

Possibilités offertes par le machine learning

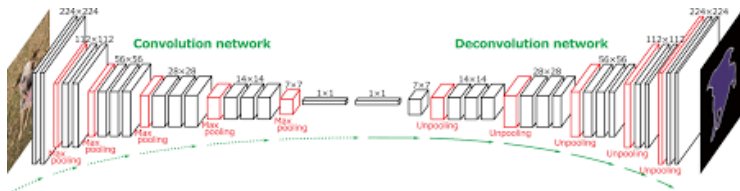
De l'importance de **l'algorithme** (et de ses paramètres)

Définition du modèle :

- Quel type d'entrée ? (image, son, texte, capteurs,...)
- Quelle dimension en entrée ? en sortie ?
- Quelle structure ? Combien mettre de couches deep ?
- Attention network ? Memory network ? Convolutions ? Softmax ?
Renforcement ? Autoencoder ?
- Quelle loss ?
- Quel méta-paramètres ?
- ...

- 1 - initialisation aléatoire du modèle
- 2 - Tant que(critère arrêt == 0)
 - Sélection aléatoire d'un **batch** de données
 - **Forward** : Passe avant du **batch** dans le modèle
 - Calcul de l'erreur par rapport aux sorties attendues
 - **Backward** : Rétropropagation du gradient de l'erreur en fonction des paramètres dans le modèle (mise à jour du modèle)
 - Calcul critère arrêt
- 3 - Calcul de l'erreur sur un échantillon de données **qui n'ont JAMAIS été vues par le modèle pendant l'apprentissage !**

Auto-encoder



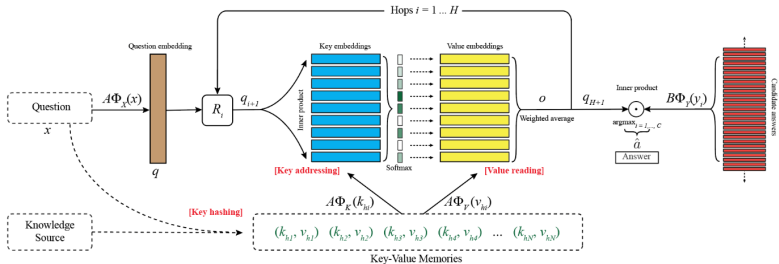
Attention Network

Machine Learning

Modèle à attention et Modèle à Mémoire

Choix de l'algorithme

Memory Network



Fonction de perte (Loss function)

\approx Distance entre la sortie et la cible

Sortie :

0.0	0.1	0.4	0.0	0.0	0.2	0.1	0.0	0.2	0.0
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Cible :

0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Fonction de perte (Loss function)

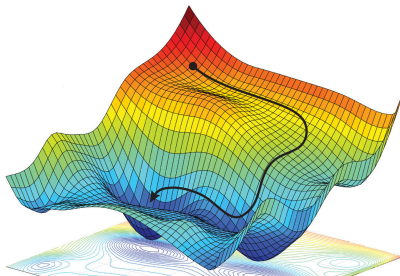
Calcul du gradient de l'erreur par rapport aux paramètres :

$$\frac{\partial Err}{\partial w_i}$$

Mise à jour :

$$w'_i = w_i - \gamma * grad$$

où : $0 < \gamma < 1$ (learning rate)



Fonction de perte (Loss function)

Regression Loss Functions :

- Erreur **absolue**
- Erreur **lissée**
- ...

Classification Loss Functions :

- Entropies-croisées
- Log vraisemblance
- Loss à marge
- ...

Embedding Loss Functions :

- L1(,L2,...)
- Distance cosinus
- ...



Exemple de descente de gradient sur un cas simple : la regression

- $y = a.X + b$
- $L = \frac{1}{2n} \sum_{i=[1..n]} (y_i^* - y_i)^2$
- $L = \frac{1}{2n} \sum_{i=[1..n]} (y_i^* - (a.x_i + b))^2$
- ...
- $\frac{\partial L}{\partial a} = \frac{1}{n} \sum_{i=[1..n]} (a.X + b - y^*).X$
- $\frac{\partial L}{\partial b} = \frac{1}{n} \sum_{i=[1..n]} (a.X + b - y^*)$

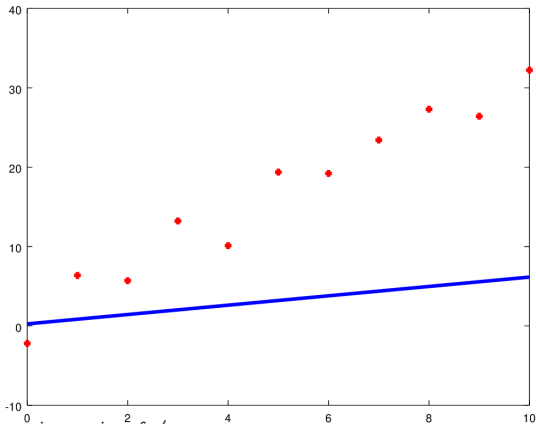
M.A.J :

- $a = a - \gamma \frac{\partial L}{\partial a}$
- $b = b - \gamma \frac{\partial L}{\partial b}$
- où $1 > \gamma > 0$ (learning rate)

Gradient et mise à jour

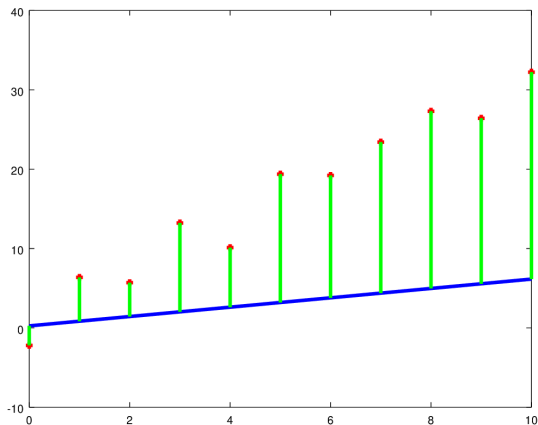
initialisation au hasard ($\gamma = 0.01$)

- $a = 0.58$
- $b = 0.25$



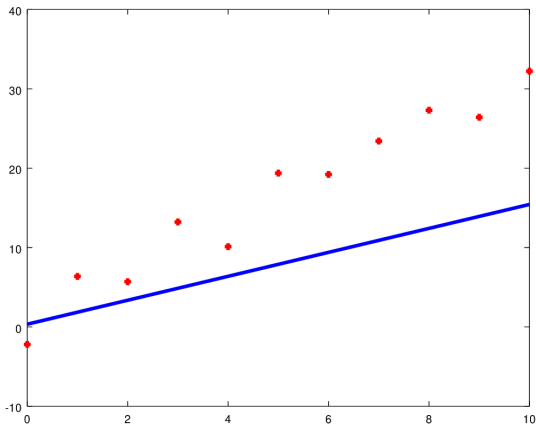
Gradient et mise à jour

- $a = 0.58$
- $b = 0.25$



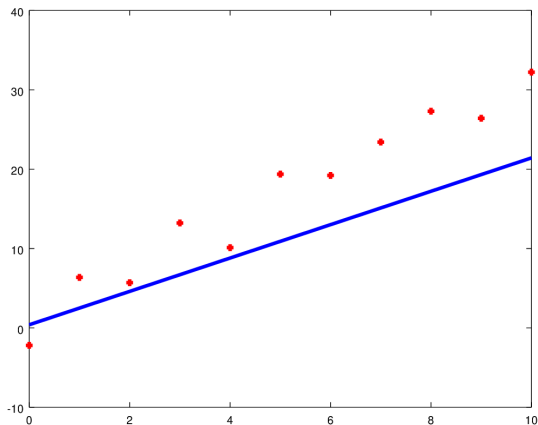
Gradient et mise à jour

- $a = 1.50$
- $b = 0.35$



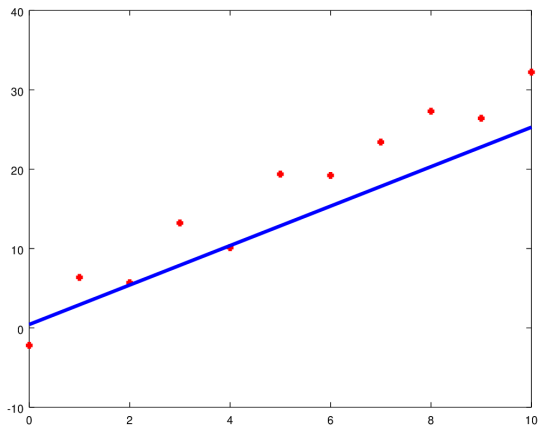
Gradient et mise à jour

- $a = 2.10$
- $b = 0.40$



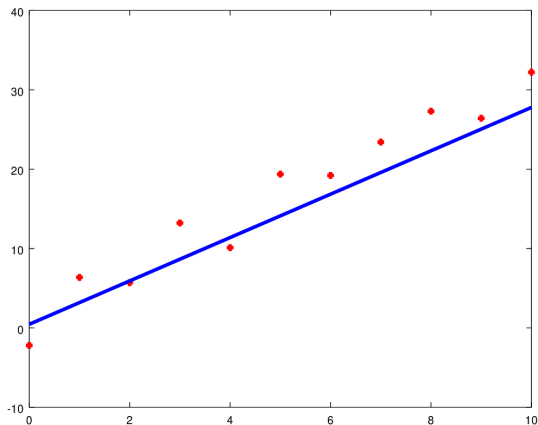
Gradient et mise à jour

- $a = 2.48$
- $b = 0.43$



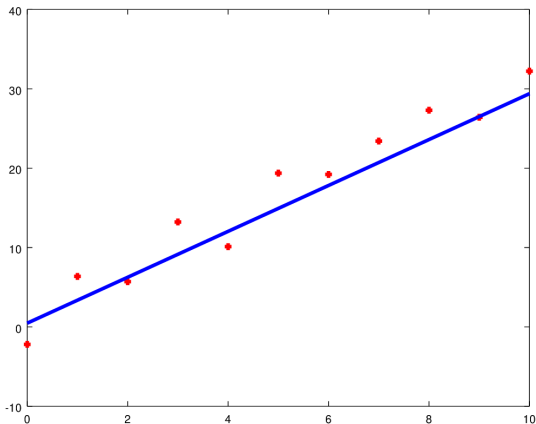
Gradient et mise à jour

- $a = 2.73$
- $b = 0.46$



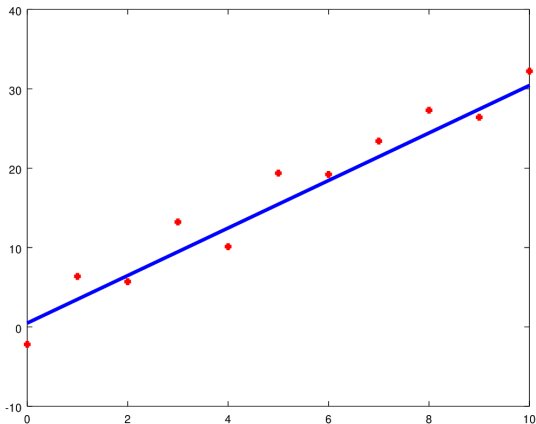
Gradient et mise à jour

- $a = 2.89$
- $b = 0.47$



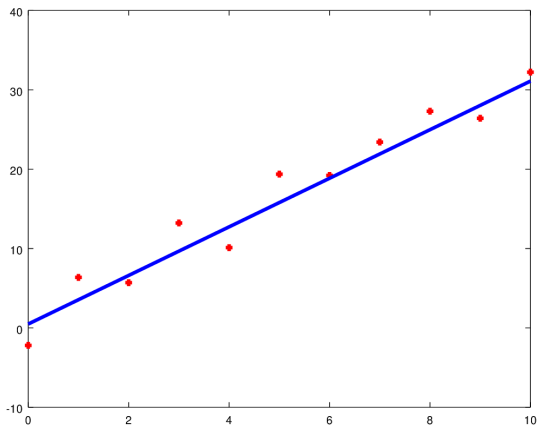
Gradient et mise à jour

- $a = 2.99$
- $b = 0.48$



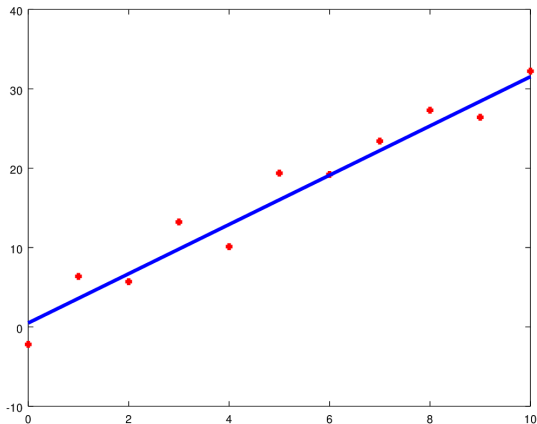
Gradient et mise à jour

- $a = 3.06$
- $b = 0.49$



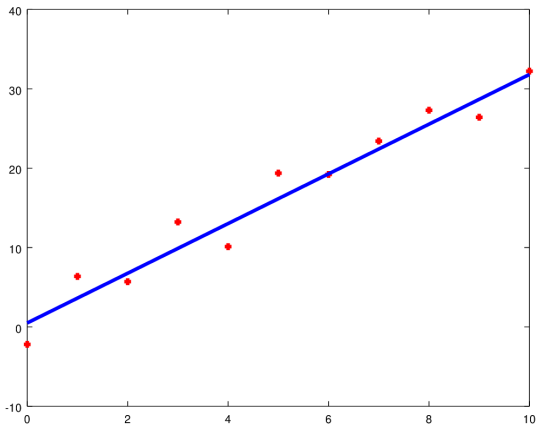
Gradient et mise à jour

- $a = 3.10$
- $b = 0.49$



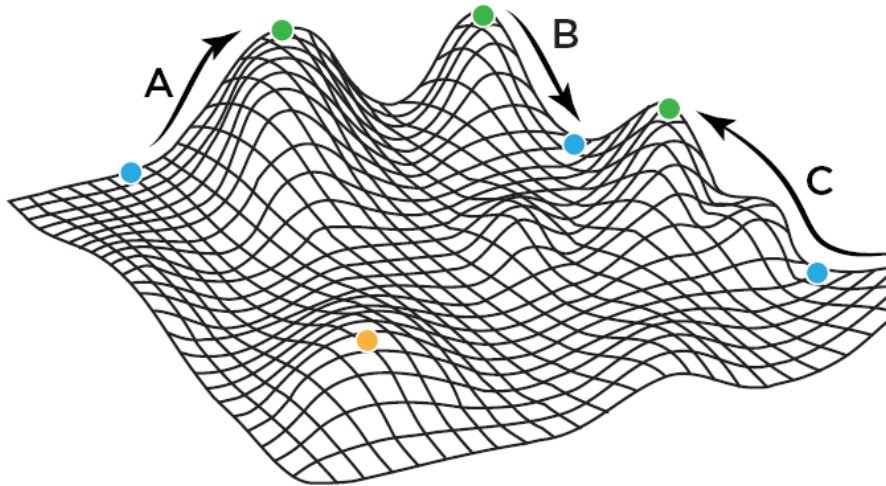
Gradient et mise à jour

- $a = 3.13$
- $b = 0.50$

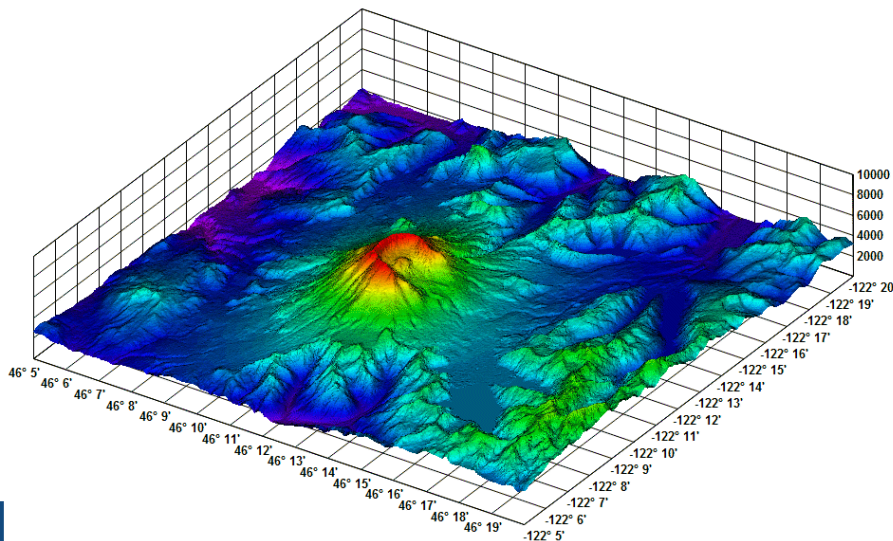


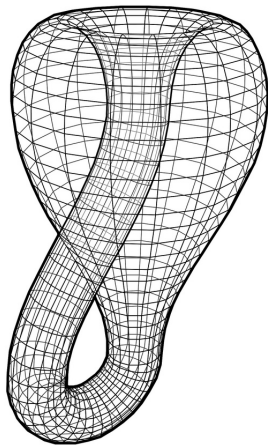
Convergence du modèle vers l'optimal ?

Choix de l'algorithme



Choix de l'algorithme





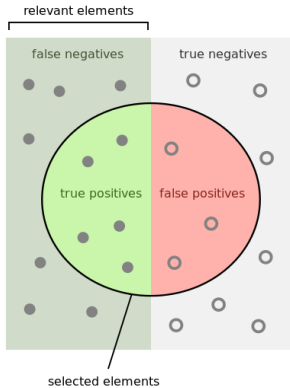
$$Précision_i = \frac{vrai\ positifs_i}{vrai\ positifs_i + faux\ positif_i}$$

$$Précision = \frac{\sum_{i=1}^n précision_i}{n}$$

$$Rappel_i = \frac{vrai\ positifs_i}{vrai\ positifs_i + faux\ négatifs_i}$$

$$Rappel = \frac{\sum_{i=1}^n rappel_i}{n}$$

Evaluation des performances



How many selected items are relevant?

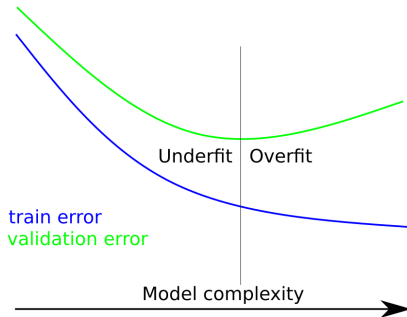
$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

Choix de l'algorithme

Quel critère d'arrêt ?



Choix de l'algorithme

sous-apprentissage/sur-apprentissage

