

DataScience & IA

Christophe Rodrigues

Plan du cours

Challenge de classification

Projet final

Réseaux de neurones

Du Perceptron au Deep Learning

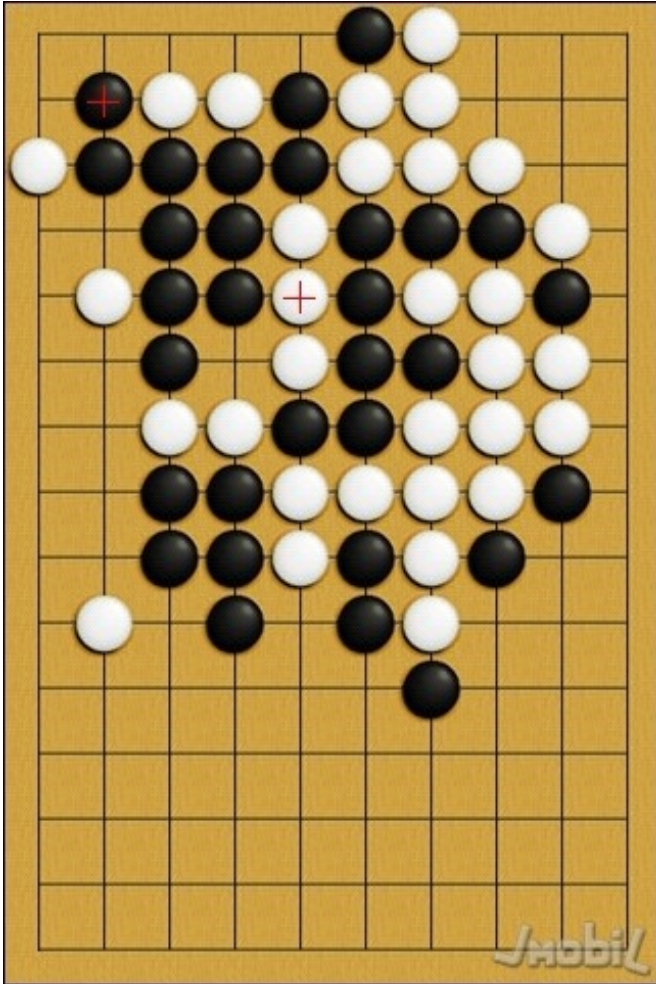
Quiz

Challenge de classification

- 1) Un dataset anonyme avec plusieurs dimensions en entrée et une variable à prédire vous sera communiqué.
- 2) vous entraînerez votre modèle sur ce dataset en divisant en deux parties : une pour l'apprentissage, une pour le test de votre modèle. Vous êtes libres de le faire comme vous le voulez, la seule contrainte est d'utiliser un K plus proche voisin.
- 3) un dataset sur le même problème vous sera fourni mais sans l'étiquette que vous devrez prédire vous-même à l'aide de votre modèle développé en étape 2).
- 4) vous devrez envoyer les étiquettes que vous avez trouver que l'on pourra comparer aux étiquettes réelles et vous attribuer un score.

Les instructions détaillées suivront

Projet final : Gomoku



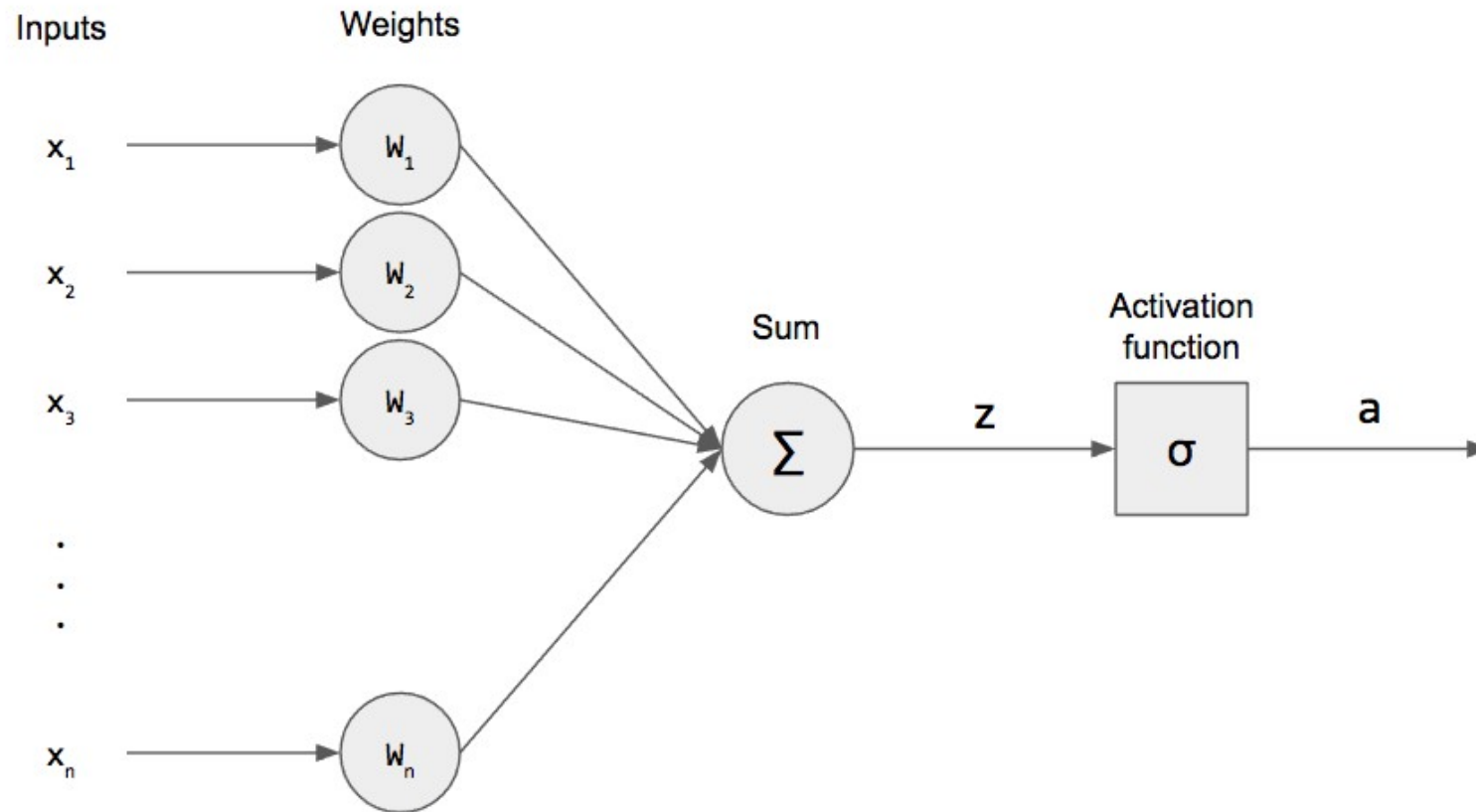
Matrice de 19x19 positions

Exactement le même principe que le Morpion, excepté qu'il faille aligner 5 pions au lieu de 3

A faire par groupe de 3/4 étudiants
Présentation/Démonstration lors de la dernière séance de TP

Réseaux de neurones

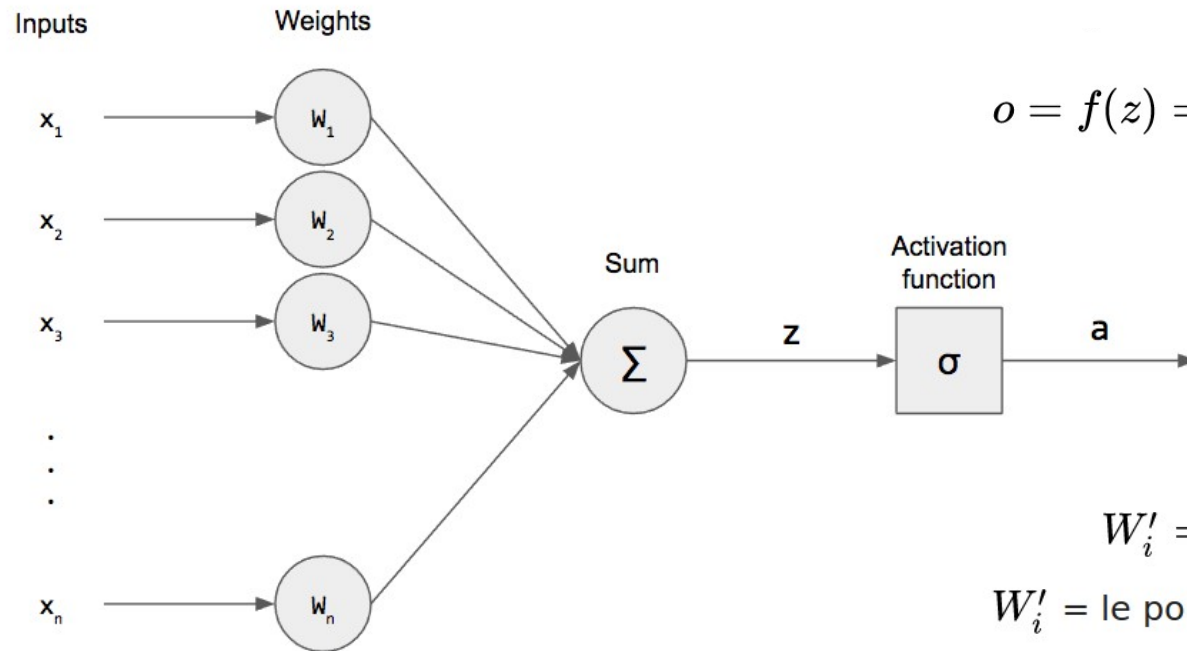
Perceptron



Rosenblatt 1952

Réseaux de neurones

Perceptron



$$o = f(z) = \begin{cases} 1 & \text{si } \sum_{i=1}^n w_i x_i > \theta \\ 0 & \text{sinon} \end{cases}$$

$$W'_i = W_i + \alpha(Y_t - Y)X_i$$

W'_i = le poids i corrigé

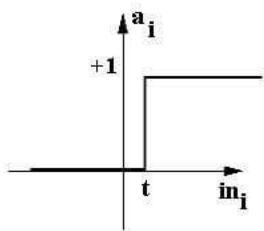
Y_t = sortie attendue

Y = sortie observée

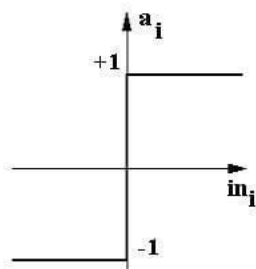
α = le taux d'apprentissage

X_i = l'entrée du poids i pour la sortie attendue Y_t

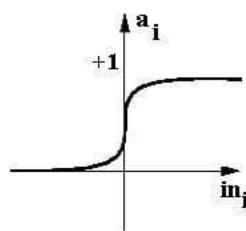
W_i = le poids i actuel



Step Function



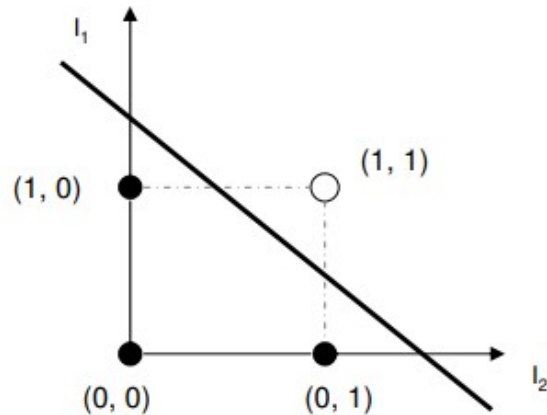
Sign Function



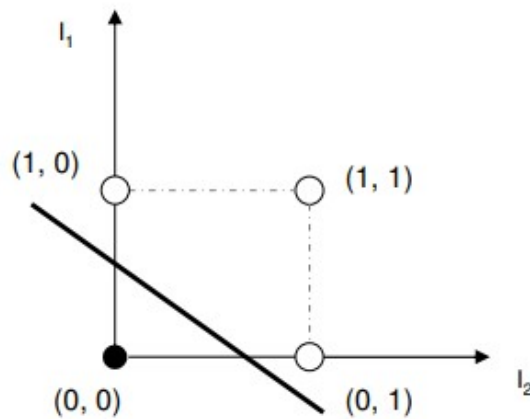
Sigmoid Function

Perceptron + XOR = L'hiver de l'IA

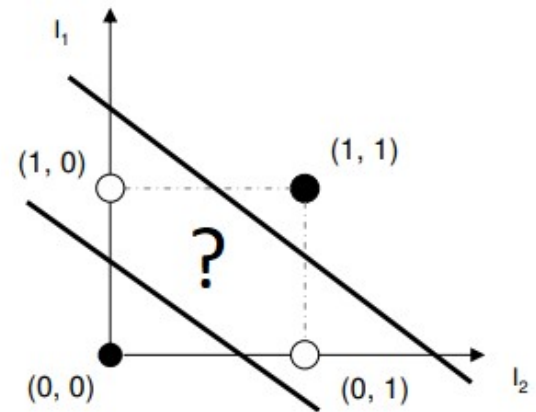
AND		
I_1	I_2	out
0	0	0
0	1	0
1	0	0
1	1	1



OR		
I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	1



XOR		
I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	0

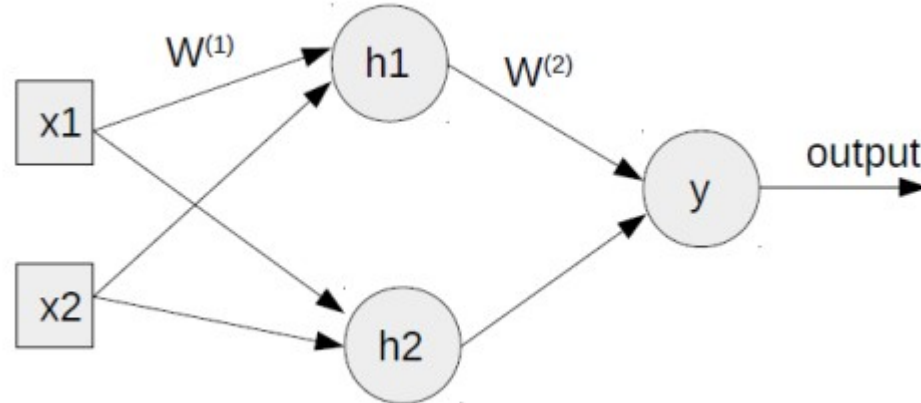


Minsky & al. 1969

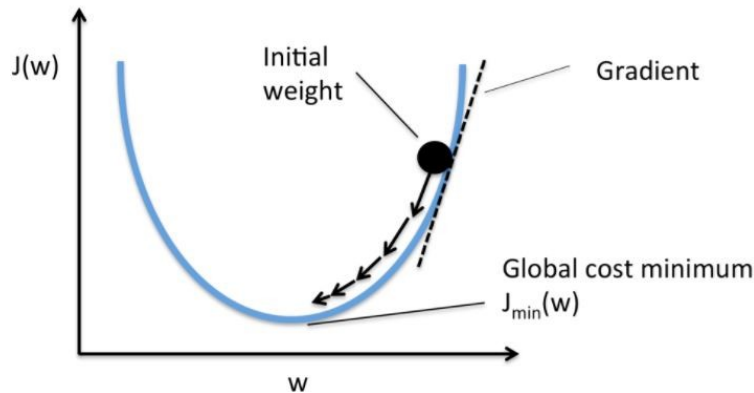
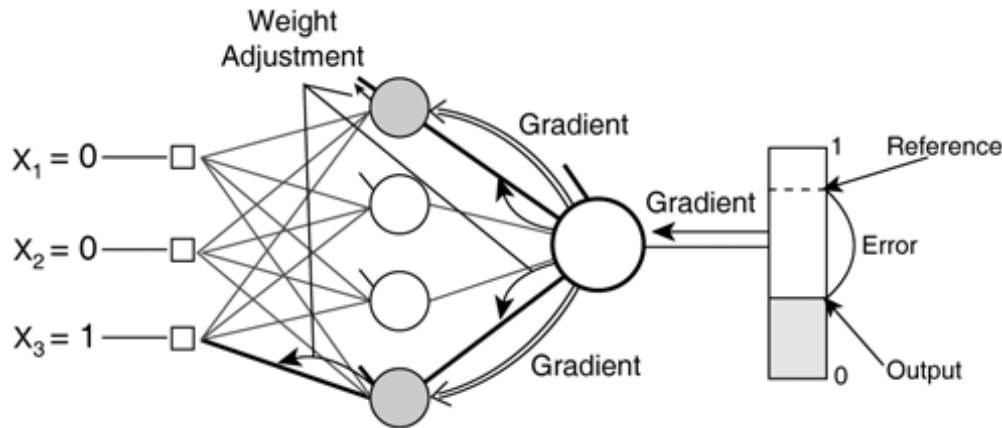
...Solution plus de 15 ans après

Perceptron multicouche avec rétropropagation du gradient (en pratique en 1986, Rumelhart)

Perceptron avec une couche cachée



Perceptron multicouches



Rétropropagation du gradient

Propager chaque exemple dans le réseau

Calculer l'erreur en sortie puis le gradient

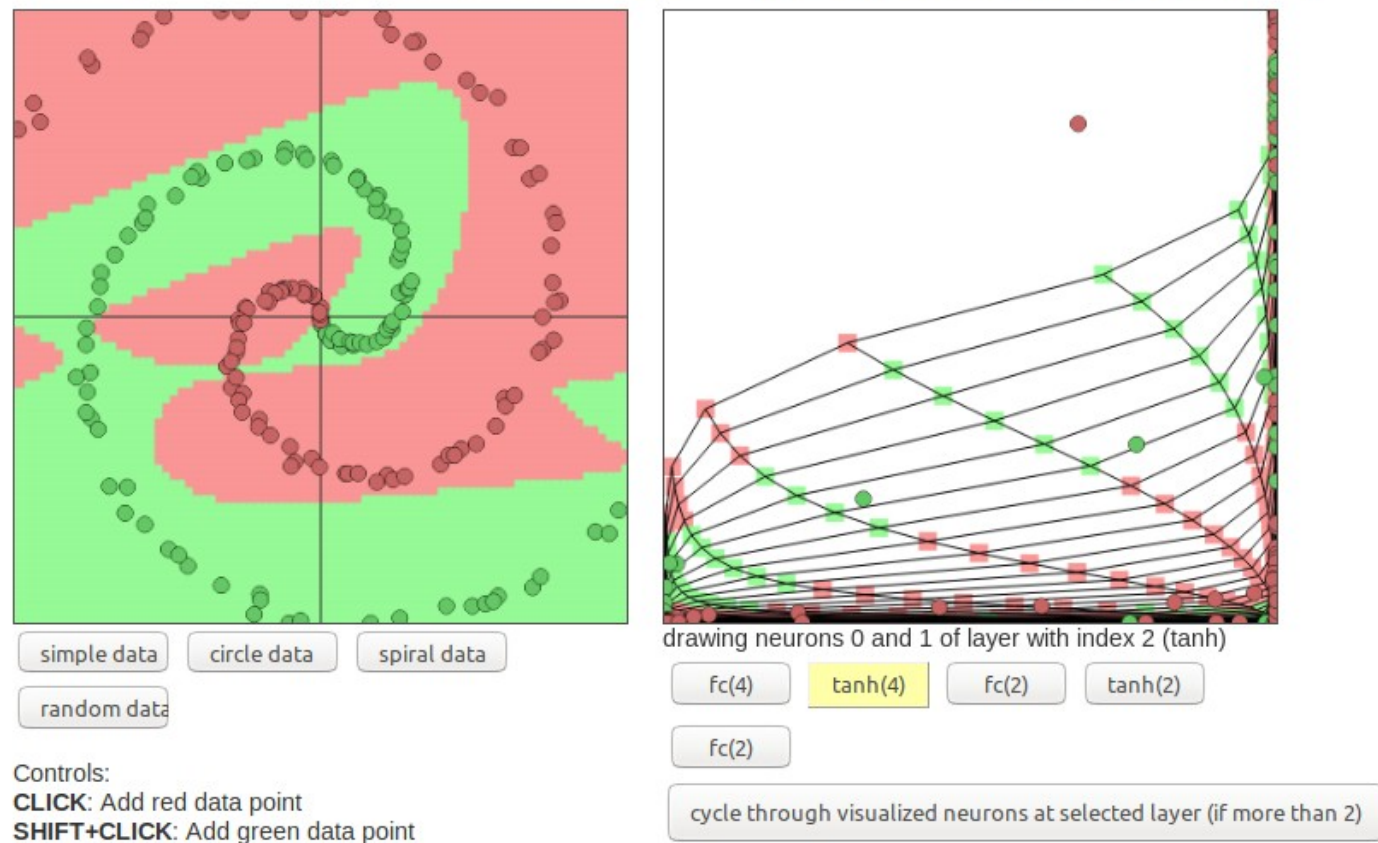
Modifier les poids dans la direction opposée au gradient

3 principales variantes :

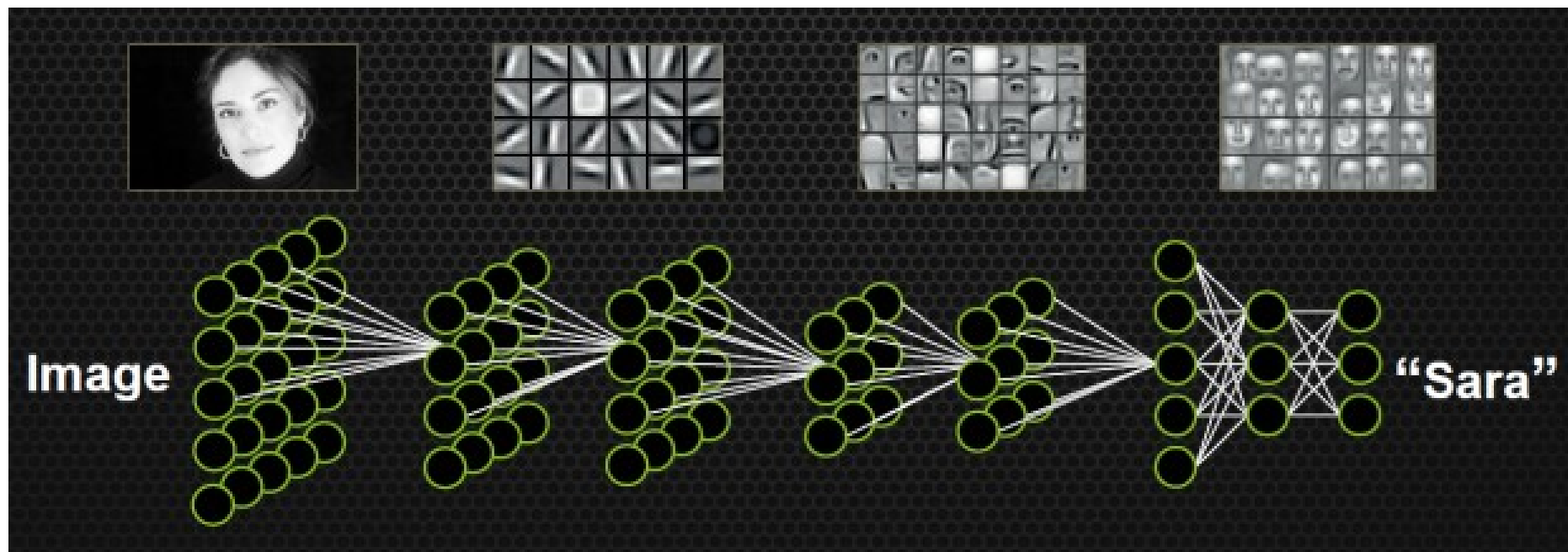
- batch (précis mais long)
- stochastic (instable mais rapide)
- mini-batch (meilleur des 2 mondes)

Inside Black Box : ConvnetJS

<https://cs.stanford.edu/people/karpathy/convnetjs//demo/classify2d.html>

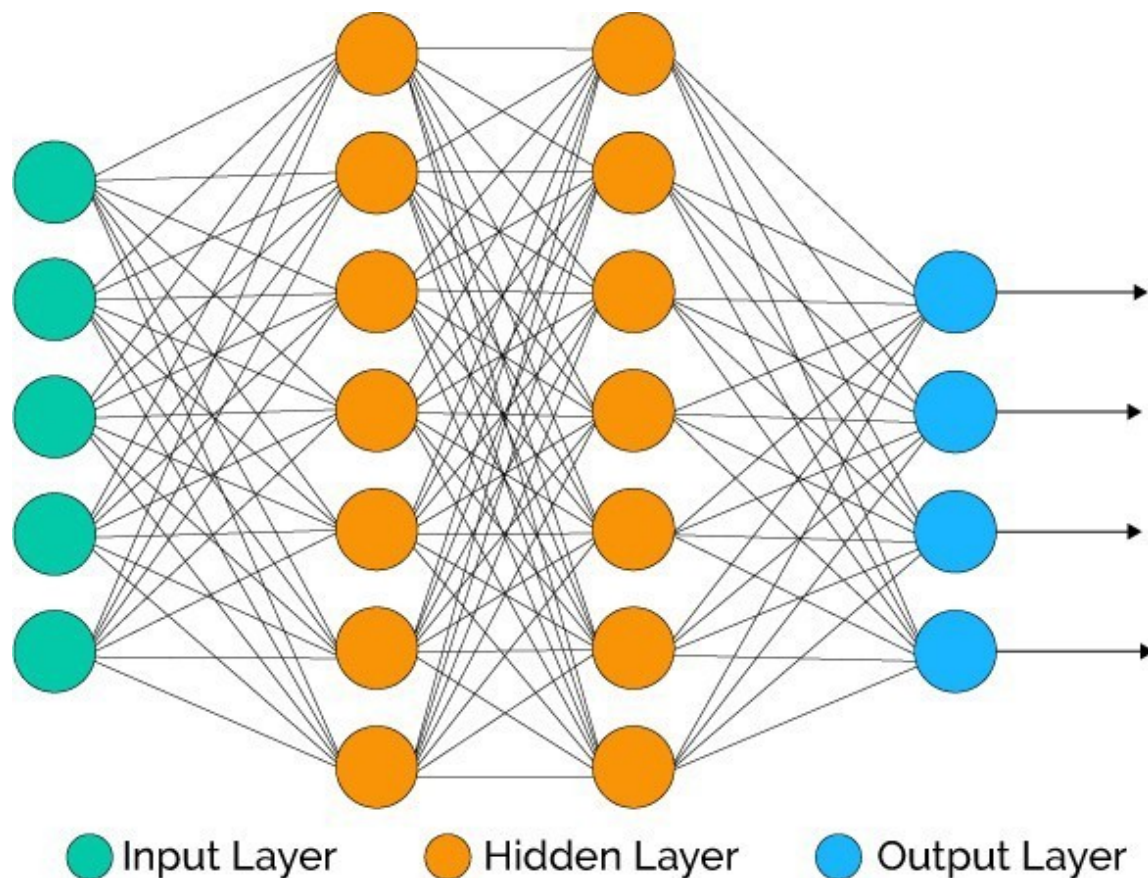


Apprentissage d'une représentation



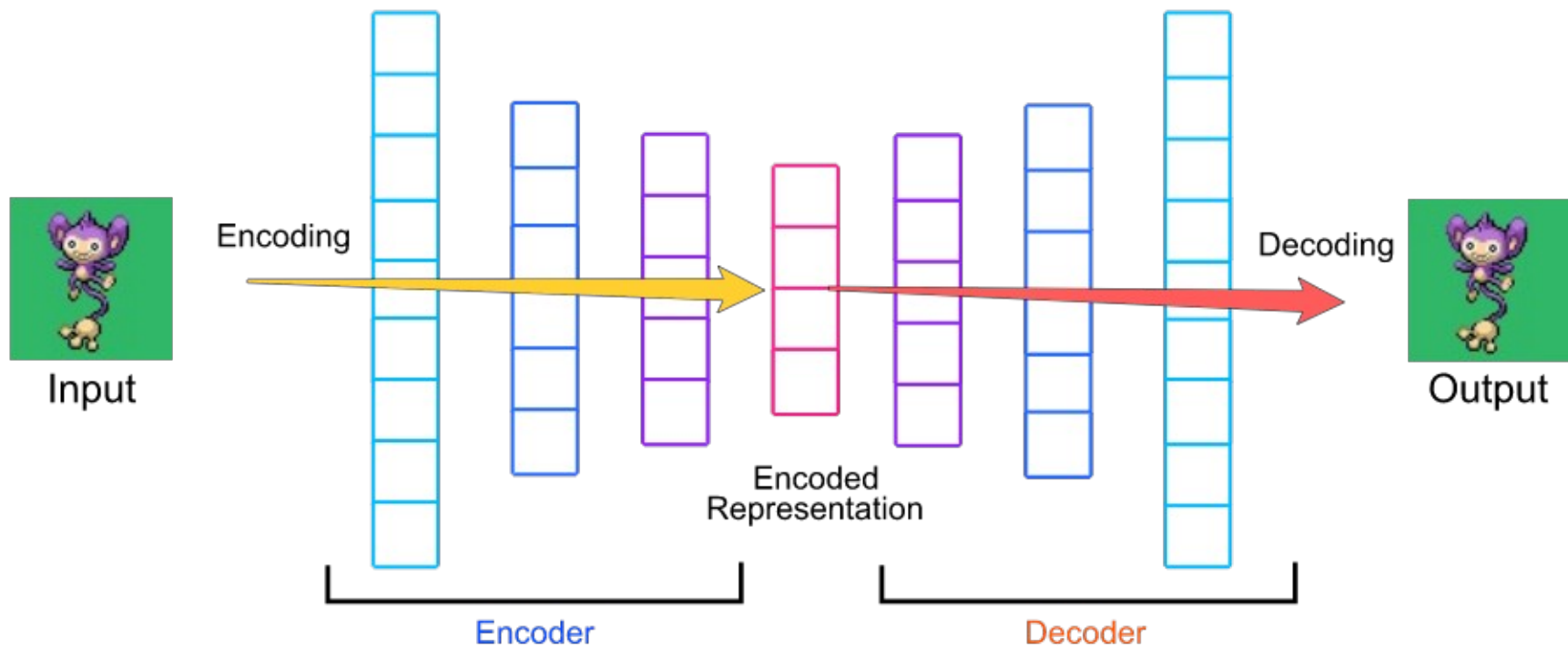
Le réseau part des pixels et apprend sa propre représentation des images

Grandes familles de réseaux



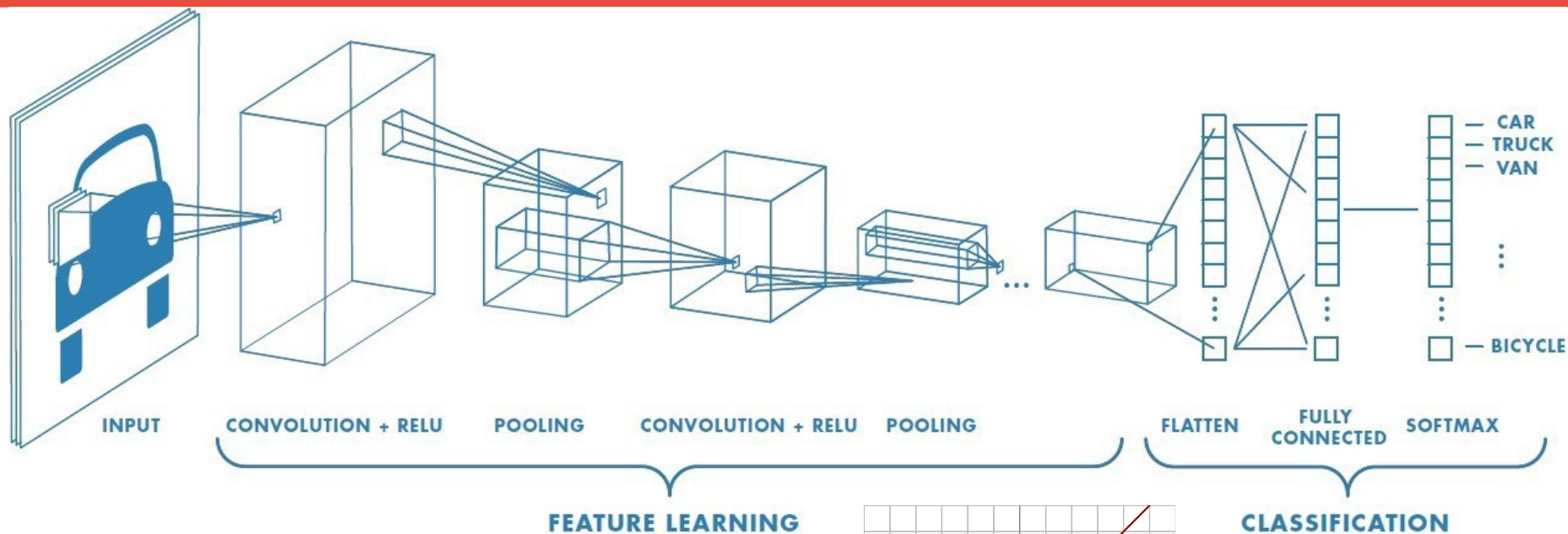
Réseau dense

AE - Auto-Encoder



Très utilisés pour le transfert de connaissances

CNN - Convolutional Neural Networks



convolution

-0.2	0.0	0.5
1.0	0.3	-0.6
0.0	0.0	0.8

0.0	0.0	0.0
0.8	-0.5	0.8
0.0	-0.2	0.0

0.4	0.2	-0.2
-0.8	0.0	0.8
0.0	-0.5	0.2



RELU

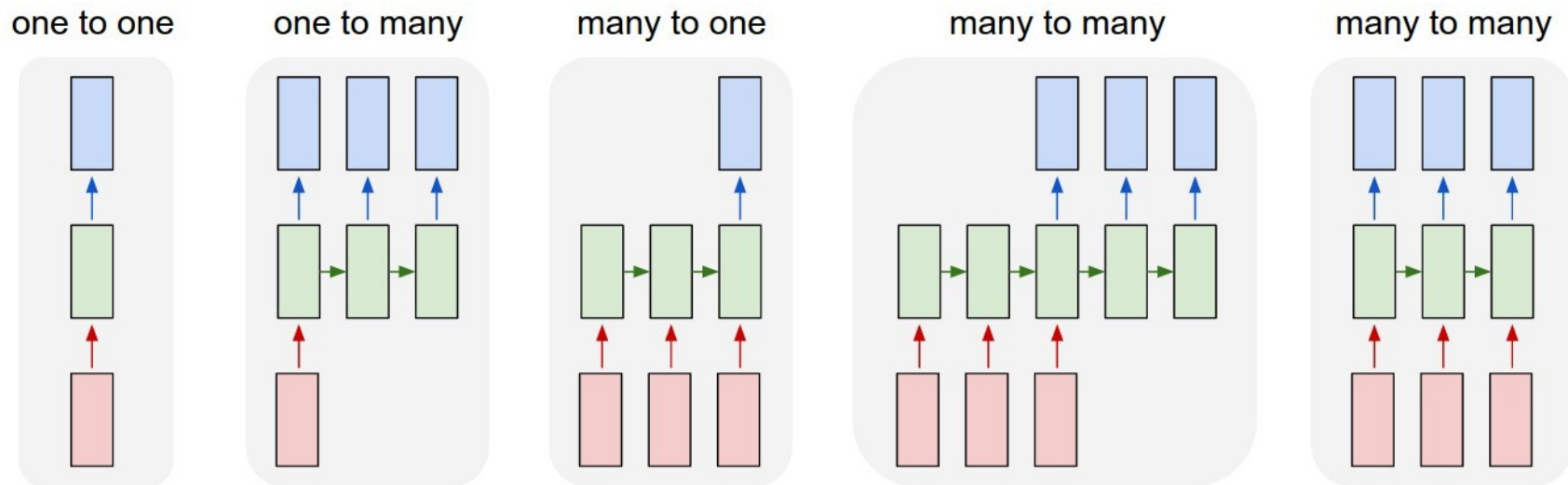


pooling



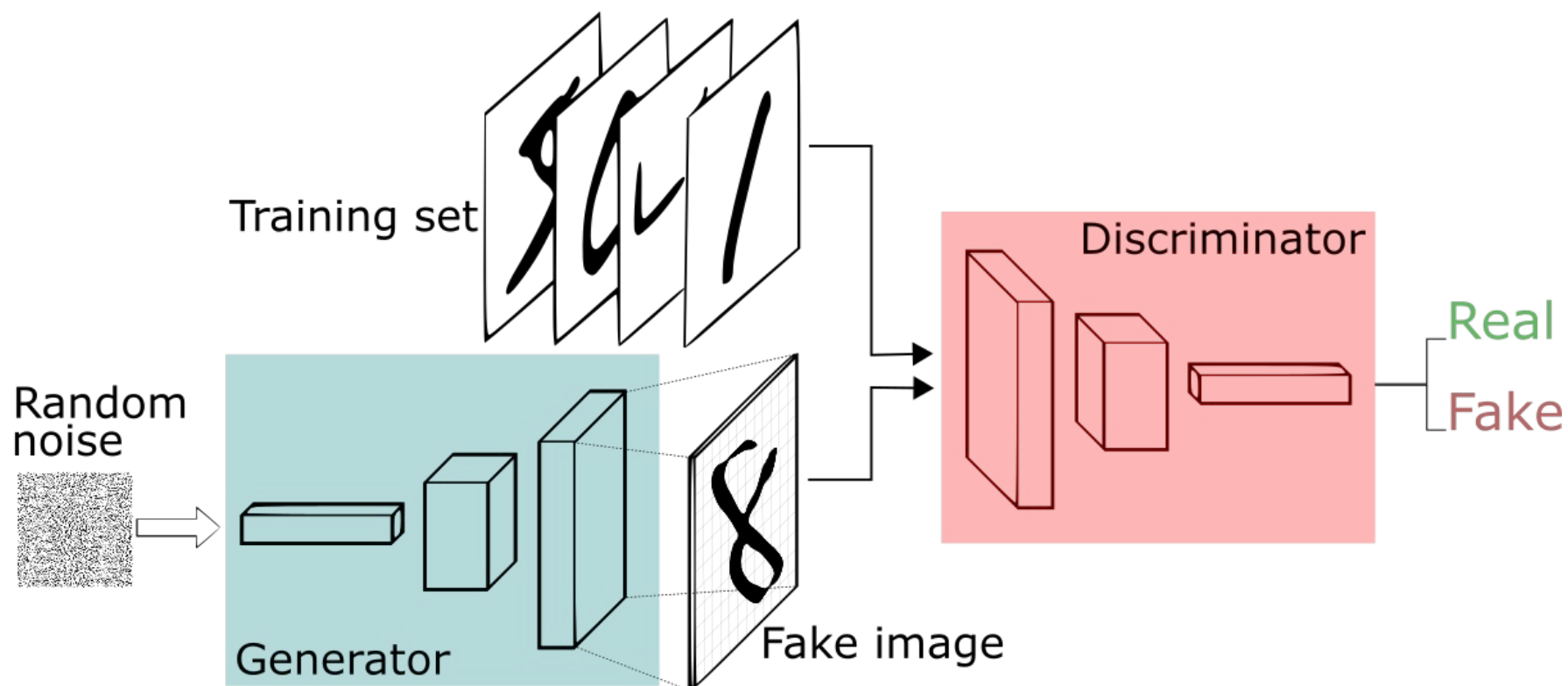
Dimensions reduced by half

RNN – Recurrent Neural Networks

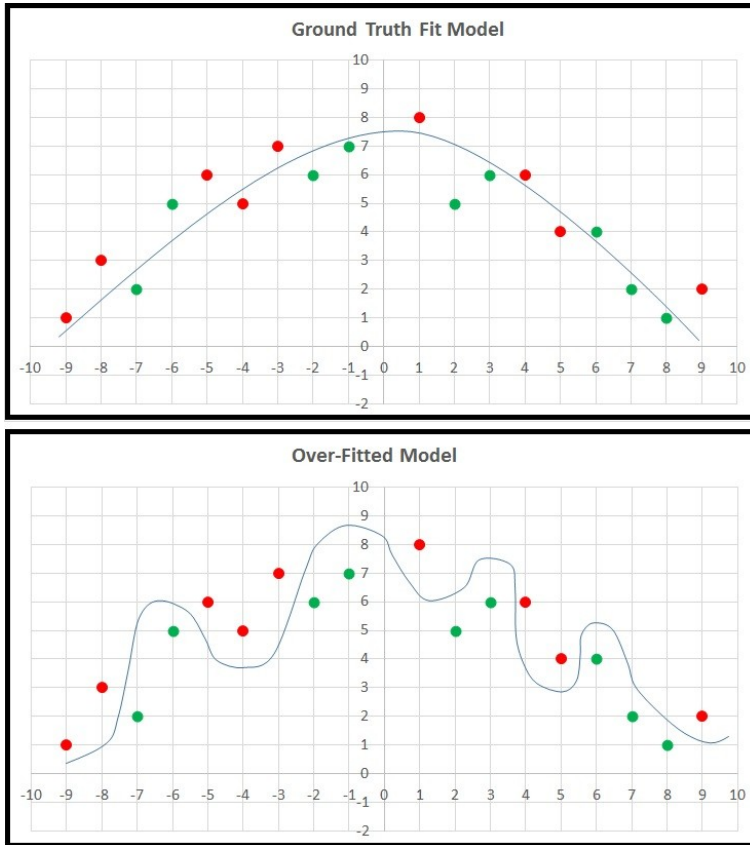


Traitement des séquences (ex : traduction de texte)

GAN – Generative Adversarial Networks

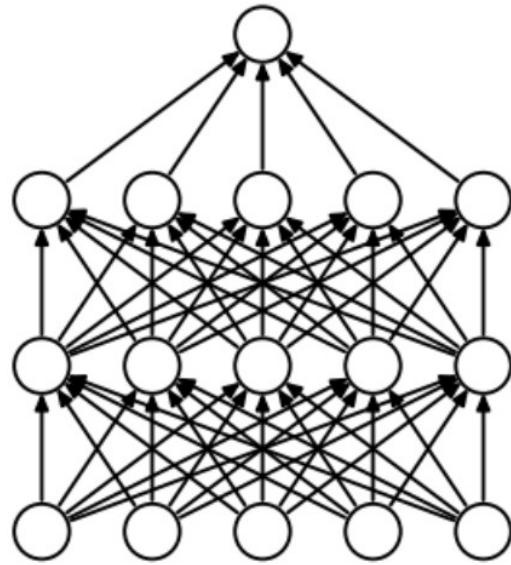


Régularisation : Contraindre les poids

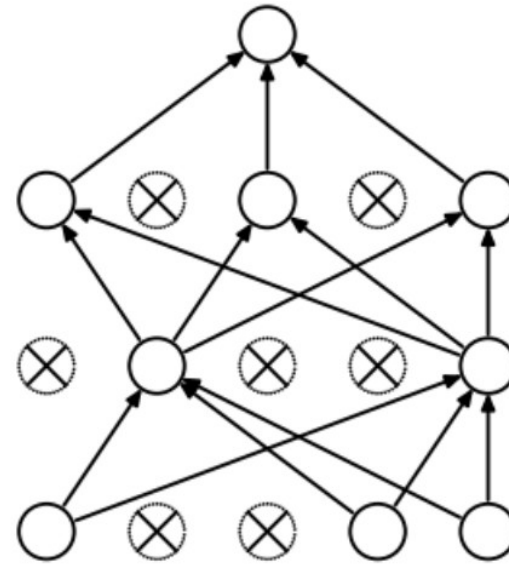


- Sur la courbe du haut, il y a 4 points mal prédits.
- Sur la courbe du bas, il n'y a aucune erreur mais elle semble faire un sur-apprentissage.
- La courbe du haut plus simple est préférable (rasoir d'Occam) car elle généralise mieux et est supposée meilleure sur des exemples jamais vus.
- Concrètement on va ajouter au calcul de l'erreur un terme restreignant la distribution des poids dans le réseau.

Régularisation : Dropout



(a) Standard Neural Net



(b) After applying dropout.

Itérativement, En apprentissage lors de la rétropropagation du gradient, certains neurones choisis aléatoirement vont être gelés et ne seront pas mis à jour.

Le but est d'éviter des co-adaptations complexes pouvant engendrer du sur-apprentissage

Eviter le sur-apprentissage

Avoir plus de données

Réduire la capacité du réseau

Régulariser les poids

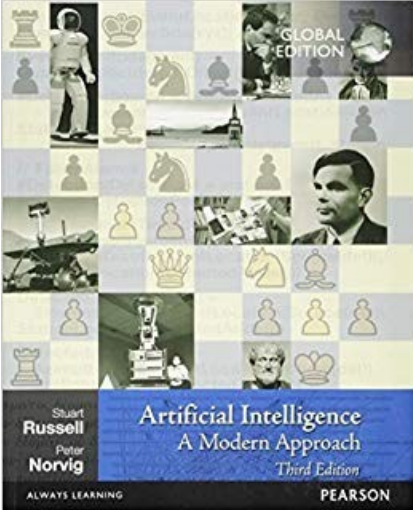
Utiliser le dropout

StarCraft - AlphaStar IA

Dans le Machine Learning de reddit : « We understand that the consensus among Starcraft 2 scene is that AlphaStar won the games through superhuman army control and that superior strategic thinking wasn't even needed. »

https://www.reddit.com/r/MachineLearning/comments/ak3v4i/d_an_analysis_on_how_alphastars_superhuman_speed/

Pour continuer...

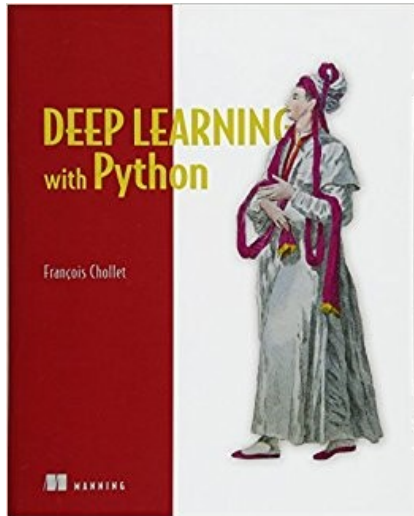


Artificial Intelligence A modern Approach
Stuart Russell, Peter Norvig
3ème édition 2011, 1152 pages

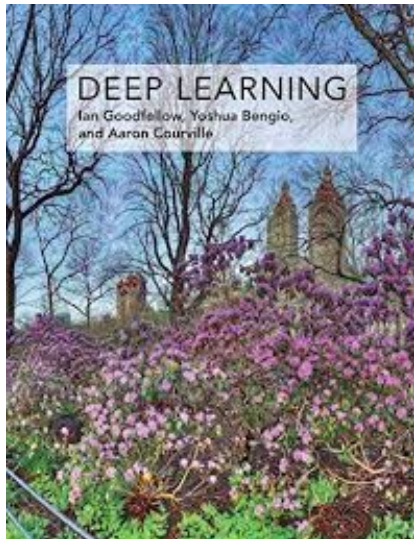


Apprentissage Artificiel
Antoine Cornuéjols, Laurent Miclet, Vincent barra
3ème édition 2018, 899 pages

Pour continuer...



Deep Learning with Python
François Chollet
2017, 384 pages



Deep Learning
Ian Goodfellow, Yoshua Bengio, Aaron Courville
2016, 775 pages

Pour continuer à l'ESILV...

Année 4 - Machine Learning par M. Ghassany

Année 4 - Deep Learning par Mme Alizadeh

Année 5 - Apprentissage par M. Oblin

Année 5 - Advanced Machine Learning for Big Data and Text Processing par M. Rodrigues