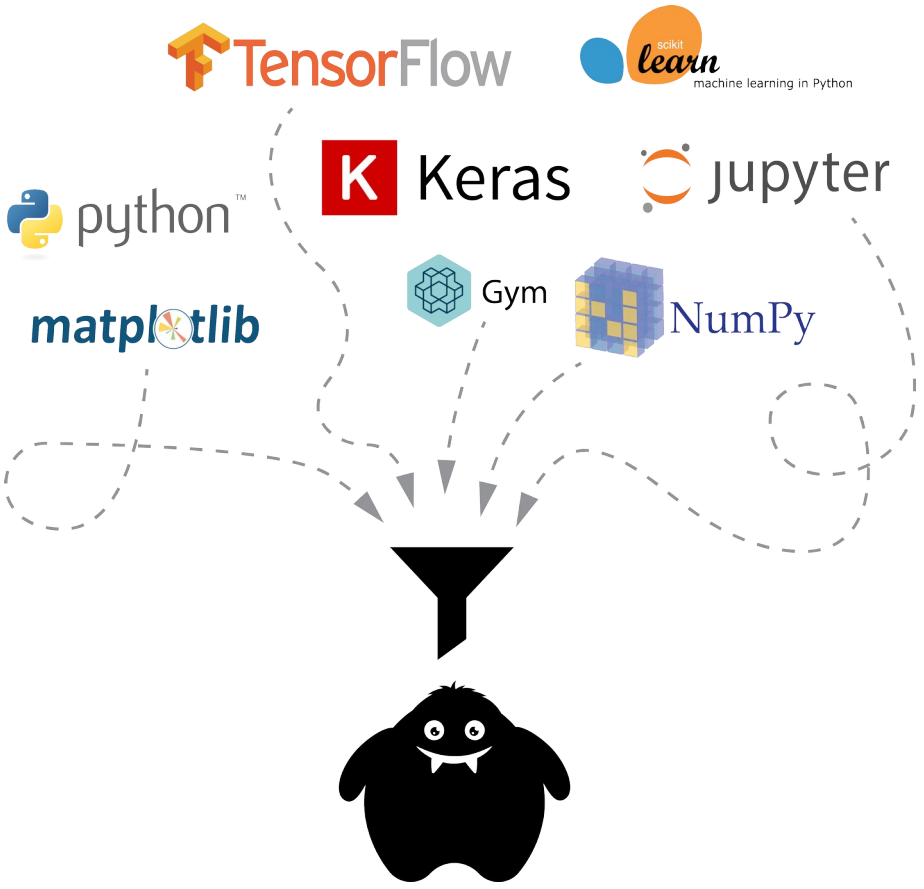
PYTORCH

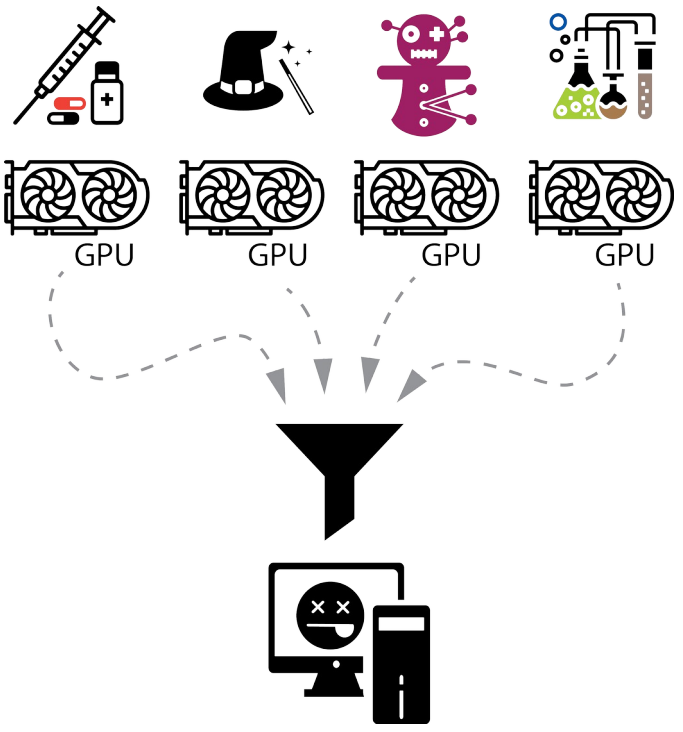
K Keras

fast.ai

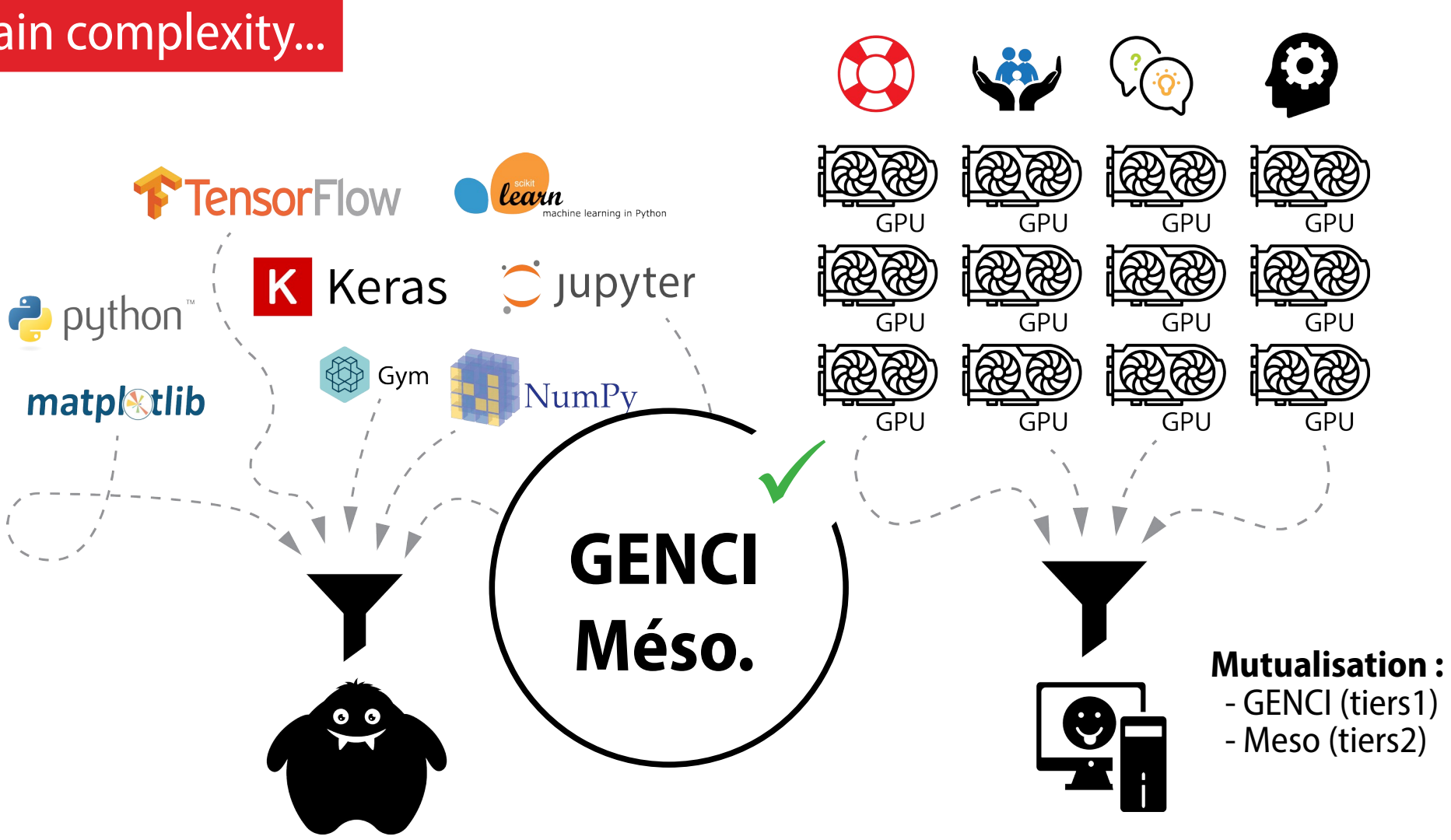
Source : Jeff Hale, « Deep Learning Framework Growth Scores 2019 » [DLPW]

A certain complexity...





A certain complexity...



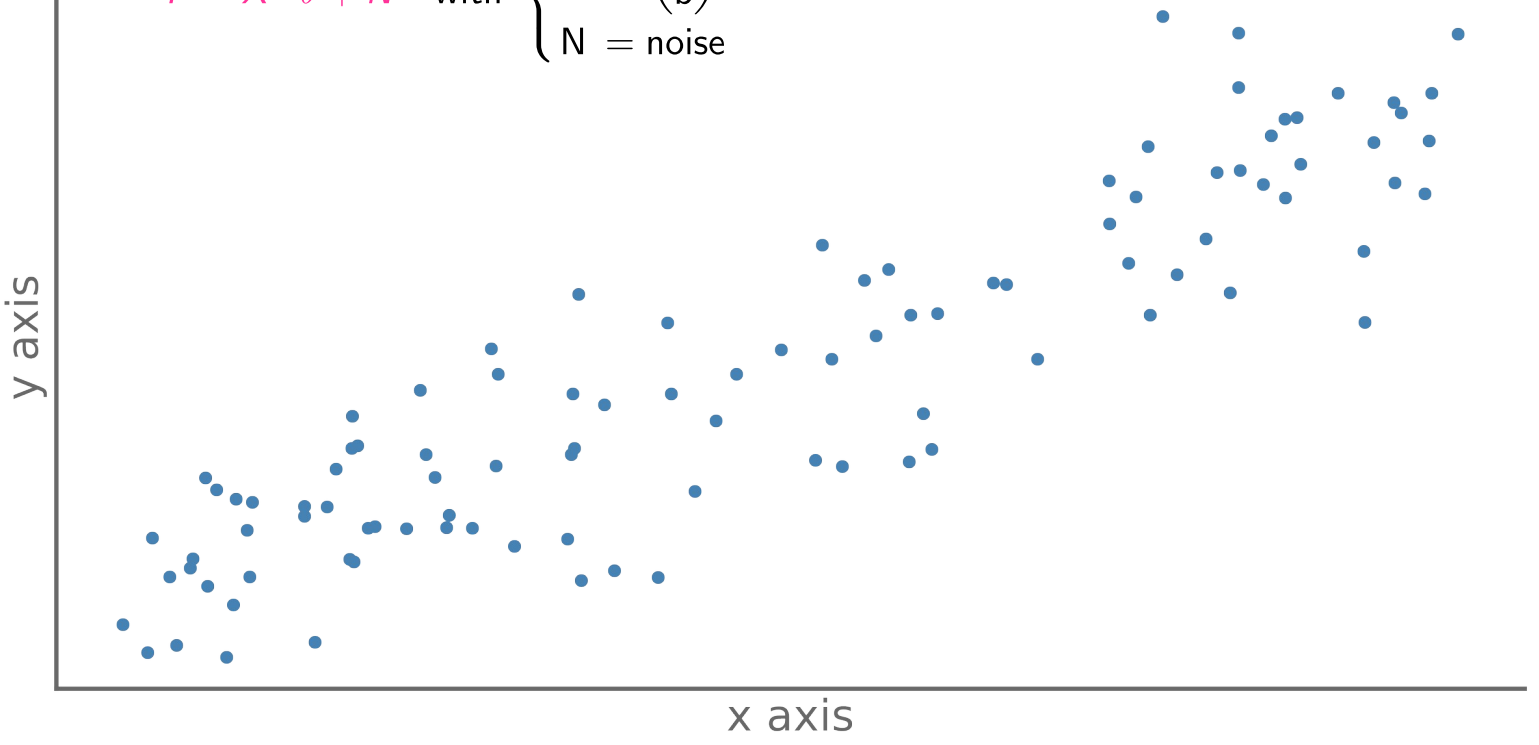
Fine, but Deep Learning What's that?

...artificial neurons ?
Ok, but what are they?

From the linear regression to the first neuron

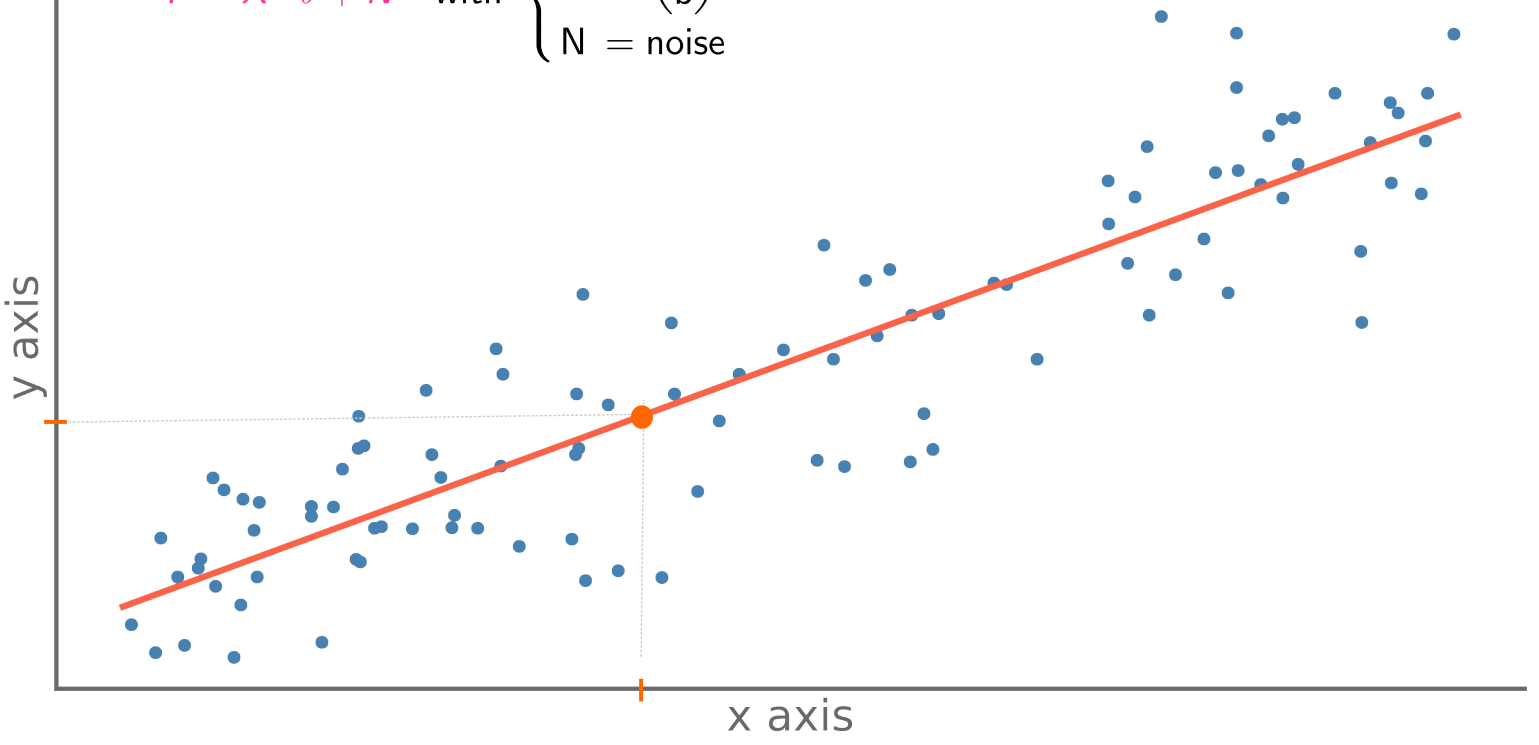
Linear regression

$$Y = X \cdot \theta + N \quad \text{with} \begin{cases} \theta = \begin{pmatrix} a \\ b \end{pmatrix} \text{ as } y = a \cdot x + b \\ N = \text{noise} \end{cases}$$



Linear regression

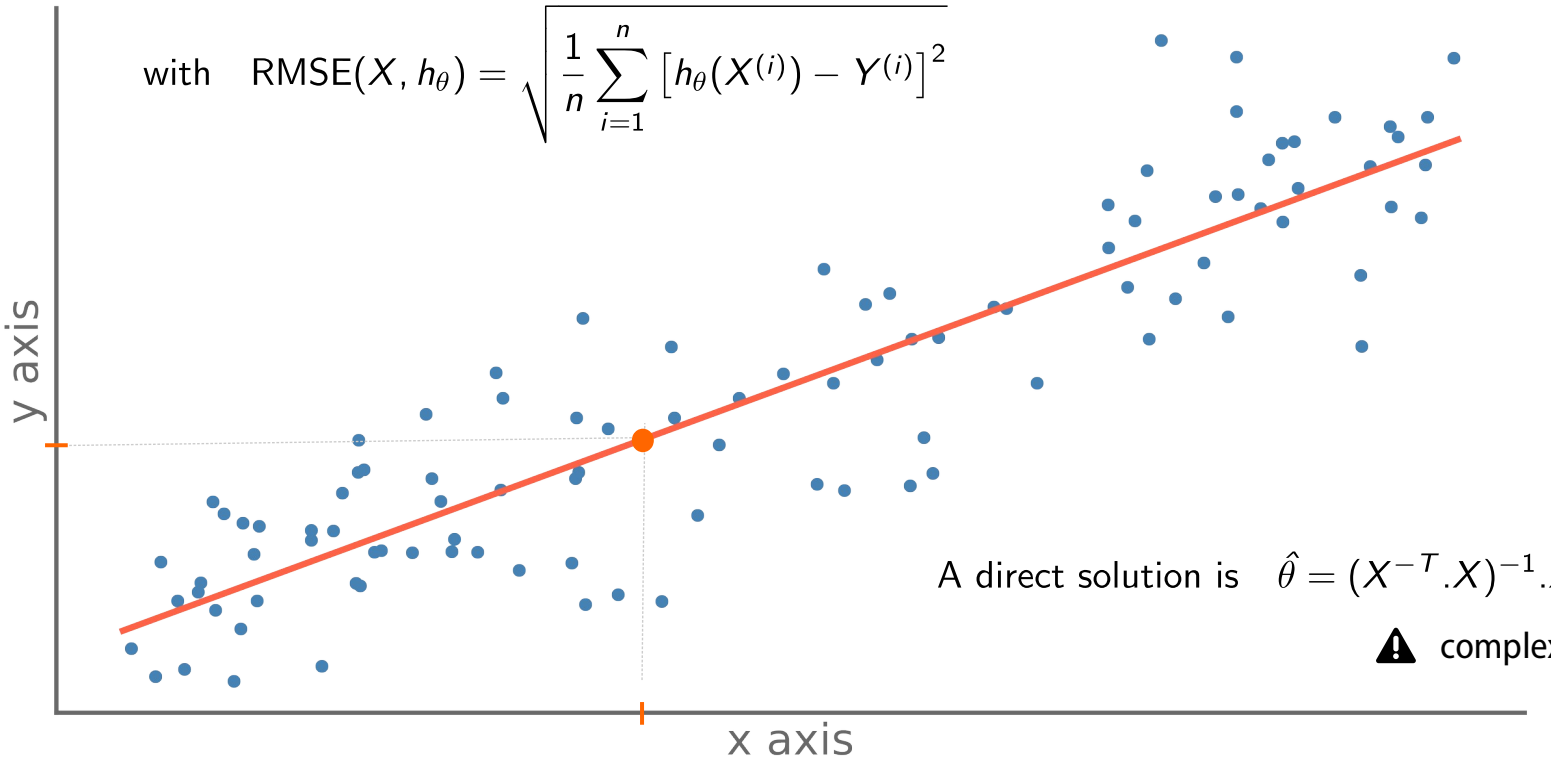
$$Y = X \cdot \theta + N \quad \text{with} \begin{cases} \theta = \begin{pmatrix} a \\ b \end{pmatrix} \text{ as } y = a \cdot x + b \\ N = \text{noise} \end{cases}$$



Linear regression

We search $\hat{\theta} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$ for which RMSE(X, $\hat{\theta}$) is minimal

with $\text{RMSE}(X, h_\theta) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[h_\theta(X^{(i)}) - Y^{(i)}\right]^2}$

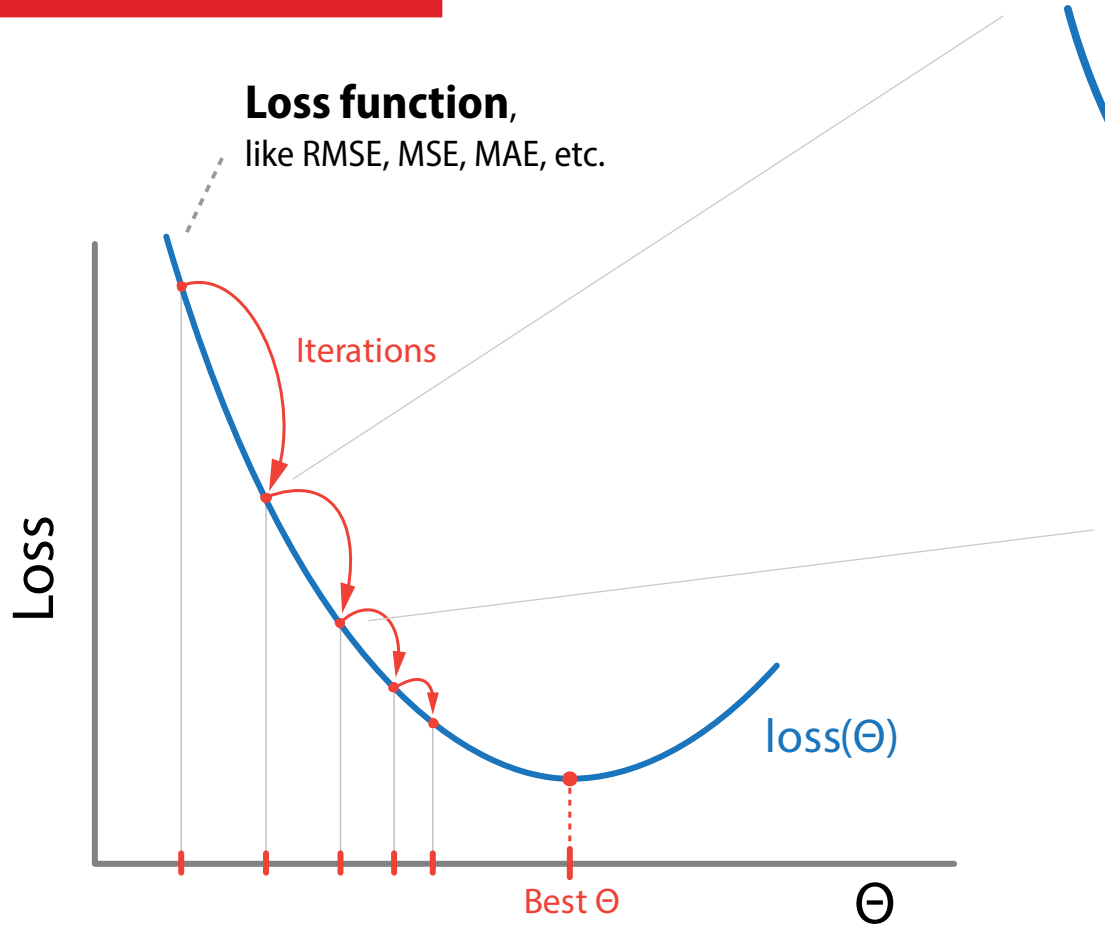


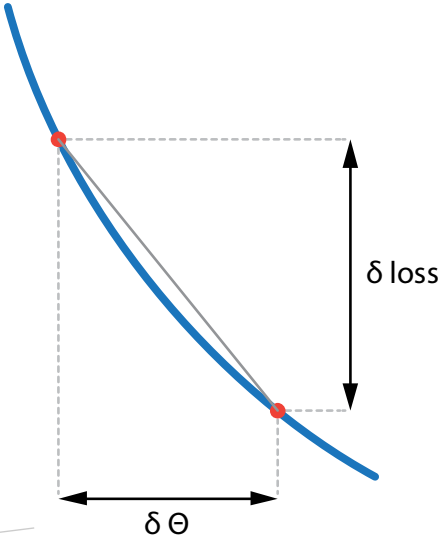
A direct solution is $\hat{\theta} = (X^{-T}.X)^{-1}.X^{-T}.Y$

⚠ complexity in n^3

RMSE : Root Mean Square Error
Erreur quadratique moyenne

Gradient descent





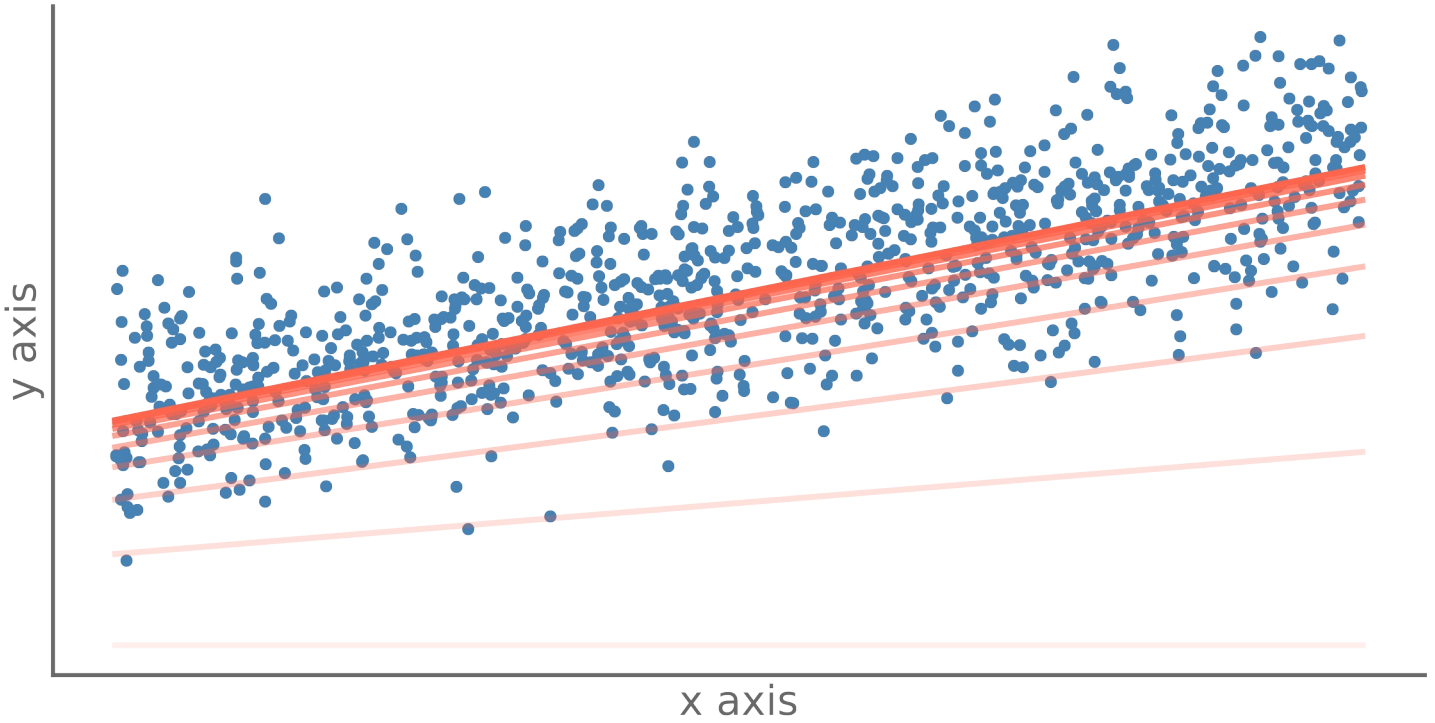
gradient $= \frac{\delta \text{loss}}{\delta \theta}$

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

where η is the learning rate

This process is called stochastic gradient descent and the function used to optimize the descent, optimization function

Gradient descent

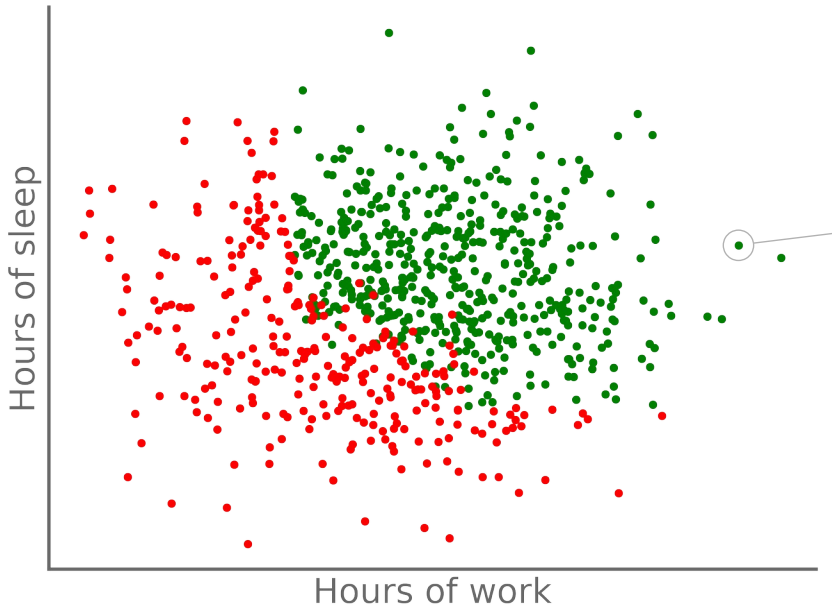


#i	Loss	Gradient		Theta	
0	+12.481	-6.777	-1.732	-3.388	+0.000
20	+4.653	-4.066	-1.039	-2.033	+0.346
40	+1.835	-2.440	-0.624	-1.220	+0.554
60	+0.821	-1.464	-0.374	-0.732	+0.679
80	+0.455	-0.878	-0.224	-0.439	+0.754
100	+0.324	-0.527	-0.135	-0.263	+0.799
120	+0.277	-0.316	-0.081	-0.158	+0.826
140	+0.260	-0.190	-0.048	-0.095	+0.842
160	+0.253	-0.114	-0.029	-0.057	+0.851
180	+0.251	-0.068	-0.017	-0.034	+0.857
200	+0.250	-0.041	-0.010	-0.020	+0.861

Logistic regression

A logistic regression is intended to provide a probability of belonging to a class.

Dataset : X Observations
y Classe



$$(X_i, y_i) \begin{cases} X_i = \begin{pmatrix} x_{i1} = \text{Hours of work} \\ x_{i2} = \text{Hours of sleep} \end{pmatrix} \\ y_i = \begin{cases} 1 & \text{belong to the class} \\ 0 & \text{don't belong} \end{cases} \end{cases}$$

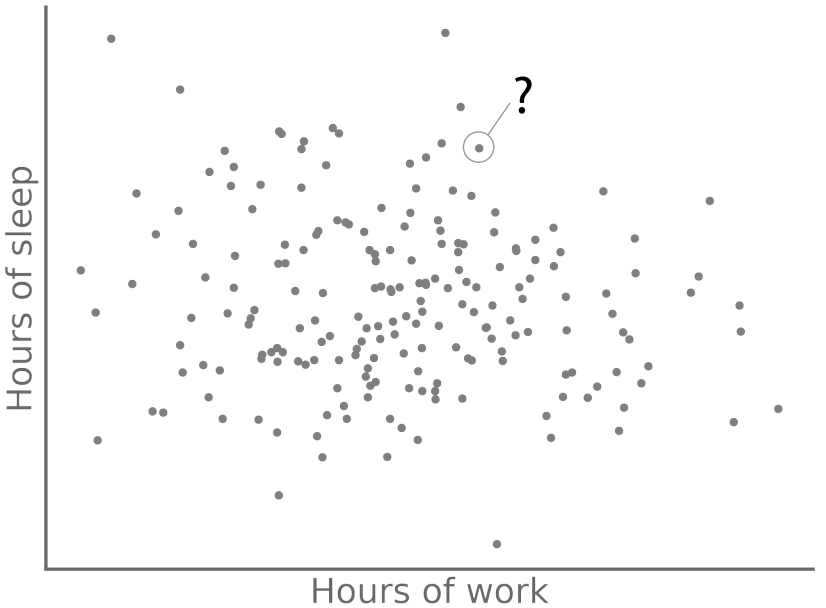
Notebook : FIDLE

Logistic regression

A logistic regression is intended to provide a probability of belonging to a class.

Dataset : X Observations
y Classe

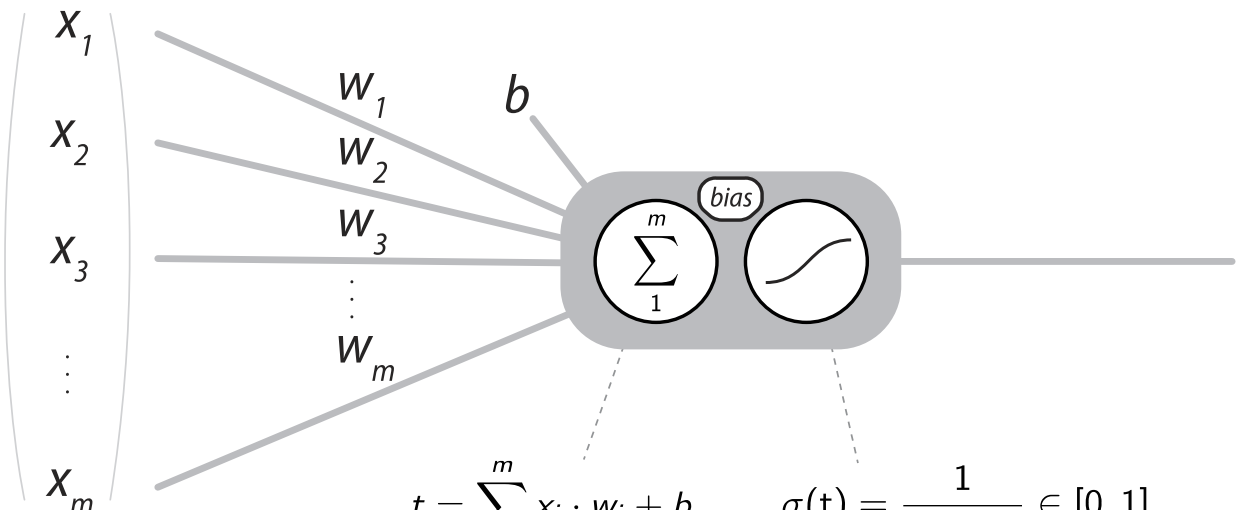
Objective : Predict the class
x given, we want to predict y
$\mathbf{y_{pred} = f(x)}$



Notebook : FIDLE

Logistic regression

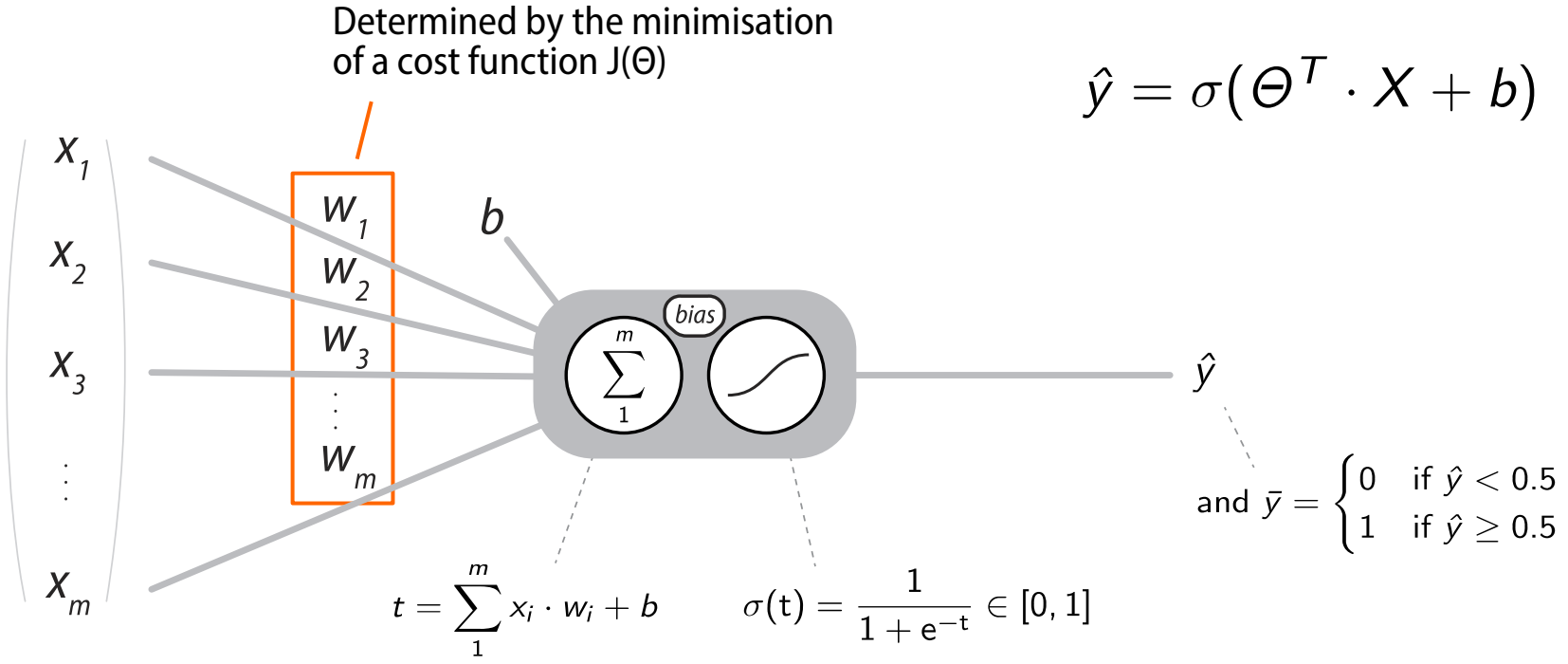
$$\hat{y} = \sigma(\Theta^T \cdot X + b)$$



and $\bar{y} = \begin{cases} 0 & \text{if } \hat{y} < 0.5 \\ 1 & \text{if } \hat{y} \geq 0.5 \end{cases}$

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

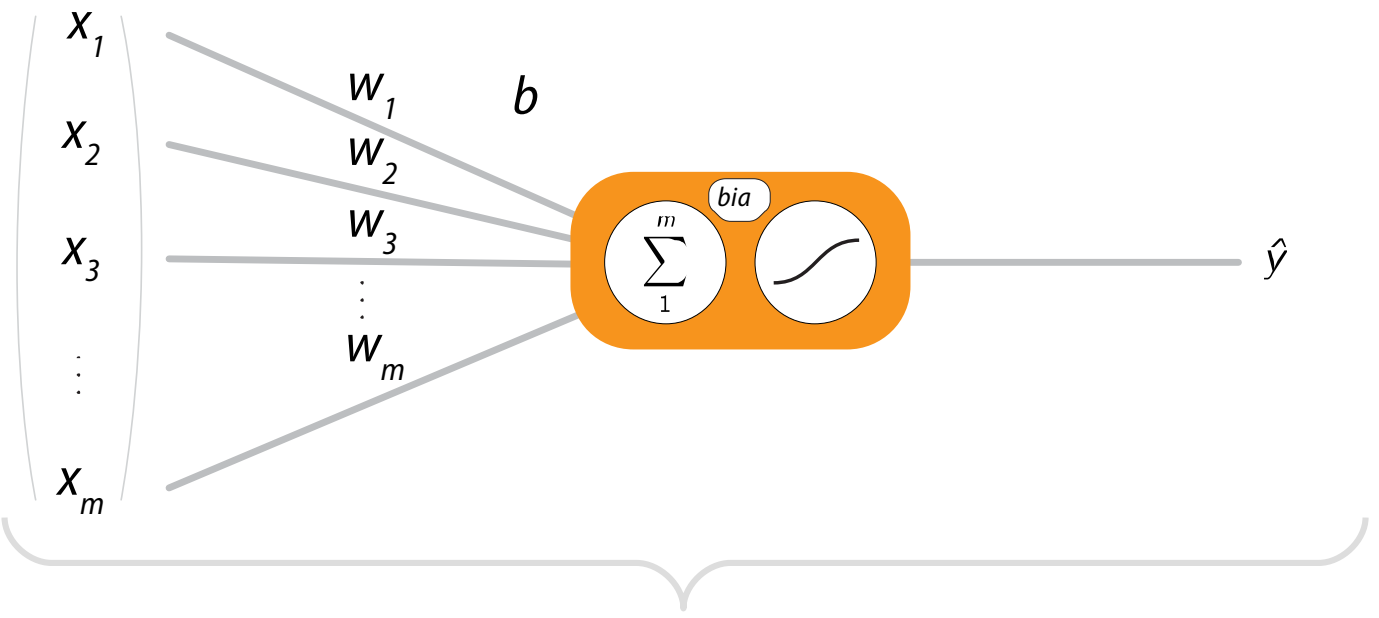
Logistic regression



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

Logistic regression

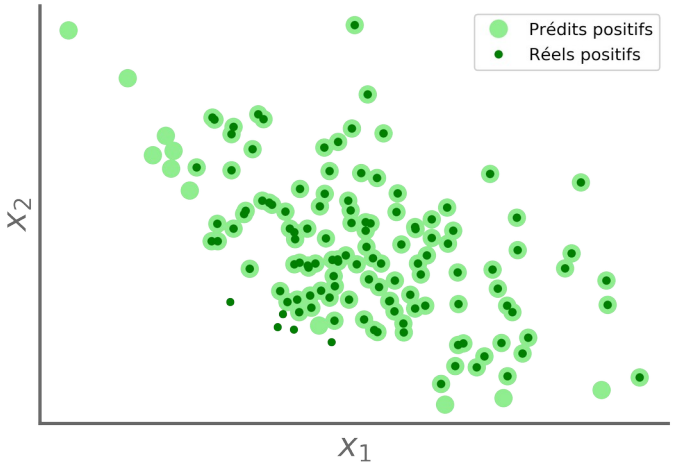
$$\hat{y} = \sigma(\Theta^T \cdot X + b)$$

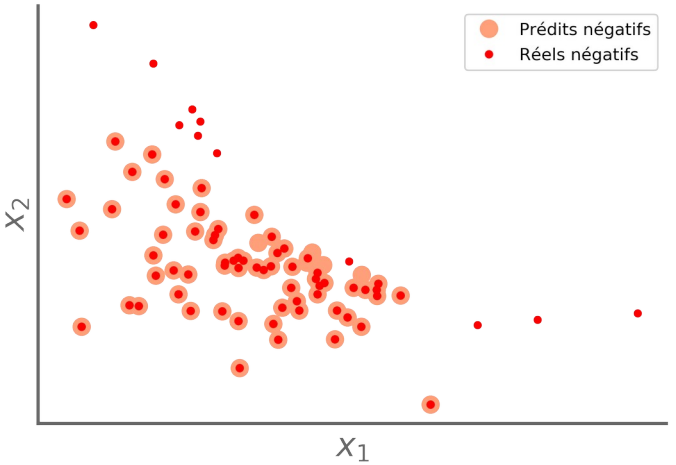


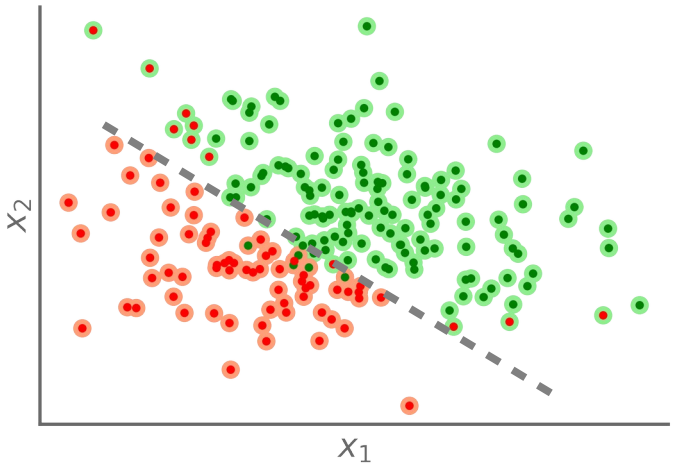
That's an « artificial neuron » !

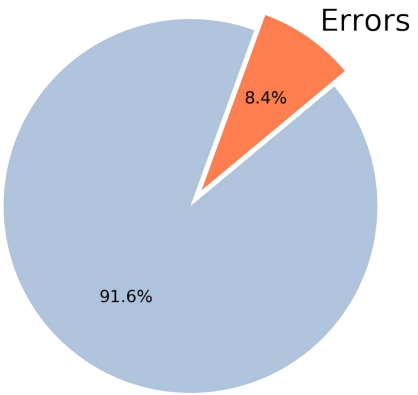
So, we have a neural network of… 1 neuron !

Logistic regression



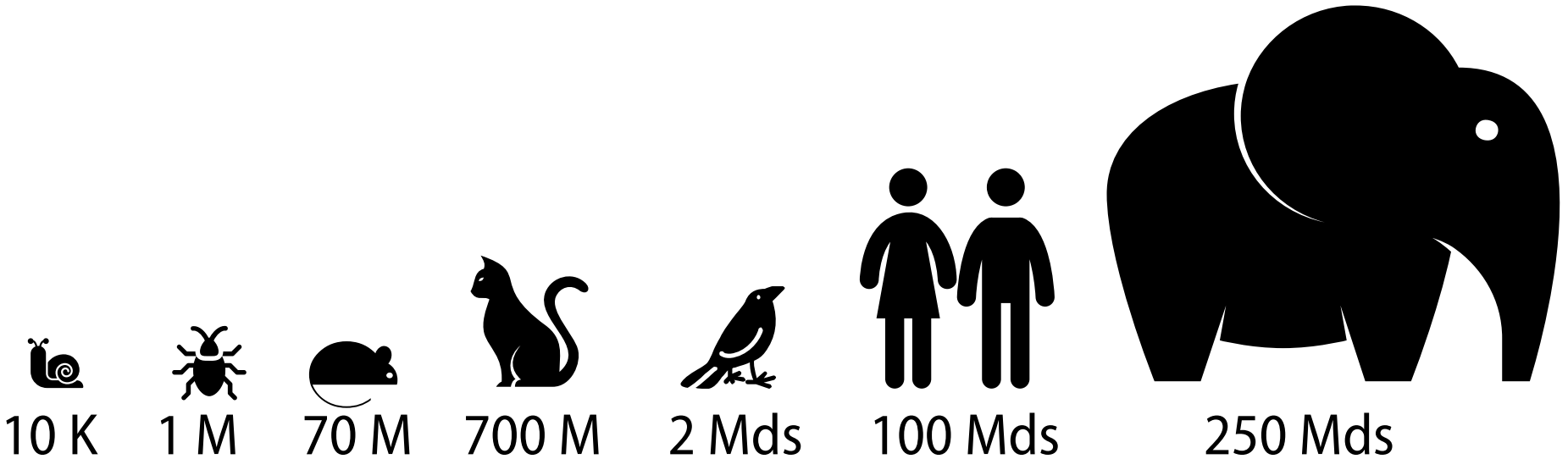


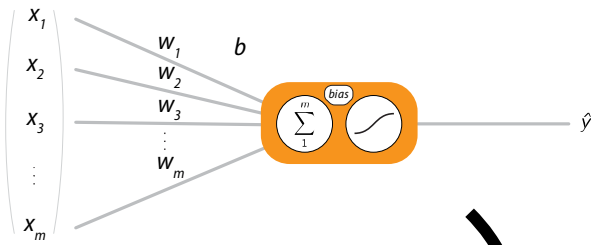


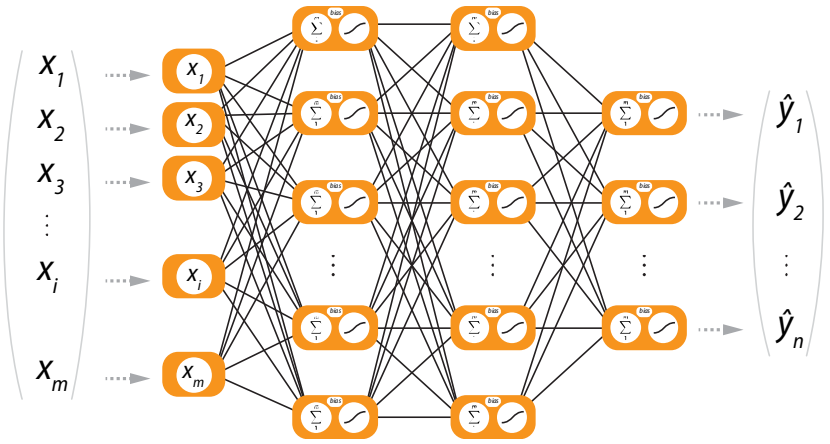


Notebook : FIDLE

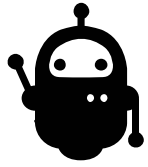
One neuron is **good**... but more than one is **better** !







- Comment passe t-on d'un neurone unique au Deep Learning ?
- Comment cette évolution s'est-elle déroulée ?
- Comment utiliser un réseau et pour quoi faire ?



Demain :
Mardi 8 Juillet
T8 / 9h30

Un autre paradigme pour les sciences numériques

..et bien plus encore !
A demain :-)

References

[JGRAY] Gray, J. (2001), from « The Fourth Paradigm: Data-Intensive Scientific Discovery » Tony Hey, Stewart Tansley, Kristin Tolle (2009). Published by Microsoft Research. ISBN: 978-0-9825442-0-4

[MCPIT] McCulloch, Warren; Walter Pitts (1943). "A Logical Calculus of Ideas Immanent in Nervous Activity". Bulletin of Mathematical Biophysics. 5 (4): 115–133. doi:10.1007/BF02478259

[DHEBB] Hebb, D. O. (1949). « The Organization of Behavior: A Neuropsychological Theory. » New York: Wiley and Sons. ISBN 9780471367277.

[FROS] Rosenblatt, Frank. (1958). « The perceptron: A probabilistic model for information storage and organization in the brain. » Psychological Review, 65(6), 386-408.

[MIPA] Minsky, Marvin; Papert, Seymour. (1969). « Perceptrons : An Introduction to Computational Geometry », MIT Press

[DRUM] Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (1986). « Learning representations by back-propagating errors ». Nature. 323 (6088): 533–536. doi:10.1038/323533a0.

[YLEC1] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, « Backpropagation Applied to Handwritten Zip Code Recognition », AT&T Bell Laboratories

[LRDN] Dominique Cardon, Jean-Philippe Cointet, Antoine Mazieres. (2018). « La revanche des neurones », Réseaux, La Découverte, 5 (211), <10.3917/res.211.0173>. <hal-01925644>

[AMAZ] Antoine Mazieres (2016) Thèse : « Cartographie de l'apprentissage artificiel et de ses algorithmes » Université Paris 7 Denis Diderot, <hal-01771655>

[TOP500] Statistics on top 500 high-performance computers. (2018) « Exponential growth of supercomputing power as recorded by the TOP500 list ». https://www.top500.org

[WKP1] Wikipedia/en. (2018) « List of datasets for machine-learning research ». https://en.wikipedia.org

[WOS1] Core database : TS=("support vector machine*" OR ("SVM" AND "classification") OR ("SVM" AND "regression") OR ("SVM" AND "classifier") OR "support vector network*" OR ("SVM" AND "kernel trick*"))

[WOS2] Core database : TS=("deep learning" OR "deep neural network*" OR ("DNN" AND "neural network*") OR "convolutional neural network*" OR ("CNN" AND "neural network*") OR "recurrent neural network*" OR ("LSTM" AND "neural network*") OR ("RNN*" AND "neural network*"))

[ALEX] A. Krizhevsky, I. Sutskever, G. Hinton. (2012). « ImageNet Classification with Deep Convolutional Neural Networks » doi: 10.1145/3065386

[ILSVRC] ImageNet Large Scale Visual Recognition Challenges http://image-net.org/challenges/LSVRC/<2012..2017>/results https://en.wikipedia.org/wiki/ImageNet

[MOBIN] Howard, Andrew G. et al. (2017) "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." https://arxiv.org/abs/1704.04861

References

[W2VEC] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean (2013), « Distributed Representations of Words and Phrases and their Compositionality », https://arxiv.org/abs/1310.4546

[GLOVE] Jeffrey Pennington, Richard Socher, Christopher D. Manning (2014) « GloVe: Global Vectors for Word Representation », http://nlp.stanford.edu/projects/glove/

[P2VEC] Ehsaneddin Asgari, Mohammad R.K. Mofrad, (2016), « ProtVec: A Continuous Distributed Representation of Biological Sequences », https://arxiv.org/abs/1503.05140

[LSTM] Sepp Hochreiter, Jürgen Schmidhuber, (1997), « Long Short-Term Memory, https://doi.org/10.1162/neco.1997.9.8.1735

[GRU] Cho, Kyunghyun; van Merrienboer, Bart; Gulcehre, Caglar; Bahdanau, Dzmitry; Bougares, Fethi; Schwenk, Holger; Bengio, Yoshua (2014), « Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation ». https://arxiv.org/abs/1406.1078

[CARTP] AG Barto, RS Sutton and CW Anderson, (1983), « Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problem », IEEE Transactions on Systems, Man, and Cybernetics, 1983

[GAN] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, (2014), « Generative Adversarial Networks » https://arxiv.org/abs/1406.2661

[WOS3] Core database : TS=('material' and ('design' or 'discovery' or 'optimization') and ('deep learning' or 'machine learning' or 'neurons'))

[AIDEX] AI Index. « A starting point for informed conversations about progress in artificial intelligence. The report aggregates a diverse set of metrics, and makes the underlying data easily accessible to the general public ». https://aiindex.org

[DLPW] Jeff Hale, « Deep Learning Framework Power Scores 2018 » and « 2019 Deep Learning Framework Growth Scores » http://bit.ly/2NagcgH and http://bit.ly/3hUGqlS

[CNIL1] Comment permettre à l'homme de garder la main ? Synthèse du débat public animé par la cnil dans le cadre de la mission de réflexion éthique confiée par la loi pour une république numérique. https://www.cnil.fr/fr/comment-permettre-lhomme-de-garder-la-main-rapport-sur-les-enjeux-ethiques-des-algorithmes-et-de

[CNIL2] Reconnaissance faciale : pour un débat à la hauteur des enjeux 15 novembre 2019 https://www.cnil.fr/fr/reconnaissance-faciale-pour-un-debat-la-hauteur-des-enjeux

Illustrations

[POTATO]	From *Die Giftpflanzen Deutschlands,* Peter Esser, 1910, via iconspng.com
[CONVO]	An Introduction to different Types of Convolutions in Deep Learning https://towardsdatascience.com/types-of-convolutions-in-deep-learning-717013397f4d
[NEURON]	Wikimedia Commons, the free media repository.
Photos	pixels.com
Icons	thenounproject.com