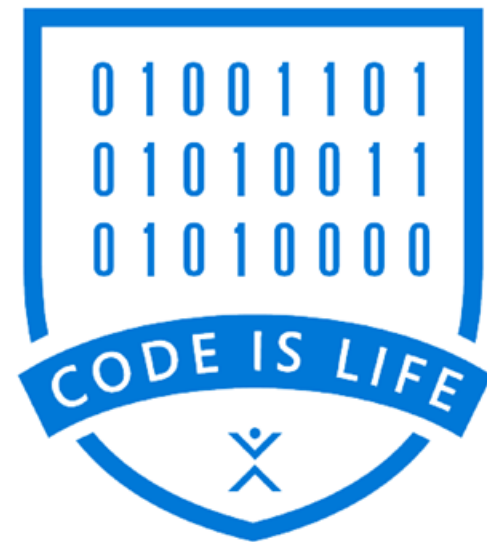


Microsoft Student Partners

# Neural Networks Comprehension



Nathan Bangwa {AI Lead}

- Neurone Biologique et Neurone Artificiel (Formel)
- Biais (bias) et Fonction d'Activation (AF)
- Apprentissage par perceptron
- Limite du perceptron mono-couche



- Perceptron Multi-couches (MLP)
  - Cost and optimization (learning)
    - Back-Propagation (retro-propagation de gradient)
- Deep Learning
  - Types des réseaux de neurones (architectures)
- What's Convolutional Neural Network



# What's Learning Model



- Un modèle d'apprentissage est soit une fonction mathématique (ou logique) simple ou composée, soit une logique(schema) à suivre pour réaliser une tâche.
- C'est à la machine de trouver ce modèle d'apprentissage (fonction ou logique à suivre) partant de données qui lui sont fournies (supervised & unsupervised) ou celles qu'elle récupère dans un environnement donné (Reinforcing).

Cfr Linear Algebra For AI [P2]

# Neurone

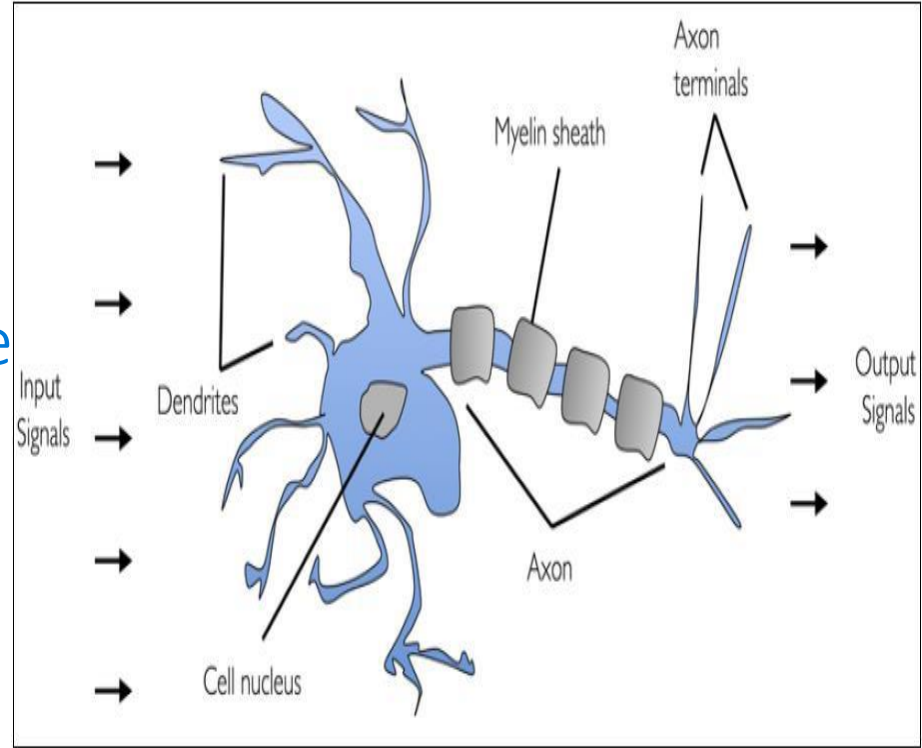
Biologique & Artificiel



# Neurone Biologique



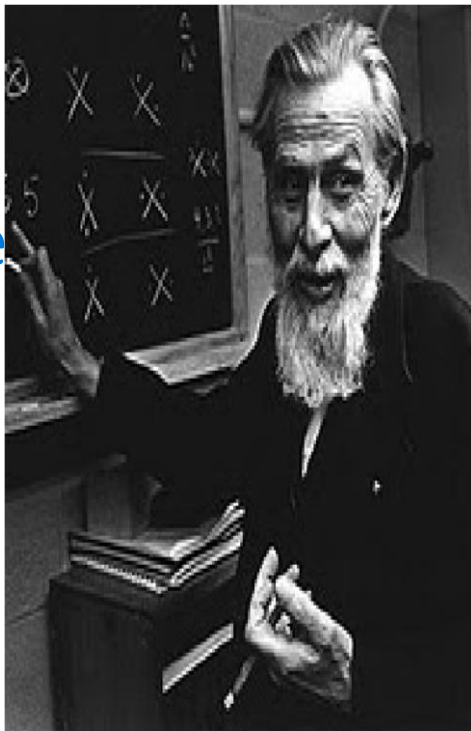
- Le neurone est la cellule qui permet la transmission de l'information et sa mémorisation (synapses):
- Principale unité fonctionnelle du système nerveux
- Le système nerveux humain contient plusieurs dizaines de milliards de neurones



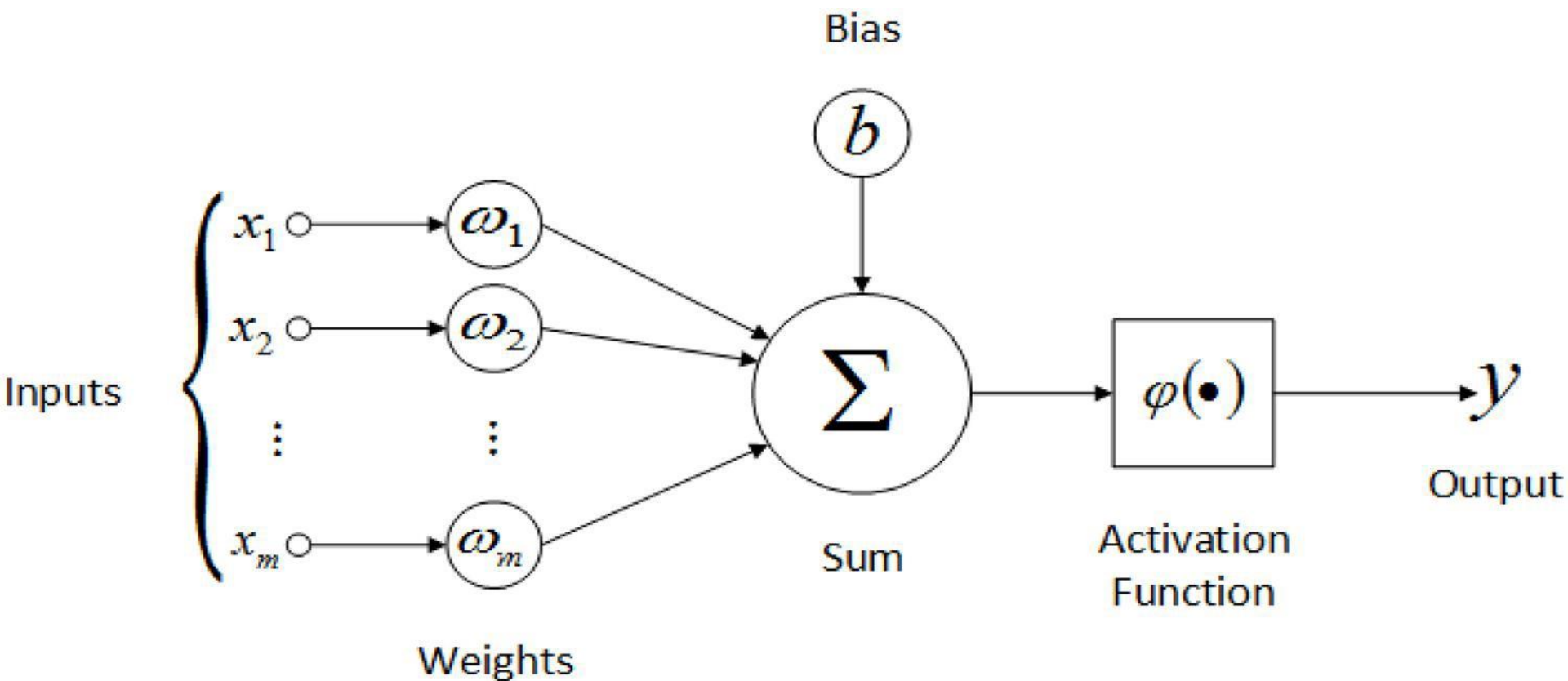
# Neurone Artificiel



- History : 1943
- J. Mc Culloch et W. Pitts
- Modèle binaire du neurone biologique.
- Les auteurs montrent que ces neurons en réseau peuvent réaliser des fonctions logiques ou arithmétiques



# Neurone Artificiel

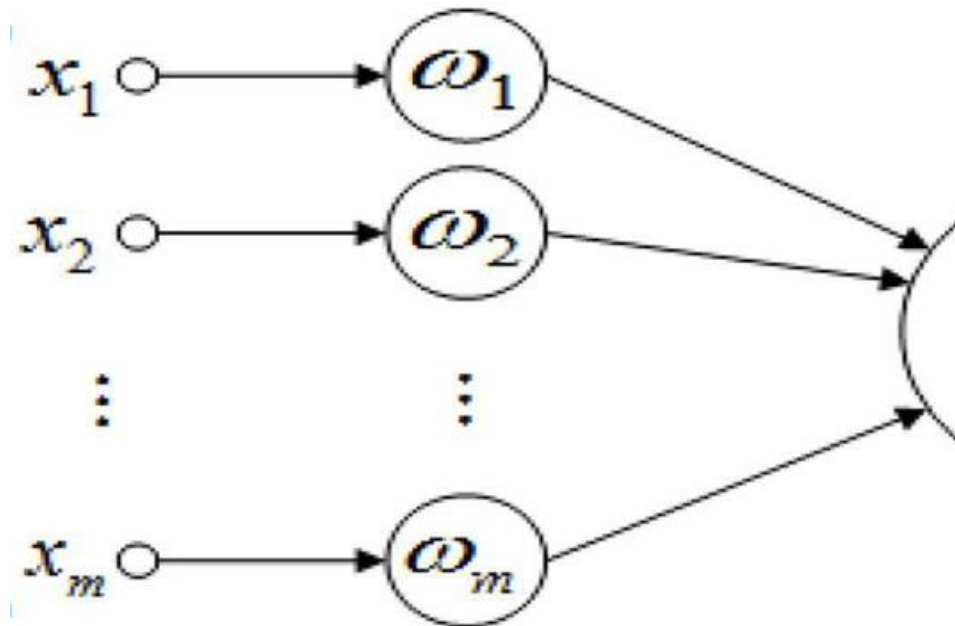




# Neurone Formel



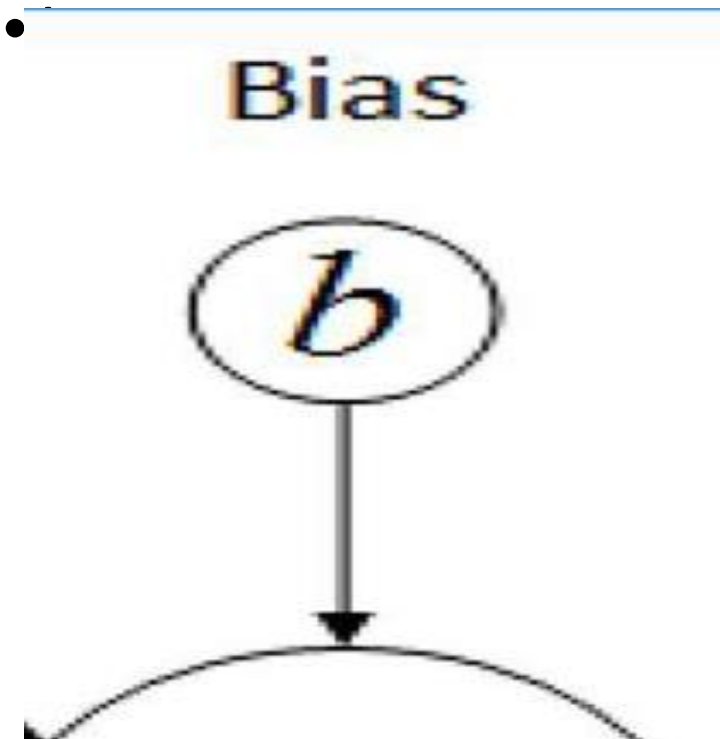
- Weight (poids)
- Ce parameter définit la sensibilité d'une synapse.
- On parle de poids synaptique



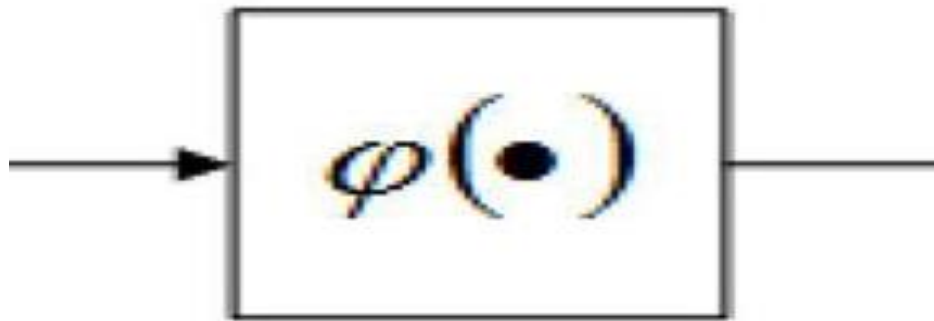
Weights

- Bias (biais)

Ce paramètre influe sur la sensibilité du neurone. C'est un paramètre externe au neurone.

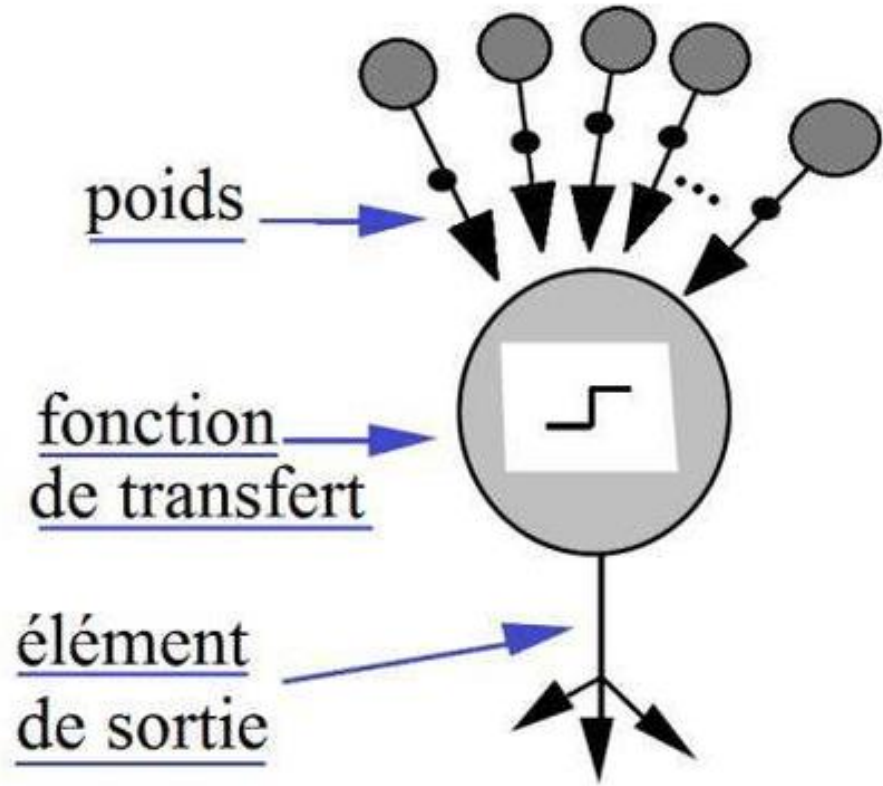
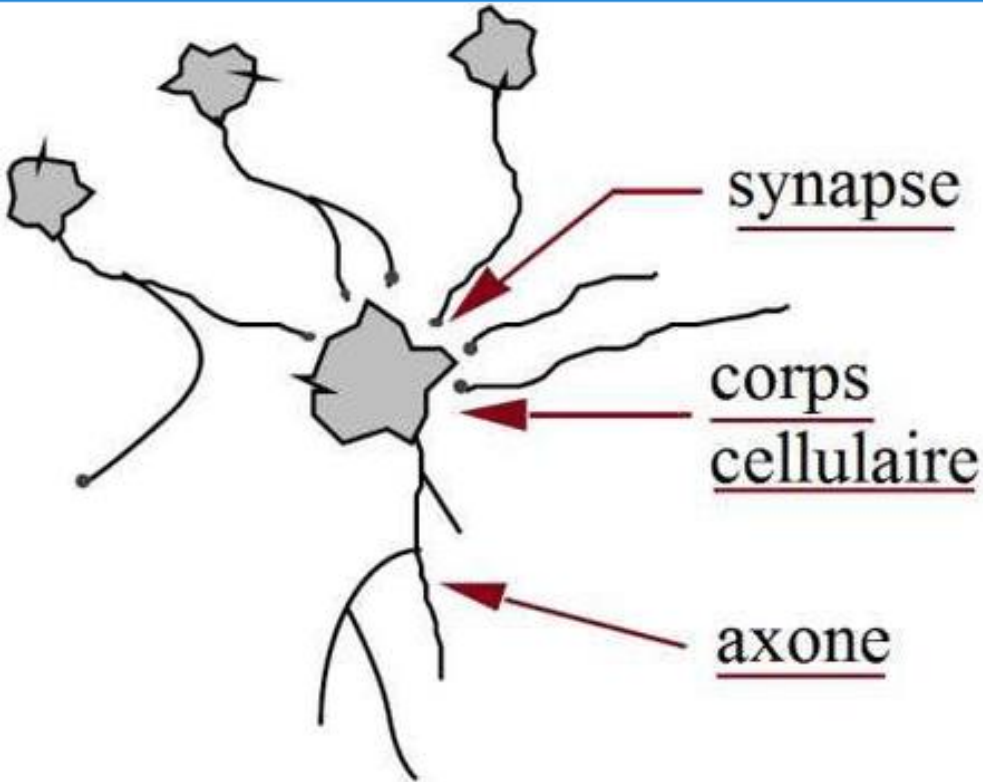


- Fonction d'activation
- c'est elle qui fixe l'état du neurone.
- Dans le cas du neurone formel, cette fonction est binaire. (fonction à seuil)
  - 1: neurone actif
  - 0: neurone inactif



Activation  
Function

# Neurone Biologique et Neurone Artificiel



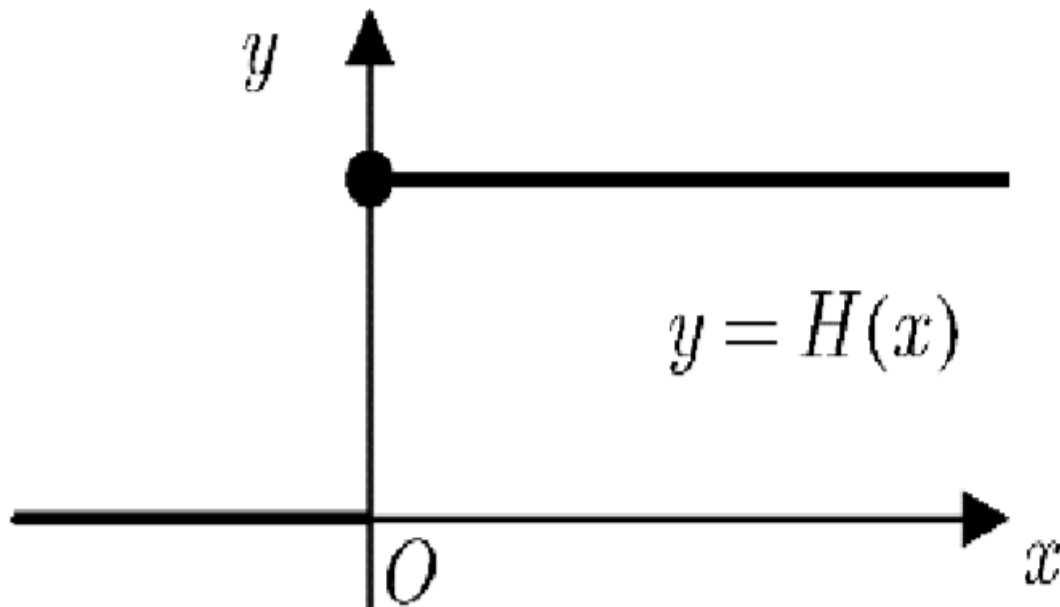
# Généralisation

Fonctions d'Activation

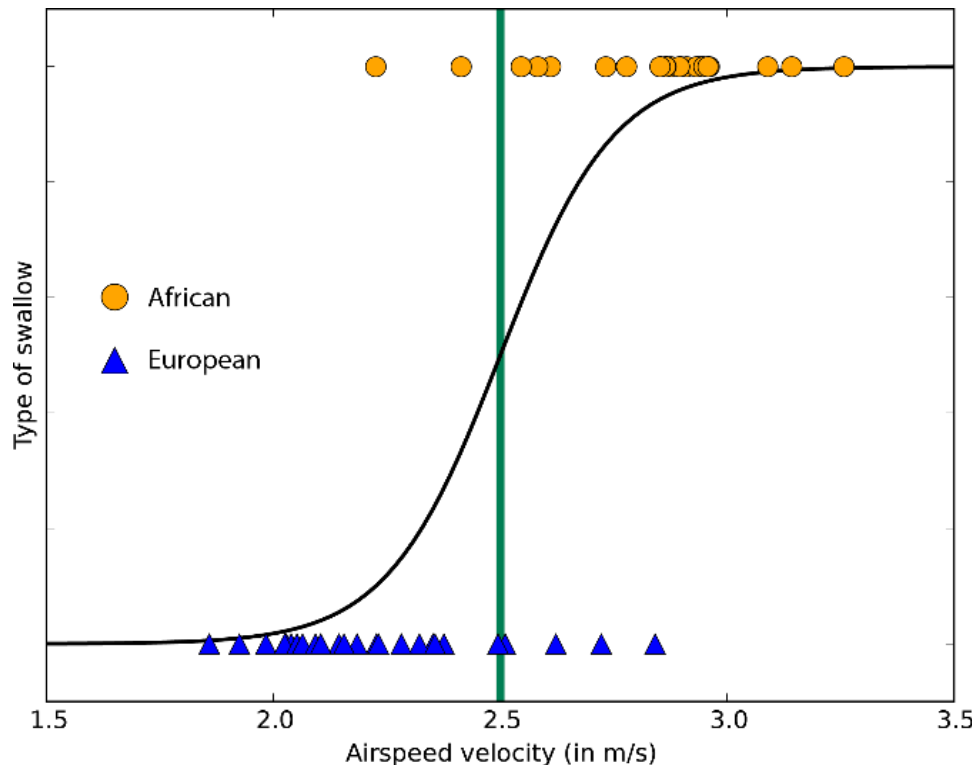


- Le neurone formel est un modèle binaire parce que sa fonction d'activation est une fonction à seuil (binaire).
  - Triste
  - Joyeux
- Pour avoir plusieurs classes on a qu'à changer de fonction d'activation car c'est elle qui impose le nombre de classes.
  - Triste
  - Joyeux
  - Neutre
  - Lover

- Fonction Heaviside
- Fonction à seuil
- 0 pour  $X \leq 0$
- 1 pour  $X > 0$

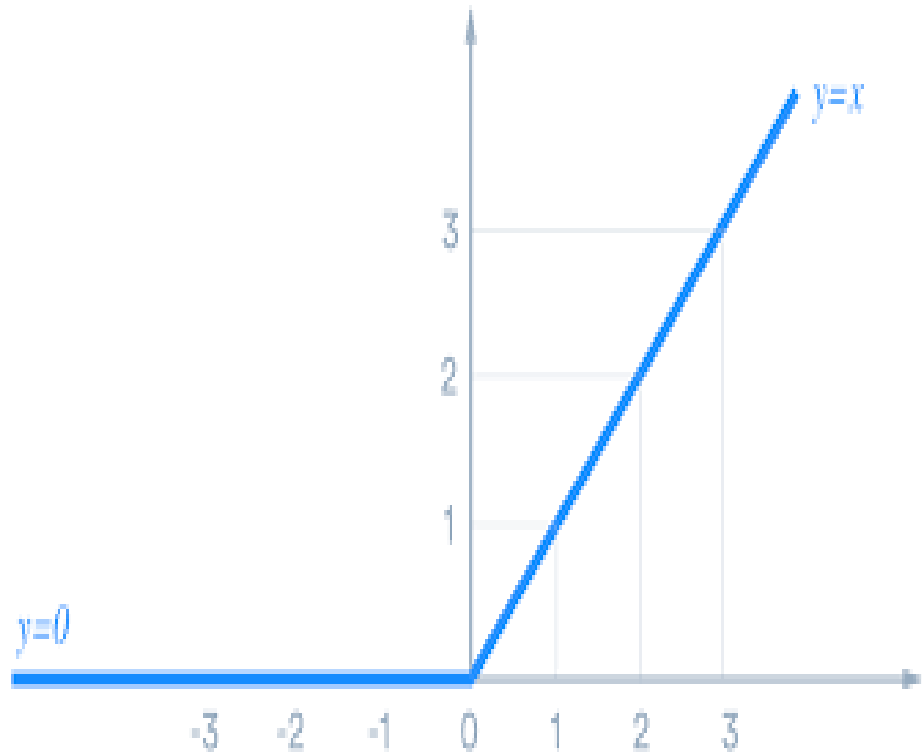


- Fonction SIGMOÏDE
- Non-linéaire
- Derivable à l'infini
- Suffert from vanishing gradient problem





- Fonction RELU(Rectified Linear Unit)
- 0 pour  $x < 0$
- $x$  pour  $x > 0$
- Solve problem from vanishing gradient
- Most widely used activation function



# Hebb Rules



- History : 1949

- D. Hebb

Si 2 cellules sont actives en meme temps alors la force de leur connexin augmente.

- Apprentissage non-supervisé



# Hebb Rules



Before conditioning



→  
response

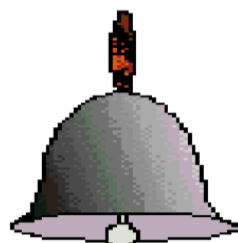


Salivation

**Unconditioned  
stimulus**

**Unconditioned  
response**

Before conditioning



→  
response

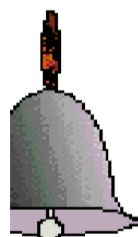


No salivation

**Neutral  
stimulus**

**No conditioned  
response**

During conditioning



+



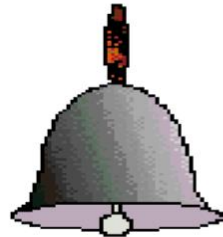
→  
response



Salivation

**Unconditioned  
response**

After conditioning



→  
response



Salivation

**Conditioned  
stimulus**

**Conditioned  
response**

# Application

Neurone Formel





# Le Perceptron

Learning method



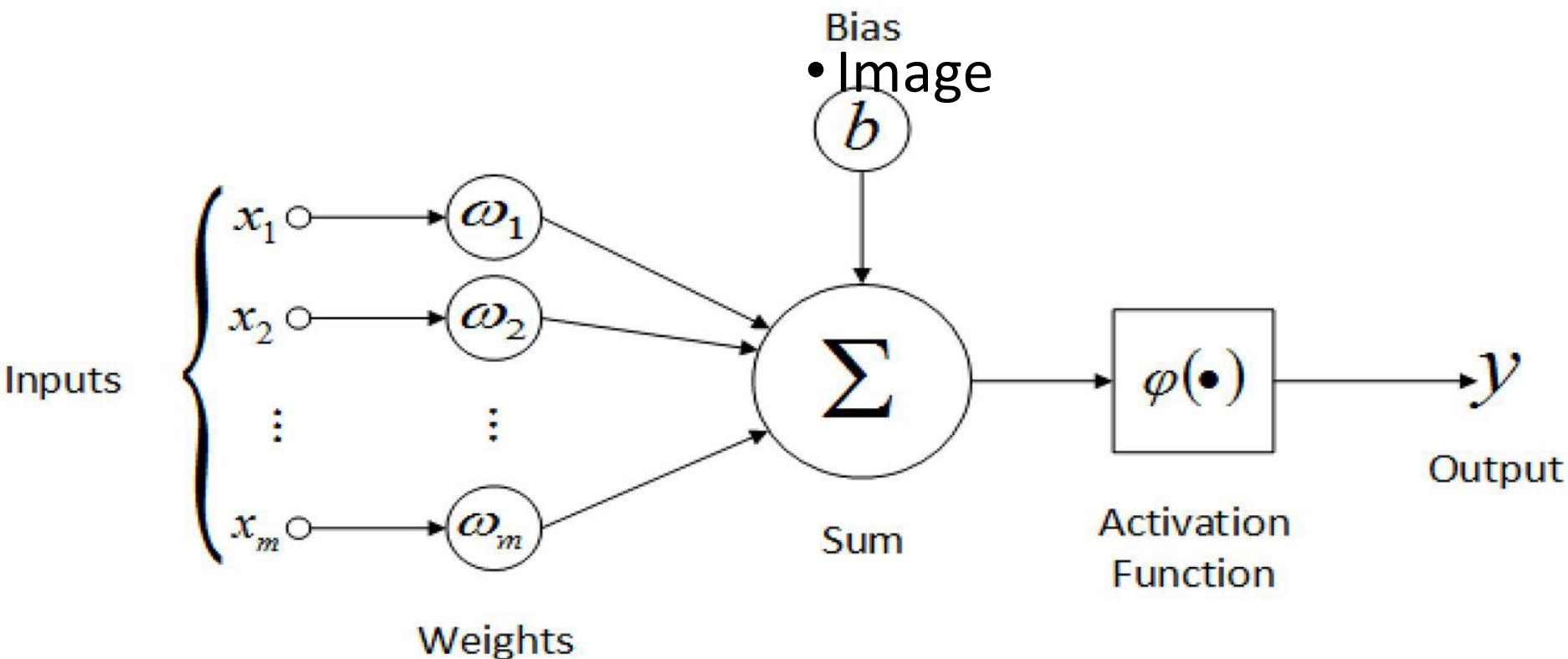
- History: 1958-1960

- F. Rosenblatt

Il invente le modèle du perceptron, basé sur le principe du cortex visuel, et destiné à la reconnaissance de formes. C'est le premier réseau «multi-couche».



# Perceptron

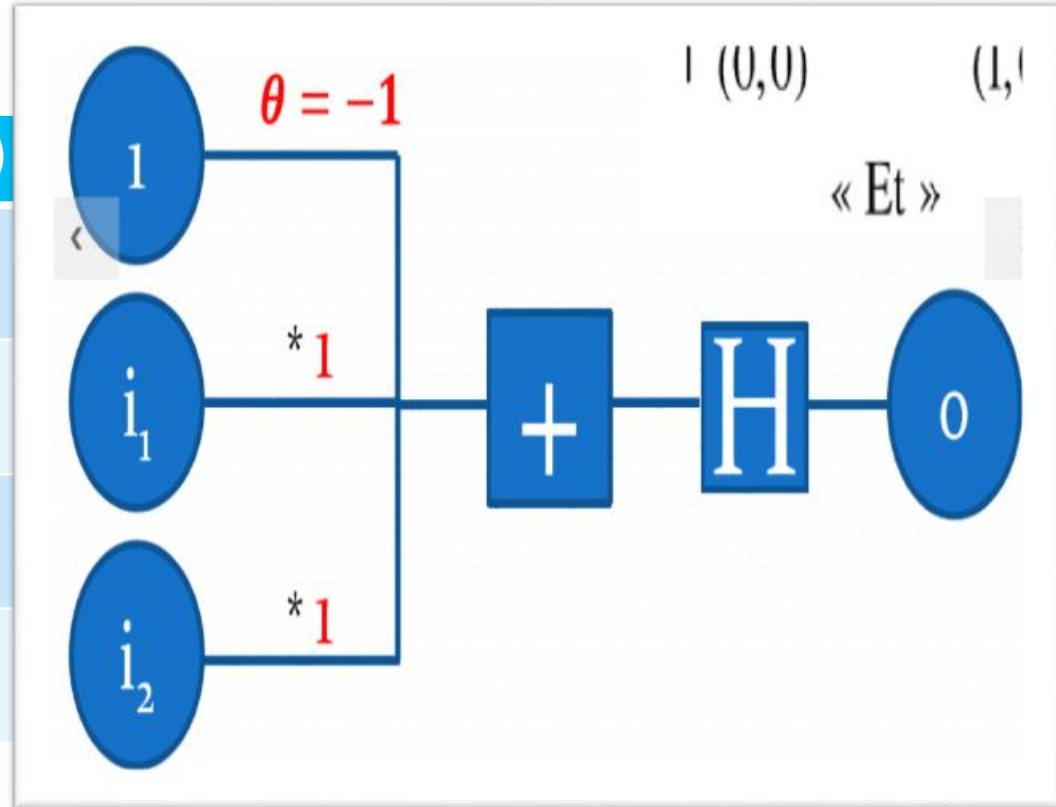


# Perceptron



## • Fonction AND

E1 (input1 )	E2 (input 2)	S (output)
0	0	0
0	1	0
1	0	0
1	1	1



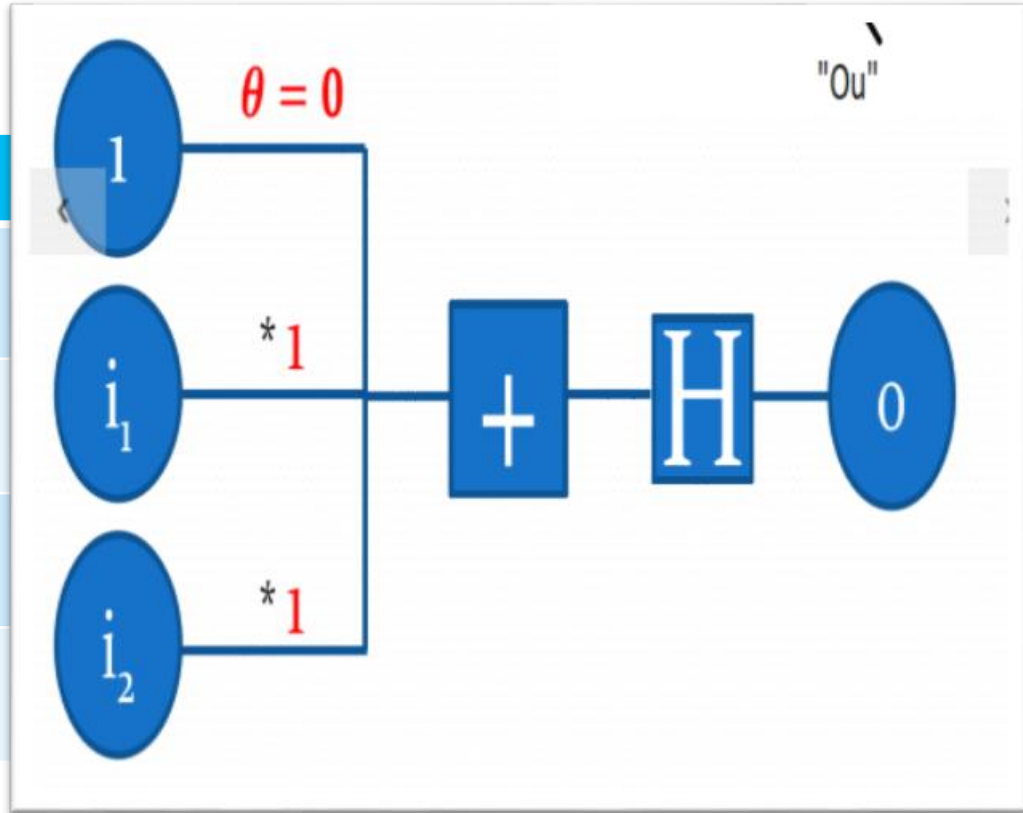


# Perceptron



## • Fonction OR

E1 (input1 )	E2 (input 2)	S (output)
0	0	0
0	1	1
1	0	1
1	1	1



# Learning Protocol

Perceptron

# Perceptron : Learning Protocol



- Initialiser tous les  $W$  (par exemple à la valeur 0)
- Pour chaque exemple d'une base d'apprentissage :
  - Comparer la réponse fournie par le perceptron avec la réponse attendue (vraie parité).

Parité attendue	Sortie	Action
-----------------	--------	--------

0	0	Aucune
---	---	--------

1	1	Aucune
---	---	--------

0	1	Retrancher 1 à tous les poids synaptiques des entrées actives
---	---	---

1	0	Ajouter 1 à tous les poids synaptiques des entrées actives
---	---	--

**On parle de  
rétropropagation de l'erreur**

# Perceptron : Learning Protocol



## • ETAPE 1

INPUT 1	POIDS 1	INPUT 2	POIDS 2	BIAIS	POIDS BIAS	OUTPUT	TARGET	ACTION
0	0	0	0	1	0		0	
0		1					0	
1		0					0	
1		1					1	

- L'algorithme d'apprentissage se poursuit jusqu'à ce que tous les éléments de la base d'apprentissage soient étudiés sans qu'aucun poids synaptiques ne soient modifiés.
  - Stabilisation du perceptron, le modèle a convergé.
- L'algorithme s'arrête si et seulement si la base d'apprentissage est linéairement séparable.

# Perceptron : Learning Protocol



## • ETAPE 2

INPUT 1	POIDS 1	INPUT 2	POIDS 2	BIAIS	POIDS BIAS	OUTPUT	TARGET	ACTION
0	0	0	0	1	0		0	
0		1					0	
1		0					0	
1		1					1	

# Perceptron : Learning Protocol



## • ETAPE 3

INPUT 1	POIDS 1	INPUT 2	POIDS 2	BIAIS	POIDS BIAS	OUTPUT	TARGET	ACTION
0	0	0	0	1	0		0	
0		1					0	
1		0					0	
1		1					1	

# Challenge

Perceptron



- Fonction XOR

E1 (input1 )	E2 (input 2)	S (output)
0	0	0
0	1	1
1	0	1
1	1	0

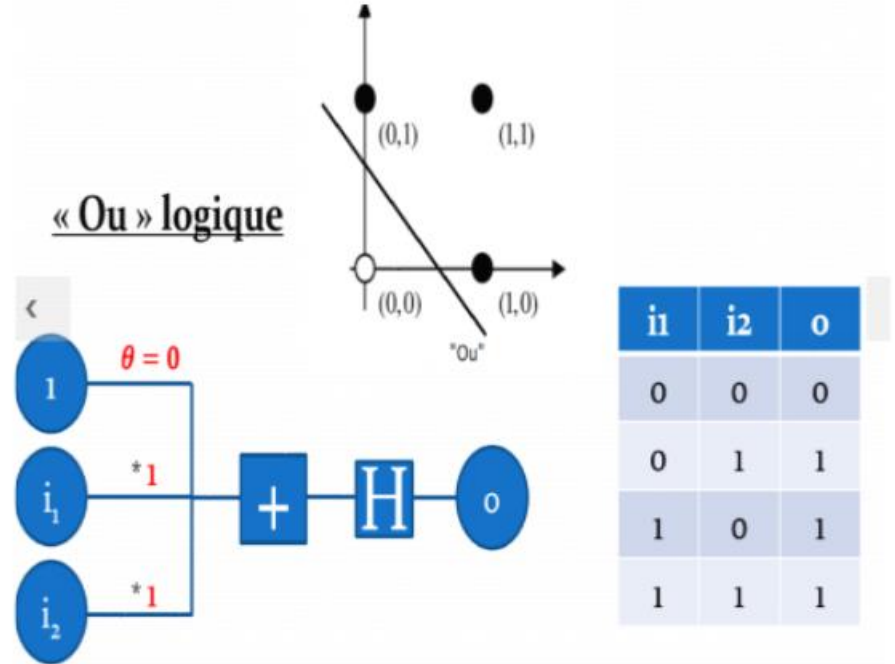
- Ecrire un perceptron qui realize la fonction OU-EXCLUSIF (XOR) don't les données d'apprentissage se trouvent dans le tableau ci-contre.

# Limite

Single Layer Perceptron

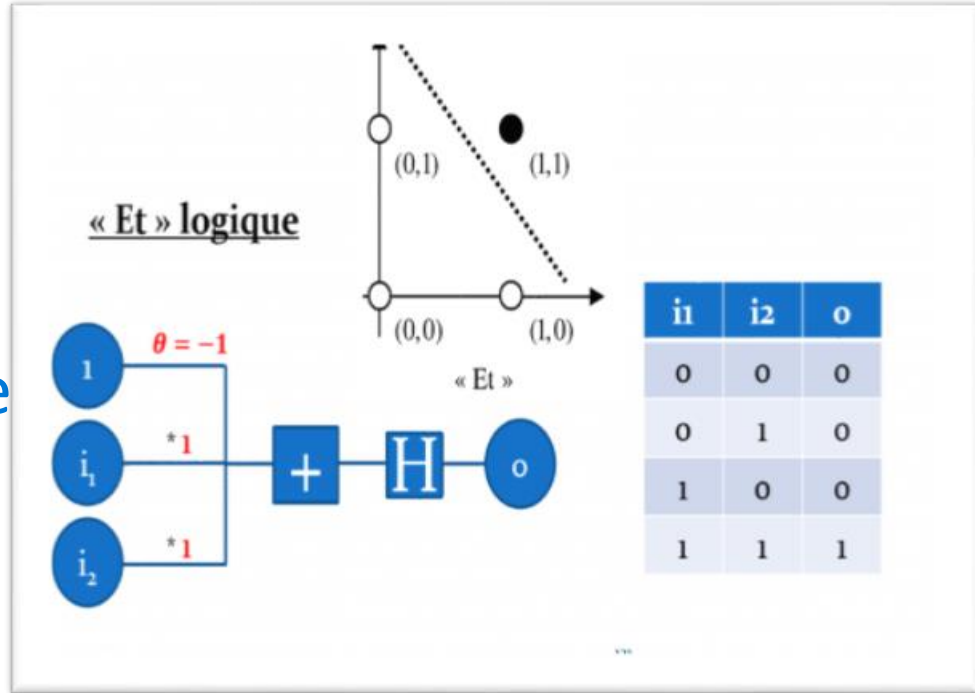
# Perceptron limit

- Un échantillon d'exemples pourra être appris par un perceptron si, et seulement si, il est linéairement séparable.
- C-à-d lorsque les données de l'échantillon se trouvant en N dimensions peuvent être séparées par un hyper-plan en N-1 dimension.



# Perceptron limit

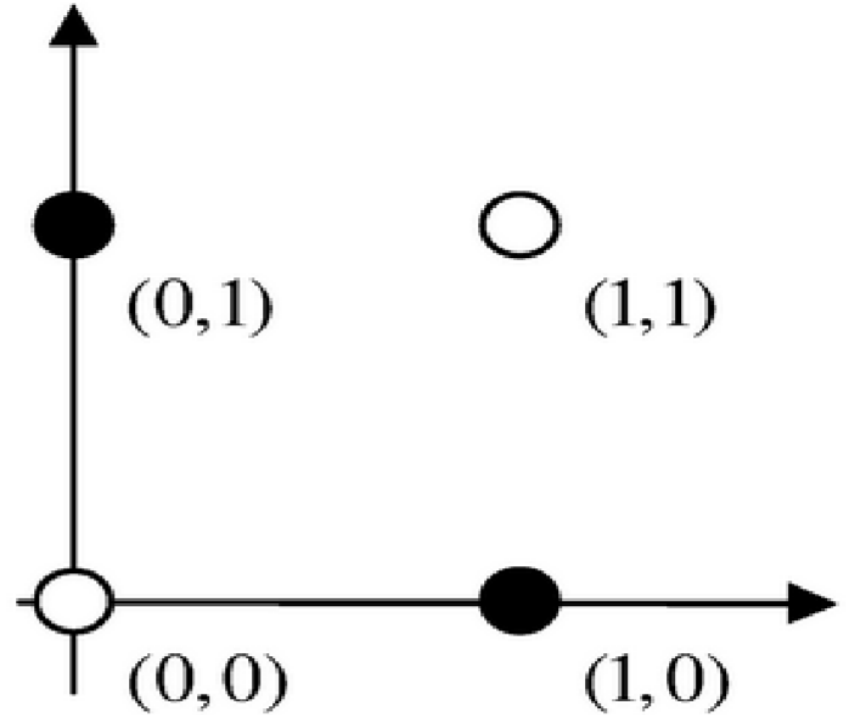
- Un perceptron ne peut modeliser que les données qui sont linéairement separable.
- C-à-d lorsque les données se trouvant en N dimensions peuvent etre séparées par un hyper-plan en N-1 dimension.



# Perceptron limit



- La fonction XOR ne peut être modélisée par le perceptron car les données (points dans un plan) de cette fonction ne sont pas linéairement séparables.



# Neural Network

Multi Layers Perceptron  
(MLP) or (PMC)

# Multi Layers Perceptron



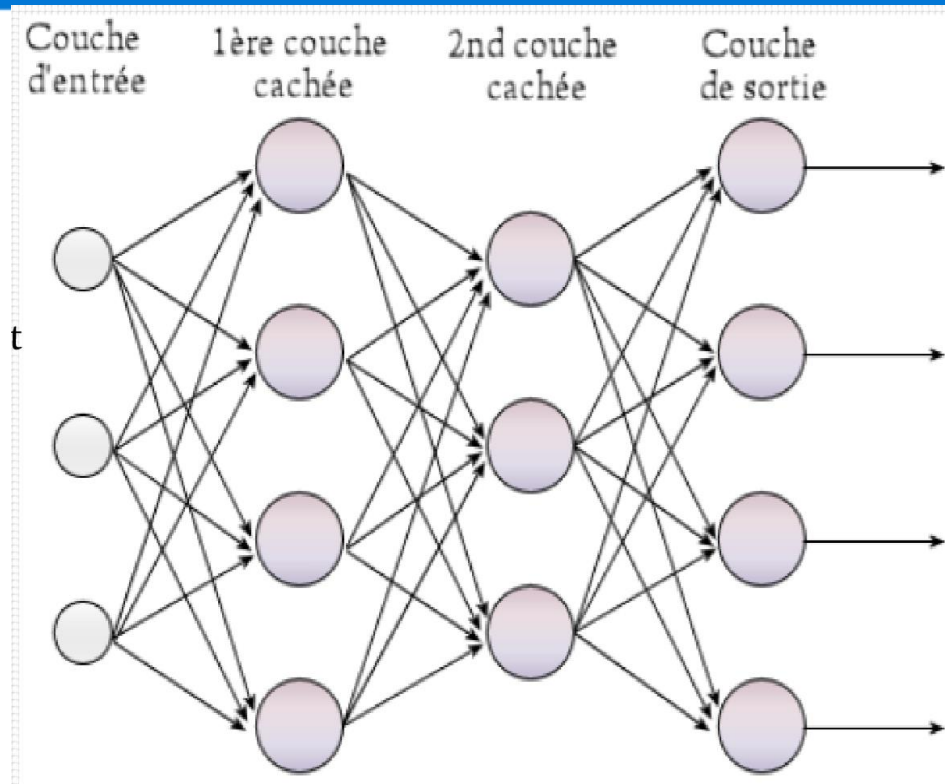
- History : 1986
- David Rumelhart
- Yann Lecun  
(collège de france)



# Multi Layers Perceptron : Principe



- On utilise plusieurs perceptrons en couche successives, les couches centrales étant dites caches.
- Chaque noeud de la couche A est connecté à tous les noeuds de la couche A+1 et l'information circule de la couche inférieure[A] à la couche supérieure[A+1].  
(FeedForward Connection)

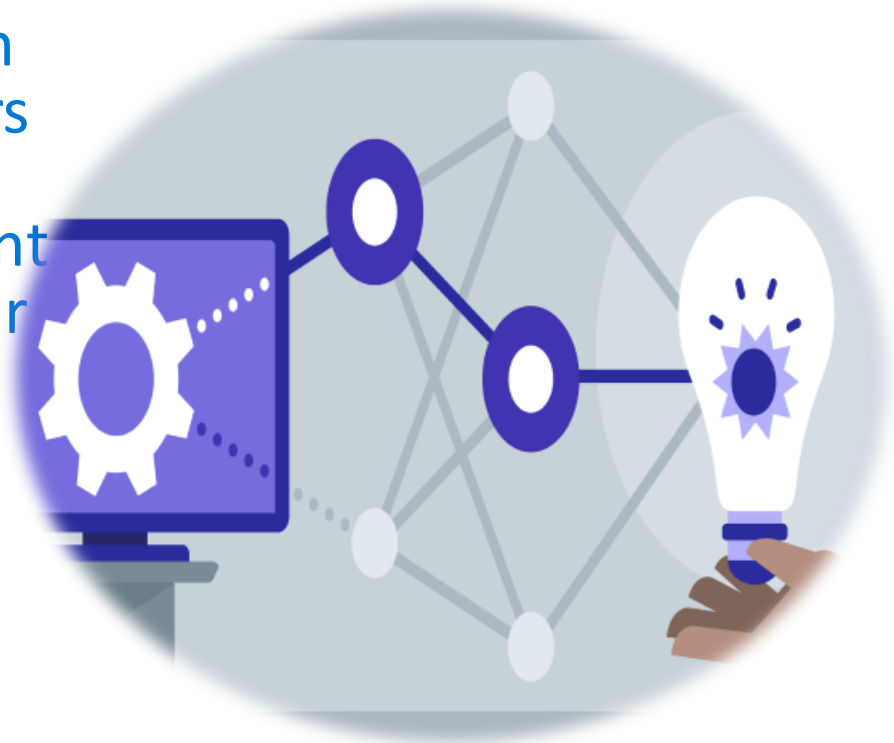




# Neural Network



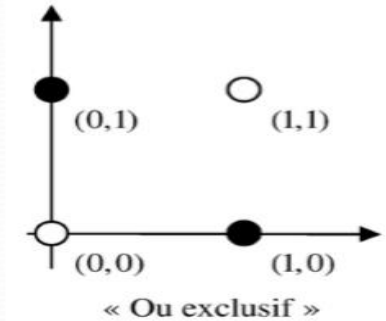
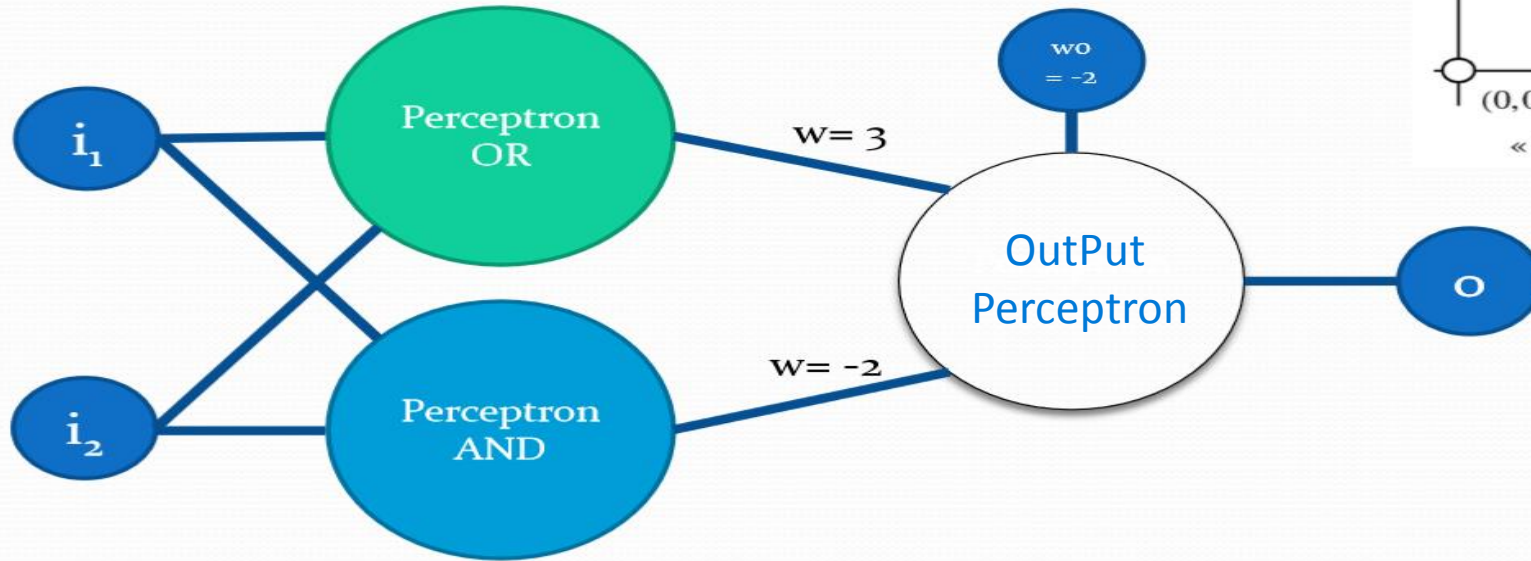
- Un réseau de neurones est un système composé de plusieurs unités de calcul simples fonctionnant en parallèle, dont la fonction est déterminée par la structure du réseau, la solidité des connexions, et l'opération effectuée par les éléments ou nœuds. [DARPA Neural Network Study Group :1988]



# Neural Network : MLP



Pour résoudre notre problème de linéarité :



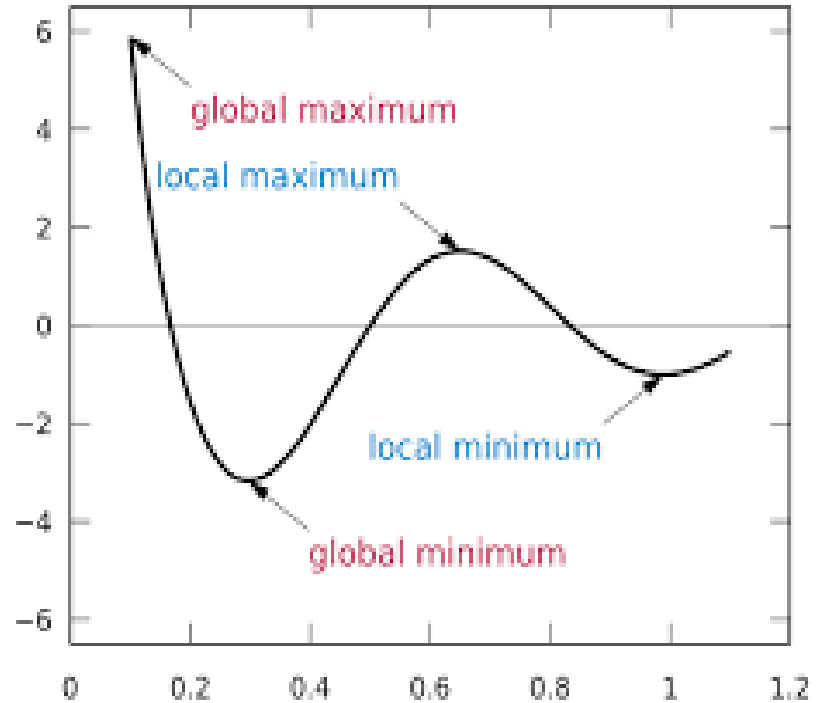
# Retropropagation de Gradient

Back-Propagation

# Back-Propagation methods



En statistiques, la rétropropagation du gradient est une méthode pour calculer le gradient de l'erreur pour chaque neurone d'un réseau de neurones, de la dernière couche vers la première. [Wikipédia](#)



**BATCH**: mise à jour des poids après la présentation de tous les exemples

- calculs et stockage plus lourds si trop d'exemples

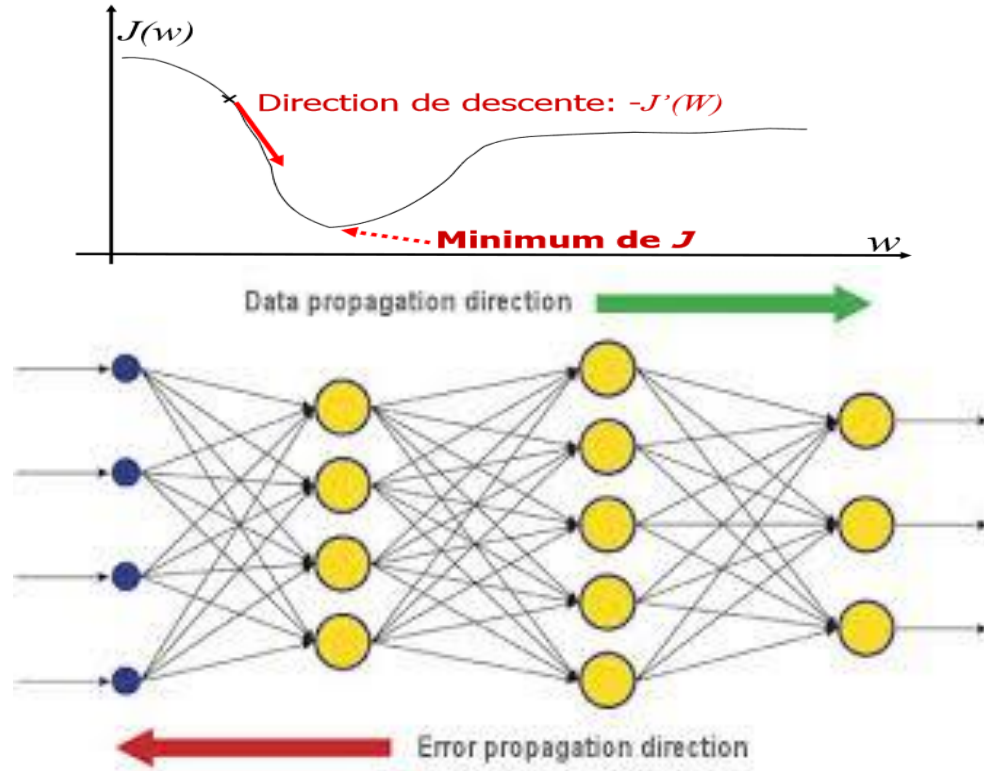
• **SÉQUENTIEL** : mise à jour des poids après chaque exemple.

- On-line method
- Gradient stochastique)
- besoin de tirer l'exemple au hasard
- problèmes de convergence

# Back-Propagation Advantage



- S'adapte à toutes les architectures neuronales
- Peut s'utiliser en parallèle avec d'autres algorithmes de Machine Learning.



# Challenge

Perceptron

- Suite de peniel

question	Response
11	1
13	3
41	4
55	

C'est quoi le mot de passe??

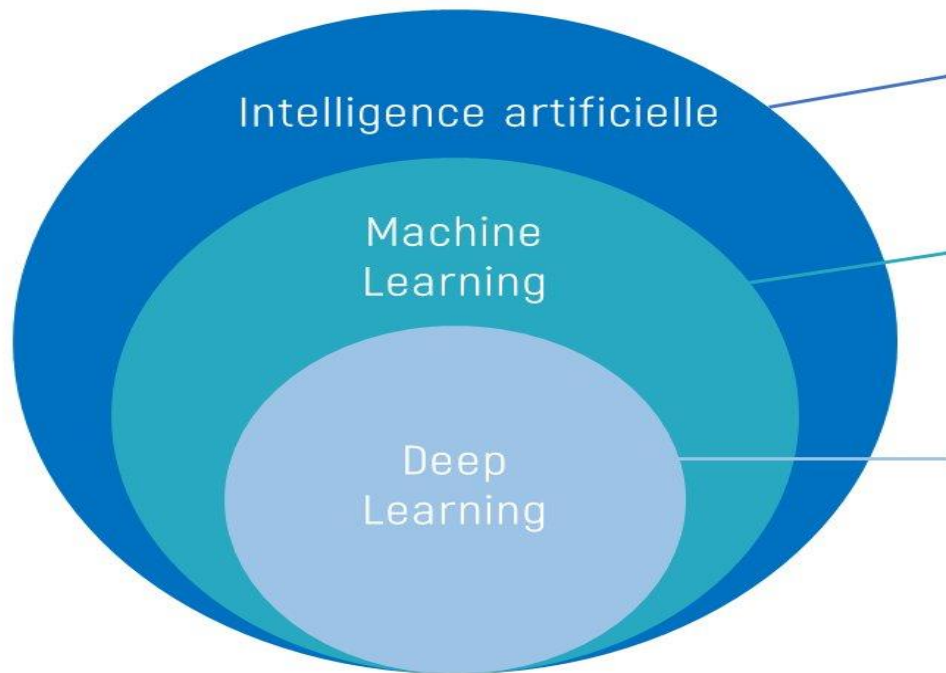
Response = 25

Response = 5 + 5



# Deep Learning

# Deep Learning



**L'INTELLIGENCE ARTIFICIELLE**, c'est l'étude de dispositifs qui analysent leur environnement et définissent des schémas d'actions pour maximiser leurs chances pour atteindre un but donné.

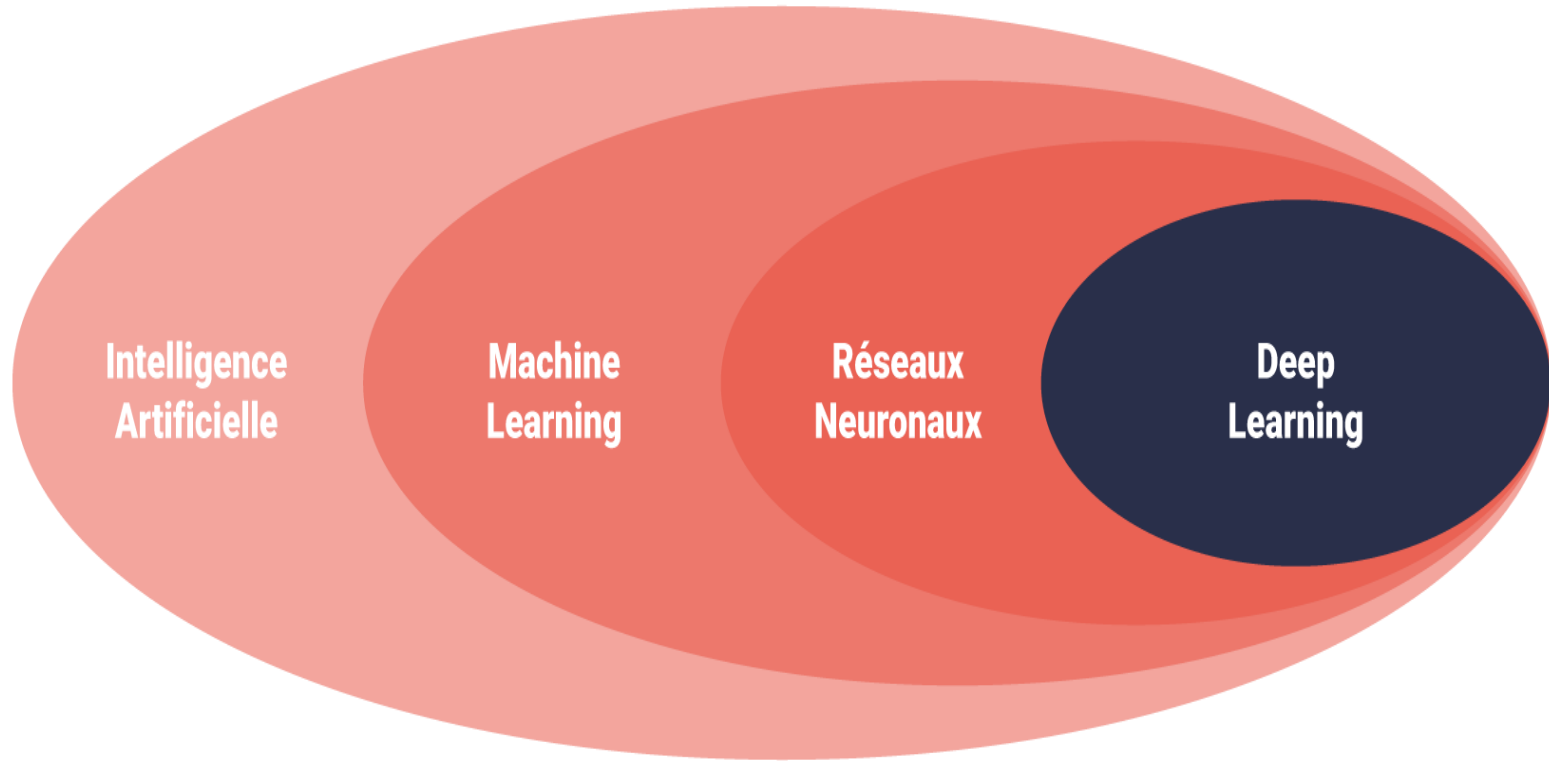
**LE MACHINE LEARNING** est une forme particulière d'intelligence artificielle, où les machines apprennent à exécuter des tâches pour lesquelles elles n'ont pas été explicitement programmées.

**LE DEEP LEARNING** est une forme particulière de machine learning basée sur une répartition des tâches entre différents algorithmes de machines learning, organisés en différentes couches, chaque couche se base sur la couche précédente. Ensemble, ces couches constituent un réseau de neurones artificiel qui va imiter la manière dont le cerveau humain résout un problème.

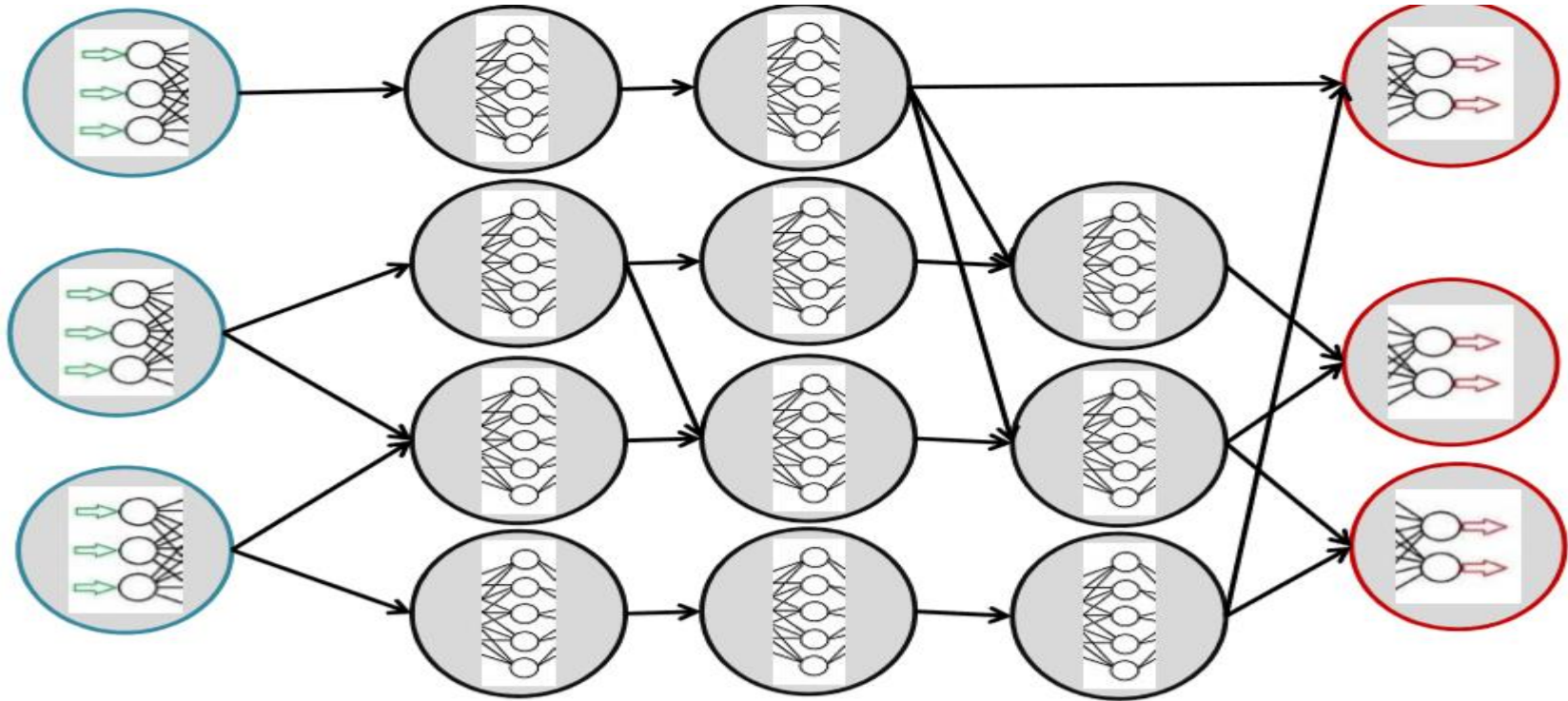
Source : [www.webfoundation.org](http://www.webfoundation.org)

deODIS

# Deep Learning



# Deep Learning



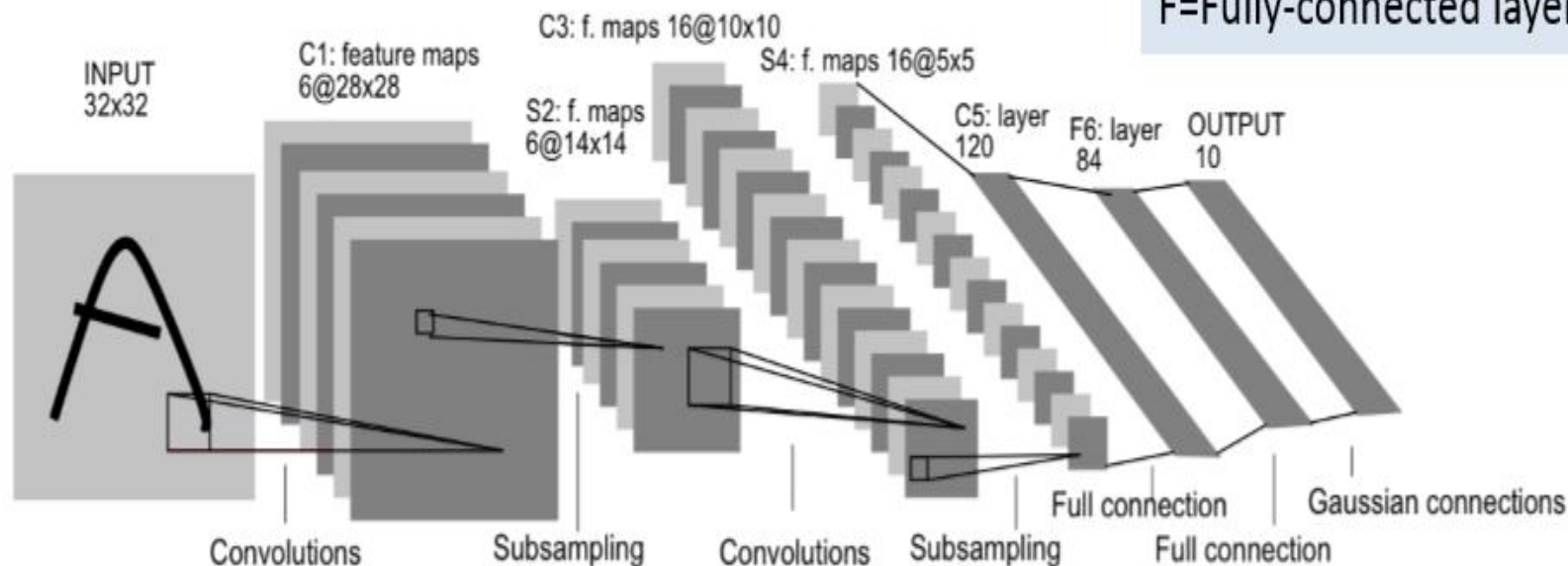
# NN Architecture

Multi Layers Perceptron  
(MLP) or (PMC)

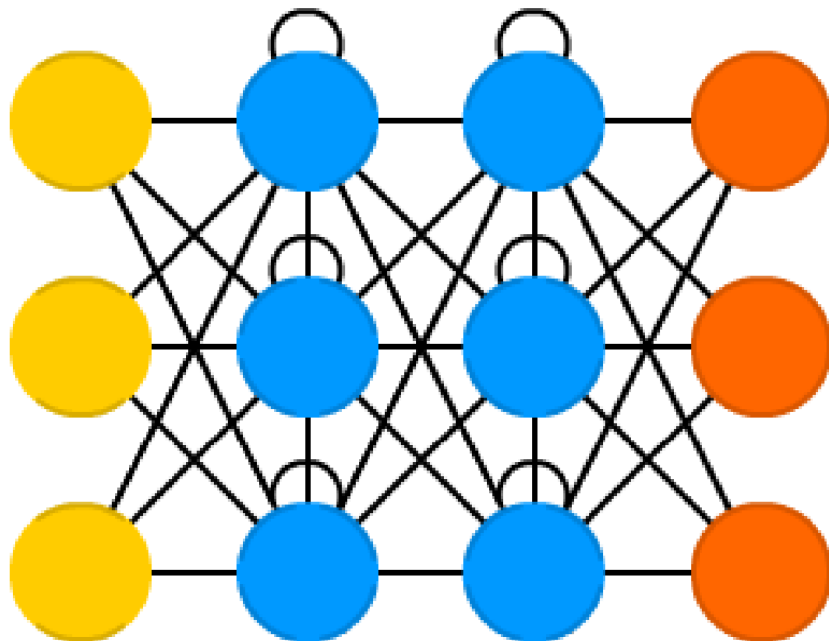
# Architecture : Convolutional NN



PROC. OF THE IEEE, NOVEMBER 1998

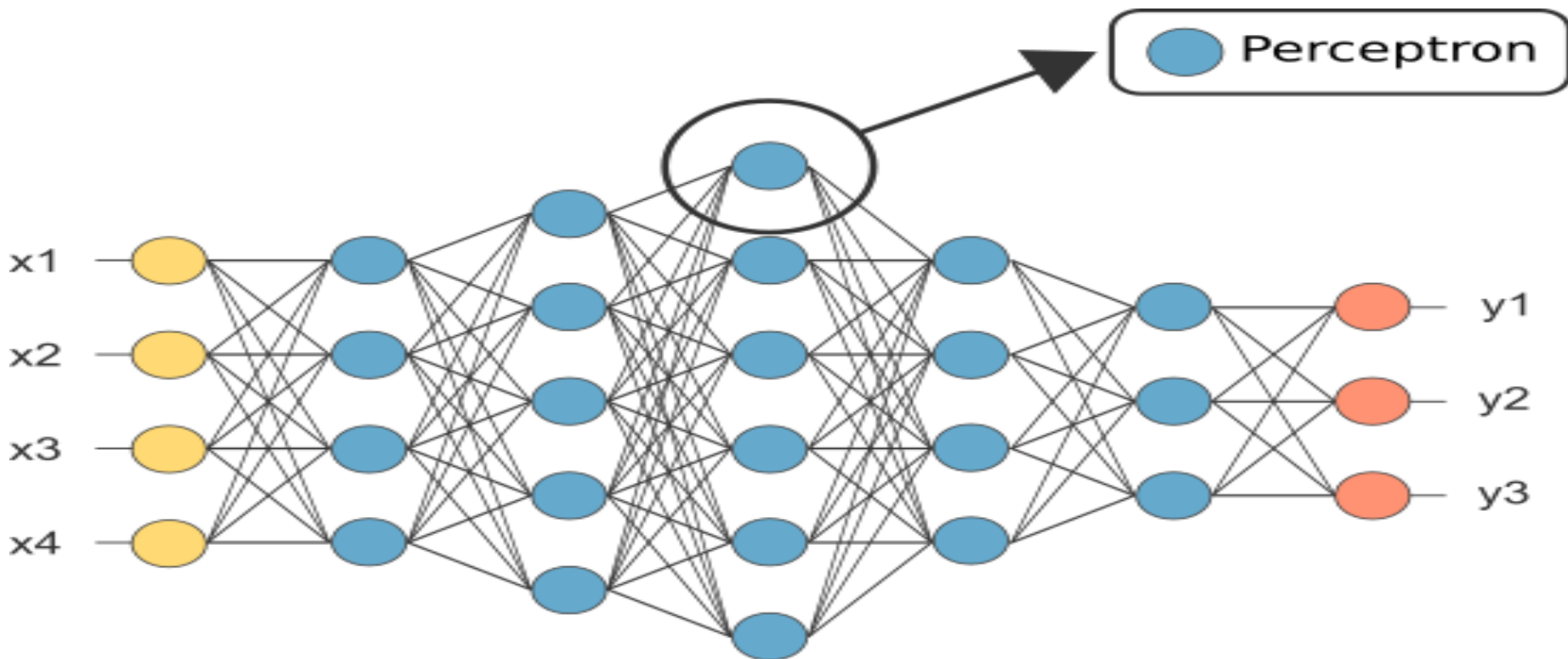


## Recurrent Neural Network (RNN)



- Backfed Input Cell
- Input Cell
- △ Noisy Input Cell
- Hidden Cell
- Probablistic Hidden Cell
- △ Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- △ Different Memory Cell
- Kernel
- Convolution or Pool

# Architecture : Deep NN

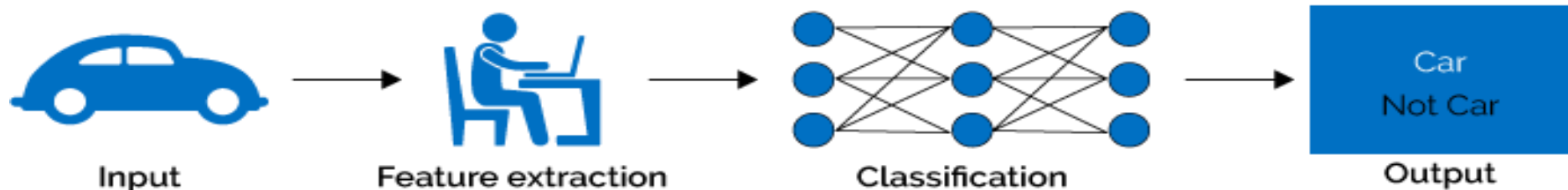




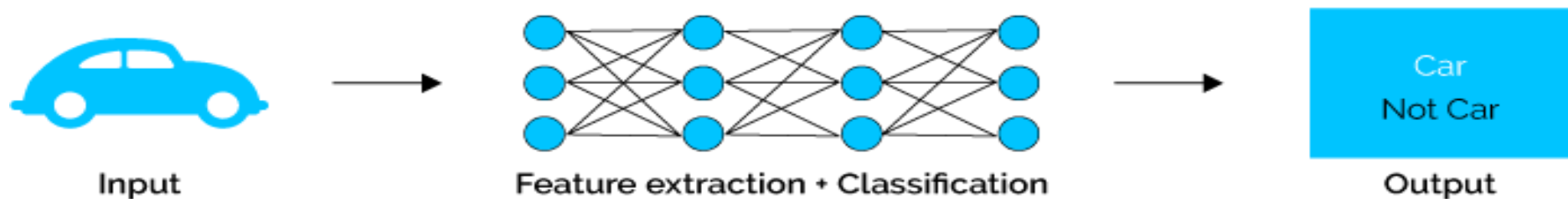
# Deep Learning : Advantage



## Machine Learning



## Deep Learning



# Convolutional Neural Network

# Convolutional Neural Network



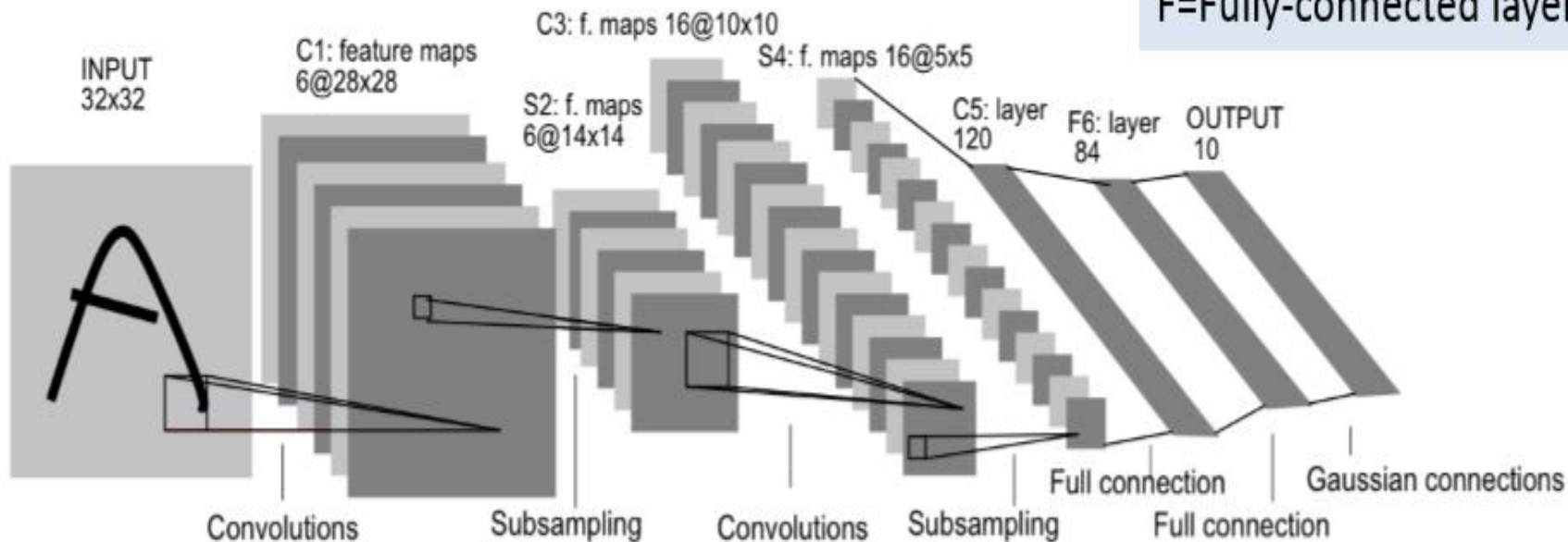
- History



# Convolutional Neural Network



PROC. OF THE IEEE, NOVEMBER 1998



# Convolutional Neural Network

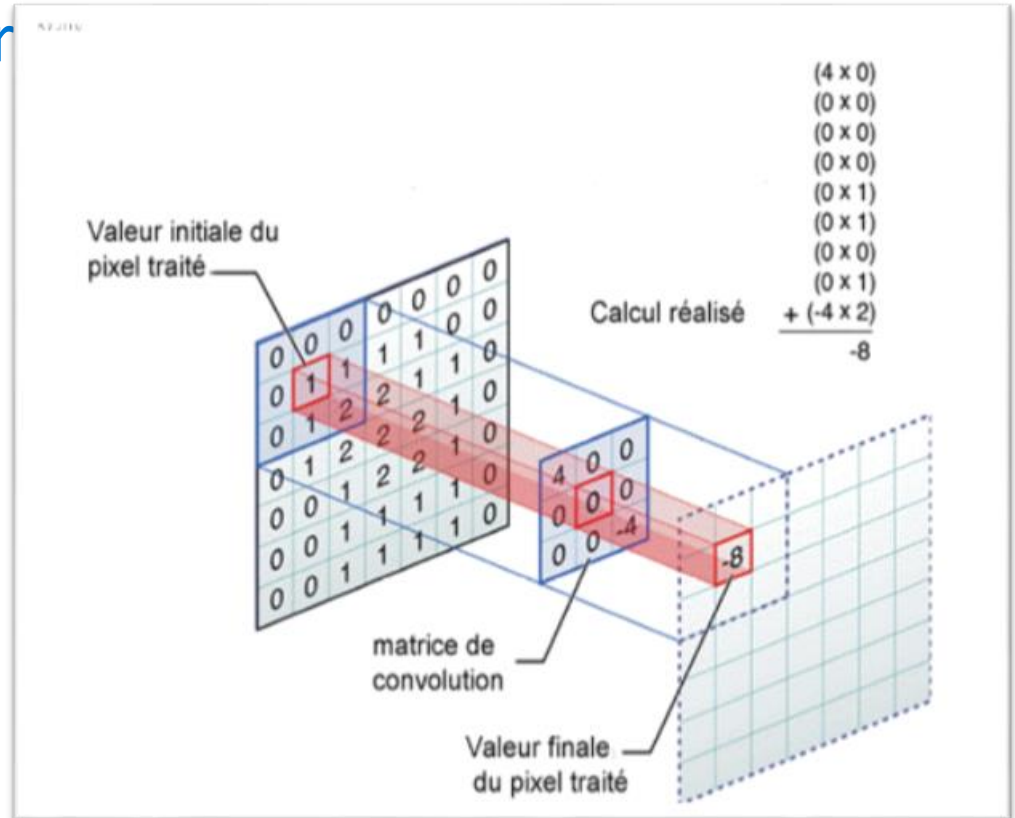


- Couche de convolution
- Couche de Sous-échantillonnage (pooling)
- Couche de transformation
- Full connected layer

# CNN : Convolution

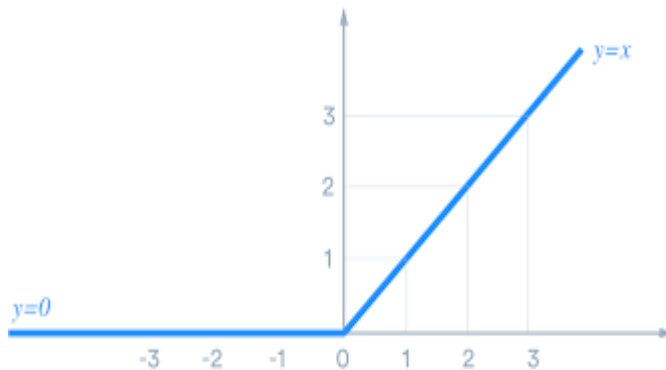


- Cette couche va appliquer plusieurs filtres de convolution.



- Cette couche va corriger les erreurs.
- La **fonction ReLU** est interprétée par la formule:  $f(x) = \max(0, x)$ . Si l'entrée est négative la sortie est 0 et si elle est négative alors la sortie est x.

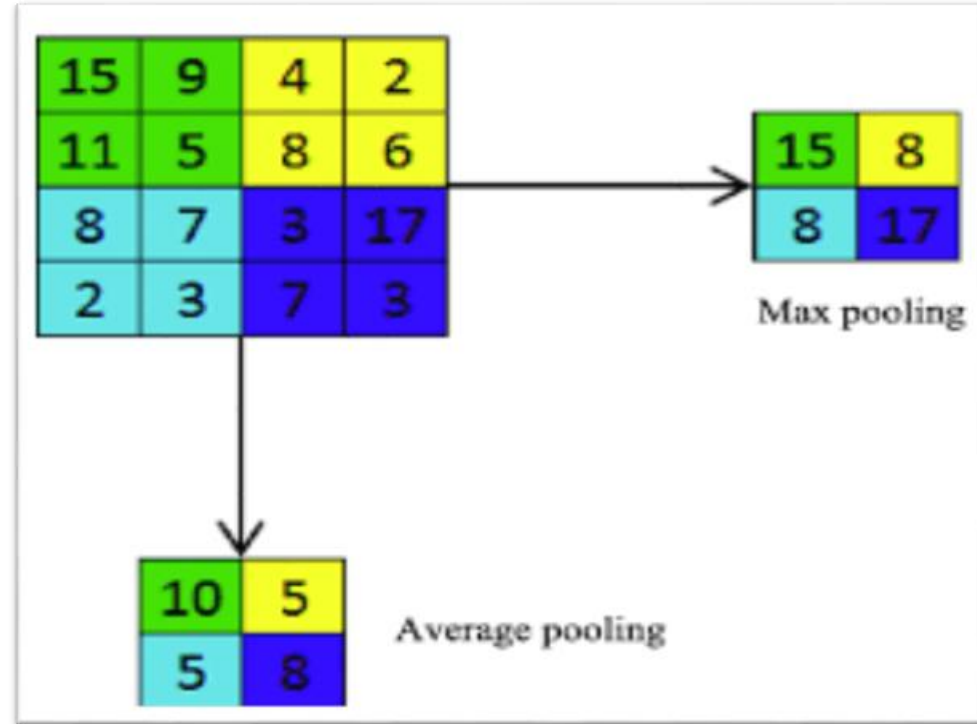
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



# CNN : Sub-Sampling



- Cette couche va échantillonner (réduire la taille de) l'image.
- On parle de pooling
  - Max
  - Average (moyenne)

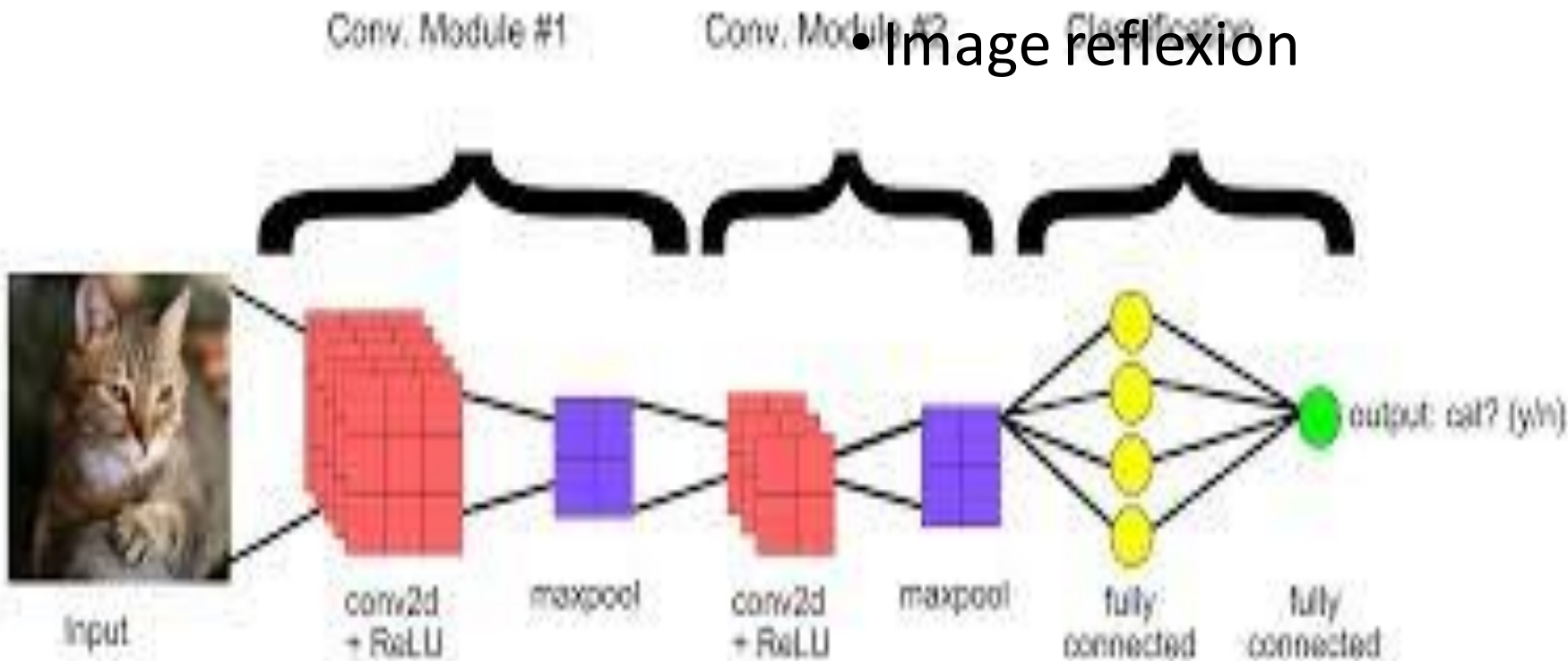




# CNN : Full Connection



- Image reflexion



## Avantages

- Moins de paramètres libres
- Complexité qui n'explose pas avec la taille de l'image
- Architecture qui facilite l'introduction de connaissances
- La rétro-propagation du gradient(Back-Prop) s'adapte facilement, et peut se superposer à d'autres algorithmes
- Possibilité de traitements en parallèle.

# Choice

Lequel prendre???

Deux types de réseaux à considérer

- Convolutional Neural Network (CNN)

Generalisation du PMC avec des couches NON NEURONALES

- Deep Neural Networks (DNN)

PMC avec un grand nombre de couches

- Recurent Neural Network (RNN)

PMC bouclé

## Deux types d'entrées à considérer

1. Entrées sans topologie (une liste de variables explicatives)

Deep Neural Network (DNN)

Regression/Classification linéaire pour des jeux de données avec un grand nombre de variables.

Domotique, marketing, actuariat

Deux types d'entrées à considérer

## 2. Entrées avec topologie (une image, un signal)

- Convolutional Neural Network & Recurrent Neural Network
- Détection/Classification de patterns dans des images et des signaux.
- Object recognition, Face detection, Face recognition -> computer vision
- Speech recognition, Voice recognition

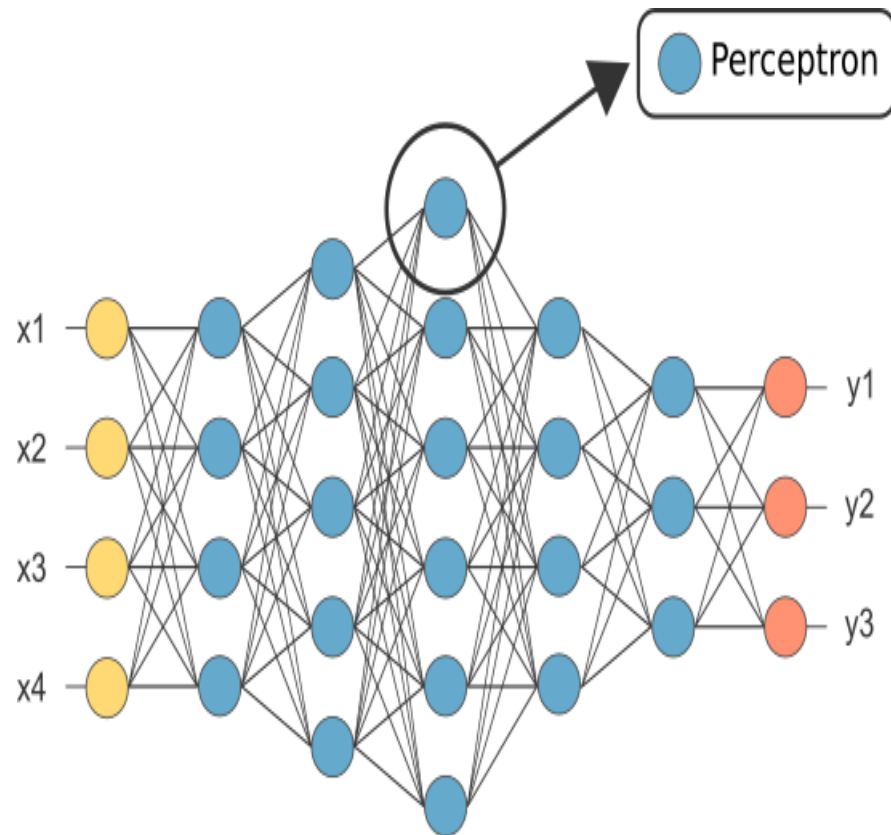
# Question ouverte

PMC BLACK-BOX

# Question ouverte



- Un Perceptron multi-couches (MLP) est une boîte noire, Un Deep Neural Network encore plus!
- Alors, comment expliquer la décision du réseau??





# AlphaGo Zero

Starting from scratch



# Microsoft Student Partners