

Plan du cours

1) Introduction au machine learning

2) Régularisation et forêts aléatoires

3) Réseau de neurones

4) Réseau de neurones convolutifs

Pour chaque séance:

1h de cours / support transparent

2h de travaux pratiques (amener un ordinateur portable)

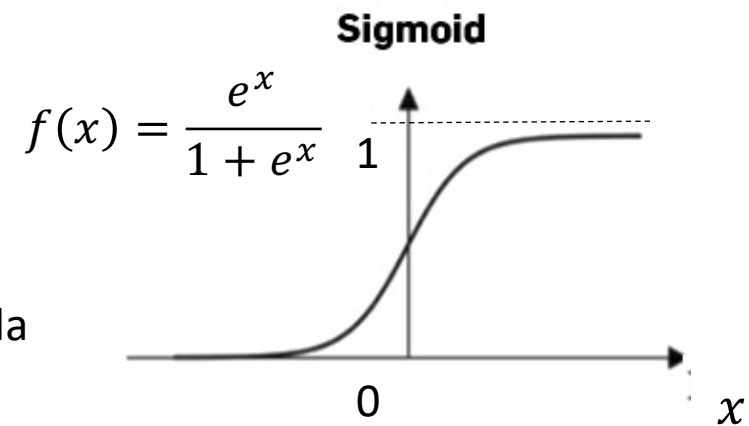
Classification

- Pour l'étude de variable qualitative en sortie y .
Exemples : on cherche à classer les images en 'Chien', 'Chat' ou 'canard'
on cherche à identifier la convection atmosphérique (avec 1 ou 0)
- L'étude de variable qualitative binaire peut se faire à l'aide d'une régression logistique.

$$f(\theta_0 + \theta_1 x_{1,i} + \dots + \theta_m x_{m,i}) = \hat{y}_i$$

$$\text{avec } f(x) = \frac{e^x}{1+e^x}$$

\hat{y}_i est interprété comme la probabilité p_i d'obtenir la variable qualitative correspondant à la valeur 1.



- La fonction loss est alors la negative cross-entropy :

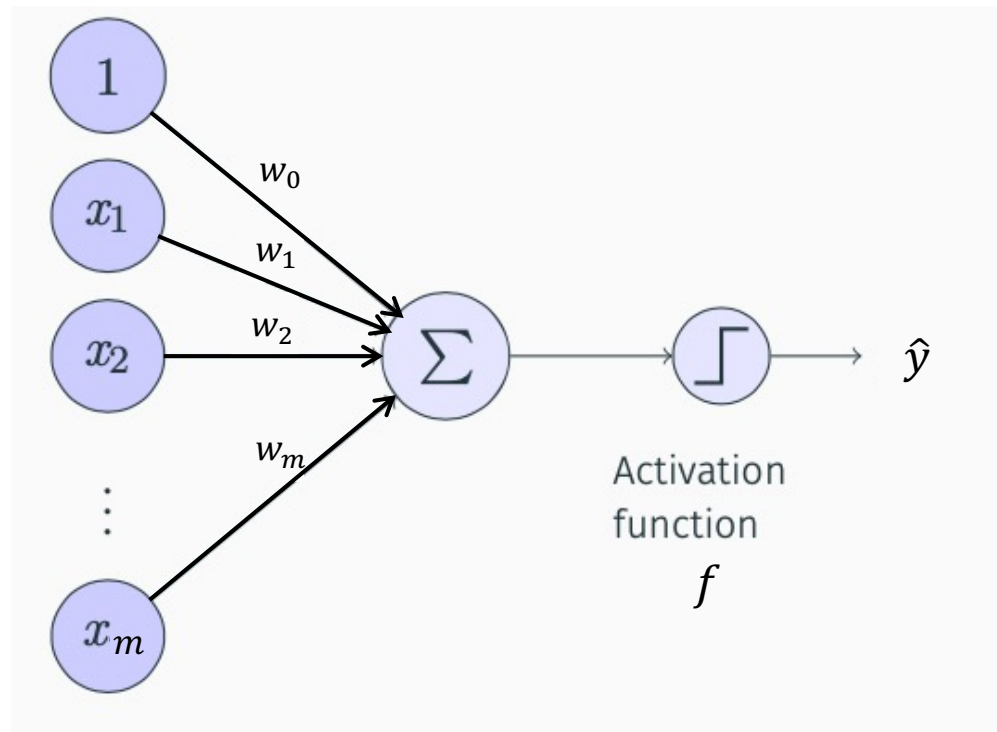
$$J(\boldsymbol{\theta}) = -(y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i))$$

Le perceptron

Biais

Fct d'activation

$$\hat{y}_i = f(\underbrace{w_0}_{\text{Biais}} + w_1x_{1,i} + w_2x_{2,i} + \dots + w_mx_{m,i}) = \underbrace{f}_{\text{Fct d'activation}}\left(w_0 + \sum_{j=1}^m w_jx_{j,i}\right)$$



Poids / Weights
 w_j

= paramètres du perceptron

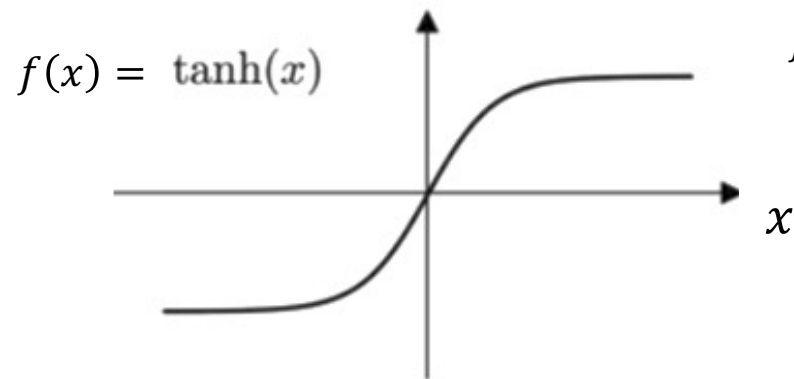
$$f(x) = ?$$

Entrées / Input

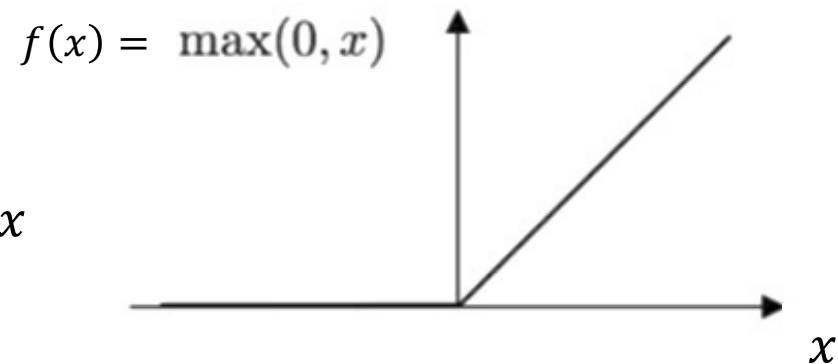
Sortie / Output

Quelques fonctions d'activation

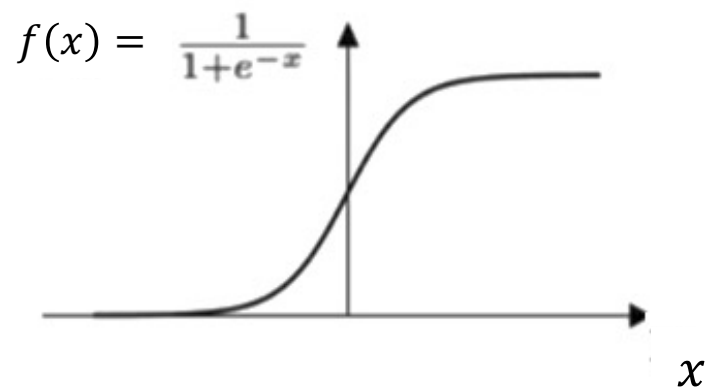
Tanh



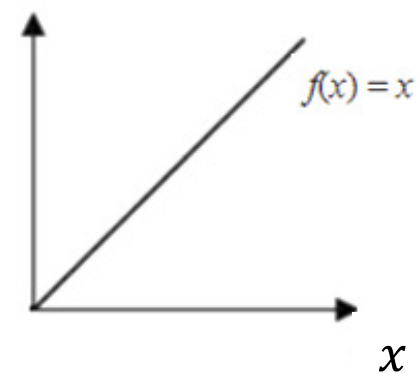
ReLU



Sigmoid



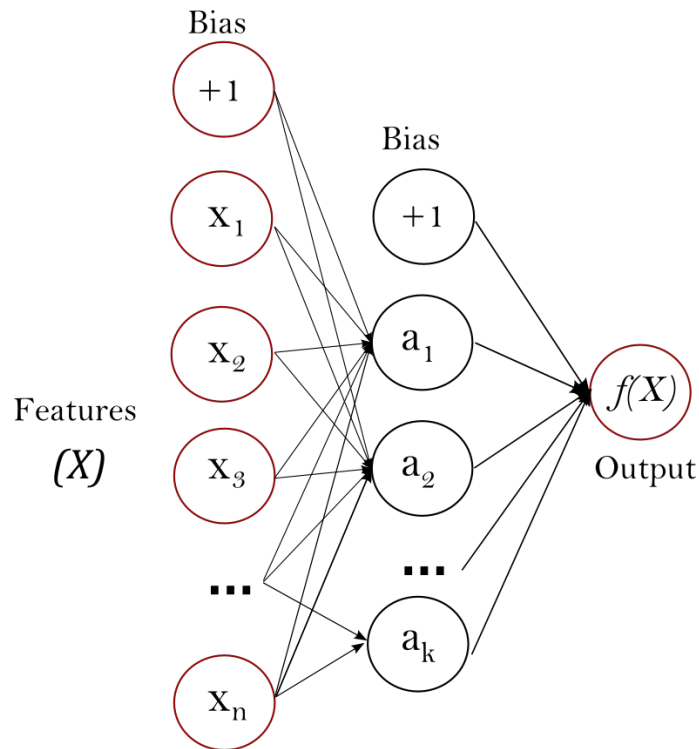
Linear



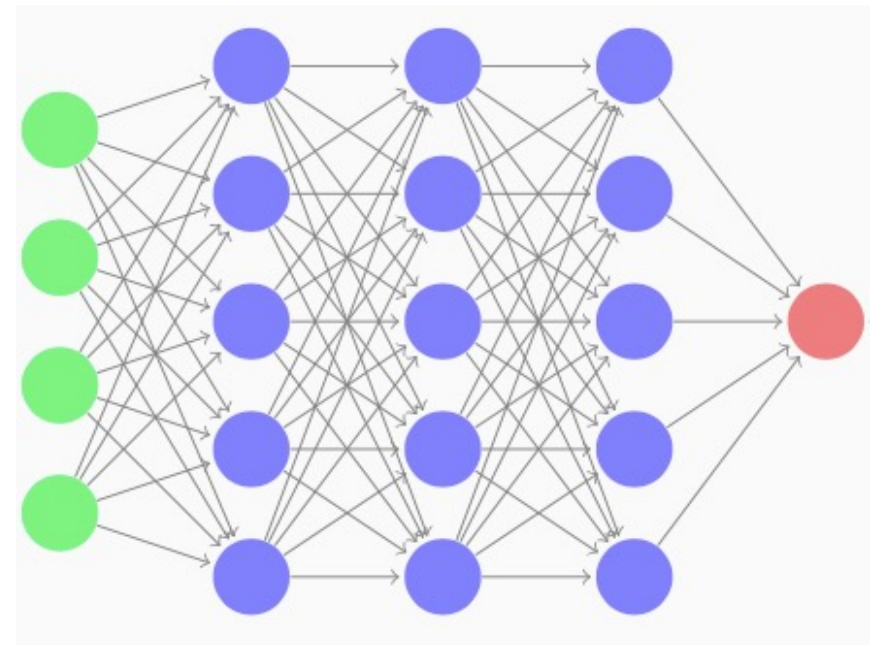
Perceptron et réseau de neurones

On peut combiner des perceptrons et former un réseau de neurones.

Rq : le biais n'est (en général) pas représenté



Une couche caché avec k neurones



Trois couches cachées de 5 neurones

Classification et régression

Regression

- Dernière couche avec fonction d'activation : linéaire ou tanh
- Fonction Loss à optimiser :

$$J(\boldsymbol{\theta}) = \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

Classification

- Dernière couche utilisant la fonction softmax

$$p_j = f_j(\mathbf{h}) = \frac{e^{h_j}}{\sum_k e^{h_k}}$$

- Fonction Loss à optimiser Entropie croisée :

$$J(\boldsymbol{\theta}) = - \sum_{i=1}^n \sum_{j=1}^k y_{i,j} \log(\hat{y}_{i,j})$$

Exemple classification

On souhaite reconnaître des images de chien et de chat.

Features : X (codé par 3 matrices donnant les niveaux RGB pour chaque pixel).

$$m = \text{Nb Pixel} \times 3$$



Cible : y (codé 0 ou 1).

			...
"Chat"	"Chien"	"Chat"	...
{1}	{0}	{1}	

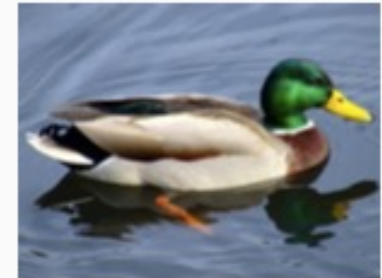
On entraîne un réseau qui estime $\hat{y}_i = f(X)$, qui s'interprète comme la probabilité que l'image soit un chat. Si $\hat{y}_i > 0.5$ alors on peut classer l'image comme celle d'un chat.

Exemple classification

On souhaite reconnaître des images de chien, de chat et de canard.

Features : X (codé par 3 matrices donnant les niveaux RGB pour chaque pixel).

$$m = \text{Nb Pixel} \times 3$$



...

Cible : y , un triplet d'entier de $\{0,1\}$

"Chat"

$\{1,0,0\}$

"Chien"

$\{0,1,0\}$

"Canard"

$\{0,0,1\}$

...

On entraîne un réseau qui estime \hat{y}_i dont la valeur est par exemple (0.1,0.7,0.3) les probabilités que l'image soit un chat, un chien ou un canard.

Travaux pratiques séance 3

- Deux grands types de banquise:

Banquise saisonnière (0-2m)



Banquise pluriannuelle (2 à 4m)



- Observation AMSR :
Radiomètre micro-onde aux fréquences 6.9, 10.65, 18.7, 23.8, 36.5 et 89.0GHz. Deux observations par jour à une résolution de 5 à 56 km

