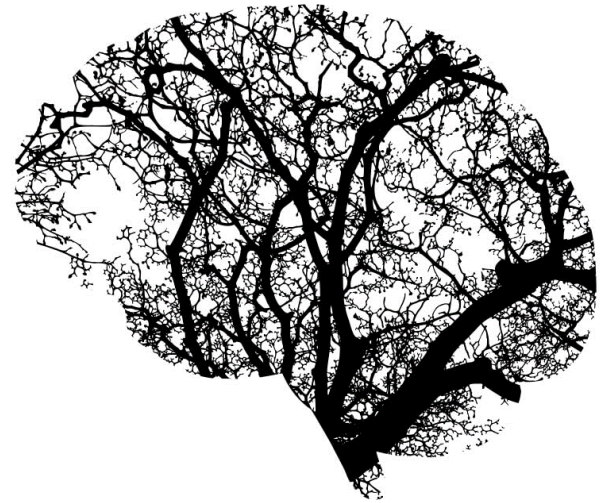


# Introduction à l'Intelligence artificielle

Des racines à l'arbre



# Sommaire

- Un peu d'histoire : Les trois grandes périodes
- Le machine learning
- Le deep learning
- Apprentissage d'un neurones
- Pour conclure
- Quiz



Un peu d'histoire : Les trois grandes périodes

## **1950-1980 - Les IA de type déterministe :**

l'idée d'intelligence artificielle émerge en 1950 quand Alan Turing se demande si une machine peut « penser ».

À ce niveau, l'ordinateur a un rôle passif, il exécute simplement des tests et retourne un résultat.

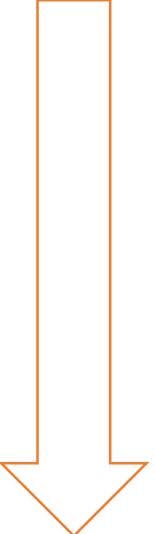
### Exemple :

Dans les critères d'attribution d'un prêt immobilier sur un site web si l'emprunteur est en CDI depuis au moins 20 mois et si son apport est supérieur à 20 % du montant d'achat, alors un accord de principe est conclu.

# Un peu d'histoire : Les trois grandes périodes

## **1980-2015 - Les IA capable d'apprendre :**

Cette fois, l'ordinateur essaye d'apprendre par lui-même à partir de données servant de référence. C'est l'époque du machine learning (l'apprentissage automatique en français).

- 
- 1980 : Les système expert tel que Mycin pour diagnostiquer les maladie infectieuse du sang.
  - 1989 : LeNet est le premier programme à reconnaitre des codes postaux manuscrits.
  - 1990 : Démocratisation du filtre anti-spam.
  - 2011 : Le programme Watson d'IBM remporte le jeu télévisé Jeopardy (traitement du langage naturel)
  - 2015 : Mise en service des Google Car en Californie.
  - 2015 : La société DeepMind crée le programme AlphaGO, le premier programme à battre un joueur professionnelle de Go.

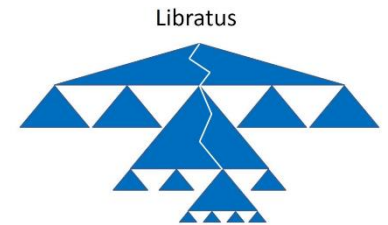
# Un peu d'histoire : Les trois grandes périodes

## **2017 à aujourd'hui – L'auto-apprentissage :**

En 2017 Le programme Libratus est vainqueur d'un tournoi de poker l'opposant à quatre joueurs professionnels humains.

Les techniques qu'il utilise devraient lui permettre de s'adapter à n'importe quelle **situation à « information incomplète »**.

En 2017 lui succède le programme AlphaGo Zero capable d'apprendre sans utiliser de base de connaissances et ceci en jouant uniquement contre lui-même. C'est la première expérience abouti **d'auto-apprentissage**.



# Le machine learning

**L'apprentissage automatique** c'est la sciences et l'art de programmer les ordinateurs de sorte qu'ils puissent apprendre à partir de données.

Exemple :

Un filtre anti-spam est un programme d'apprentissage automatique qui peut apprendre à identifier les email frauduleux à partir de spam (signalé par les utilisateurs) et de courriers normaux. Le programme va :

1. Examiner les courriers frauduleux : contient « pour vous », « gratuit », « bravo »,...
2. En déduire des règles de décisions
3. Calculer la probabilités, à chaque fois qu'un nouveaux mail lui est soumis, qu'il soit frauduleux.

*A chaque fois que les pirates feront évoluer leur mail spam pour contourner le filtre, le programme mettra à jour ses règles de décisions automatiquement grâce au machine learning.*

# Le machine learning

A chaque problème de machine learning un modèles mathématiques est associés.

Pour notre filtre anti-spam c'est la régression logistique qui semble le modèle le plus adapté. En effet ce peut prendre en entré des variables quantitatives, qualitatives ou binaire et produit en sortie la probabilité d'appartenir à un groupe et pas à un autre.

$$P(Y = 1|X) = \frac{1}{1 + \exp(W^T X + b)}$$

- $X$  est le vecteur qui contient l'ensemble des caractéristiques du mail dont on cherche à prédire si c'est un spam ou non.

$$X = [x_1, x_2, \dots, x_d]$$

- $W$  est le vecteur qui contient les poids du modèles. Ceux-ci sont calculé l'or de l'apprentissage.

$$W = [w_1, w_2, \dots, w_d]$$

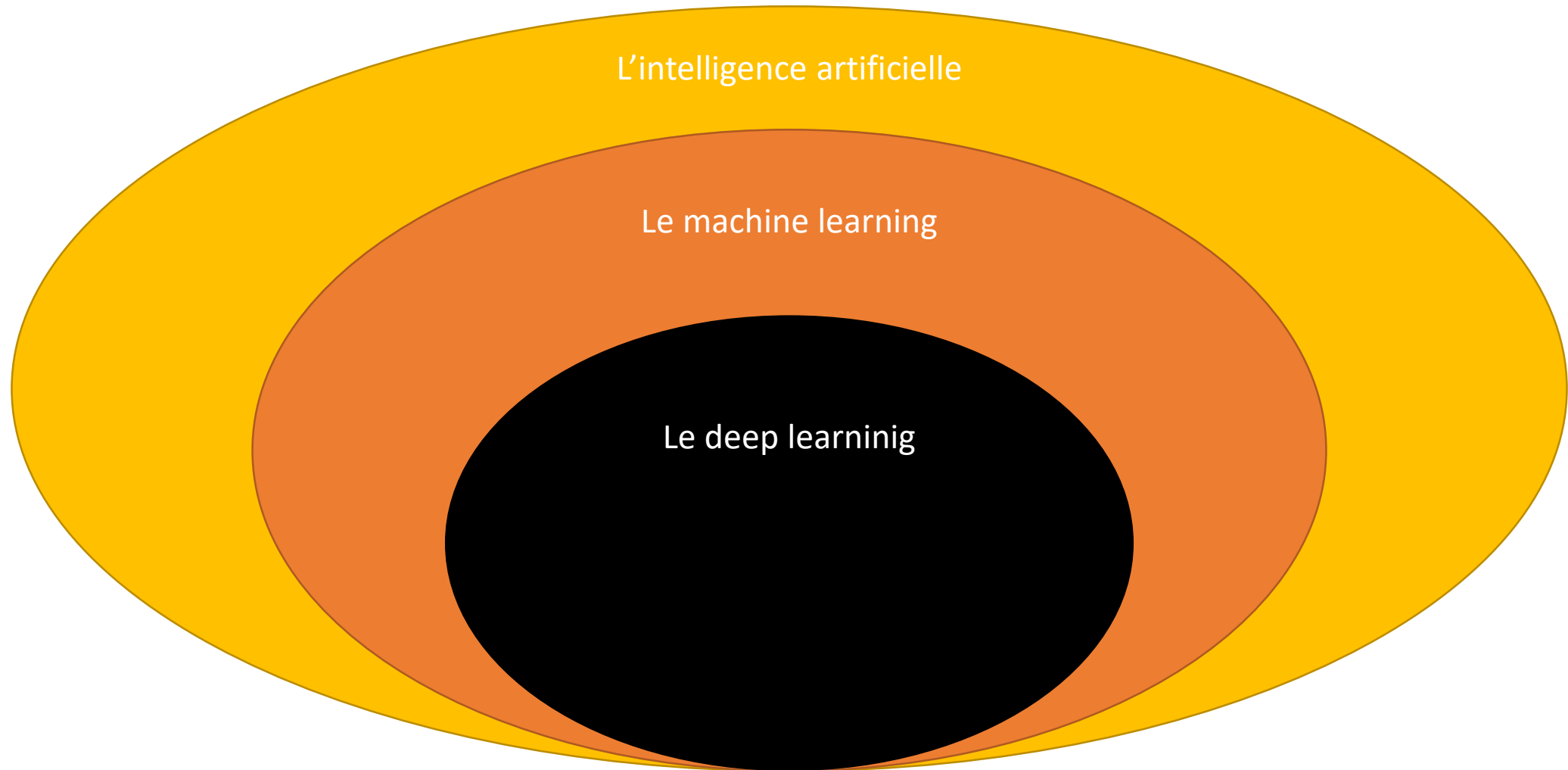
- $Y$  est la variable cible qui prendra la valeur 1 si c'est un spam, 0 sinon.

# Taxinomie des algorithmes de machine learning

Algorithme	Mode d'apprentissage	Type de problème à traiter
Régression linéaire multiple	Supervisé	Régression
Régression polynomial	Supervisé	Régression
Régression régularisée	Supervisé	Régression
Naive Bayes	Supervisé	Classification
Régression logistique	Supervisé	Classification
Modèle ARIMA	Supervisé	Série temporelle
Clustering hiérarchique	Non supervisé	-
Clustering non hiérarchique	Non supervisé	-
Analyse en composantes principales	Non supervisé	-
<b>Réseaux de neurones</b>	Supervisé et non supervisé	Régression, classement, ...



# Le machine learning



# Le machine learning

## Cas d'application de modèle simple dans le domaine du web :

- Analyse de votre personnalité pour vous proposer les produits qui vous corresponde le mieux.
- Sélection des publicités qui attirerons votre attention.
- Analyse du comportement pour s'assurer que vous n'êtes pas un fraudeur.
- Filtre mail anti spam.

## Application autre que le web :

- Prédiction de série temporelle comme le cours de la Bourse ou du Bitcoin.
- Détection d'anomalies de production.
- Optimisation de centre d'appel.
- Classification automatiquement de documents.

# Le deep learning

Le deep learning c'est quoi ?

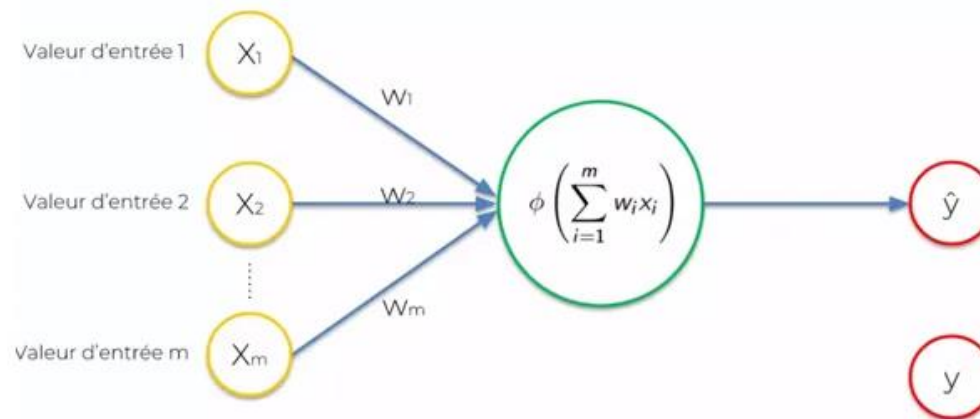
L'apprentissage profond fait appel aux réseaux de neurones profond, qui requièrent souvent de gros volume de données et une immense puissance de calcul.

On dit qu'il sont profond car ils sont organisés en de nombreuses couches de neurones successives.

# Le deep learning

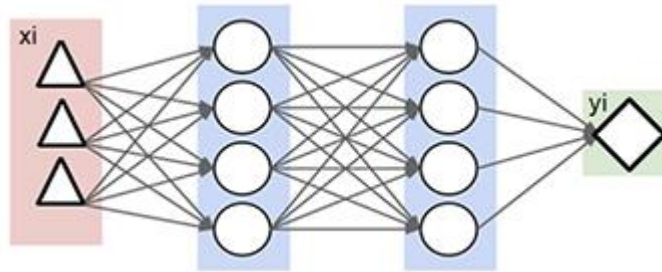
Principe de fonctionnement d'un neurone :

- Chaque neurone contient un modèle mathématique du type de la régression logistique ou autre.
- Des milliers d'entrées dont le rôle consiste à véhiculer l'information vers le centre du neurone,
- Une seule sortie dont le rôle consiste à diffuser l'information à d'autres neurones.



# Le deep learning

Voici un schéma représentant un réseau à huit neurones. Les neurones sont représentés par des cercles. À gauche, nous trouvons la **couche** des entrées (input layer en anglais).



- Chaque entrée, contient les mesures d'une valeur  $x_i$  (ex : température, la taille et le poids).
- Les neurones de la couche de gauche sont connectés aux entrées.
- La dernière couche contient une fonction d'erreur. Cette fonction récupère l'ensemble des résultats de la couche précédente et calcule une unique valeur de sortie sous la forme d'un réel. Plus ce chiffre est bas, meilleur est le score.

# Le deep learning

Réseaux de neurones convolutif (CNN)

Réseaux de neurones récurrent (RNR)

Apprentissage génératif avec autoencodeurs (GAN)

Apprentissage par renforcement

# Le deep learning

## Cas d'application de modèle de deep learning:

- Détection de tumeurs sur les scanners cérébraux (CNN).
- Classification automatique d'article de presse (TALN et RNR).
- Identification automatique de commentaire inapproprié sur les forums (TALN et RNR).
- Mise en place d'un chatbot (TALN et CLN).
- Interface de commande vocal (RNR et CNN).
- ...

# Apprentissage d'un neurones

Faire apprendre un neurone c'est calculer les poids associé à son modèle linéaire qu'il contient.

Pour calculer les poids  $W$  de notre modèle linéaire la méthode le plus couramment utilisé est la descente de gradient.

Dans un premier temps on va définir une fonction de coût MSE:

$$MSE(W) = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$



# Apprentissage d'un neurones

Le but est de trouver les valeurs de  $W$  qui vont minimiser cette fonction

Etape de l'algorithme de descente de gradient :

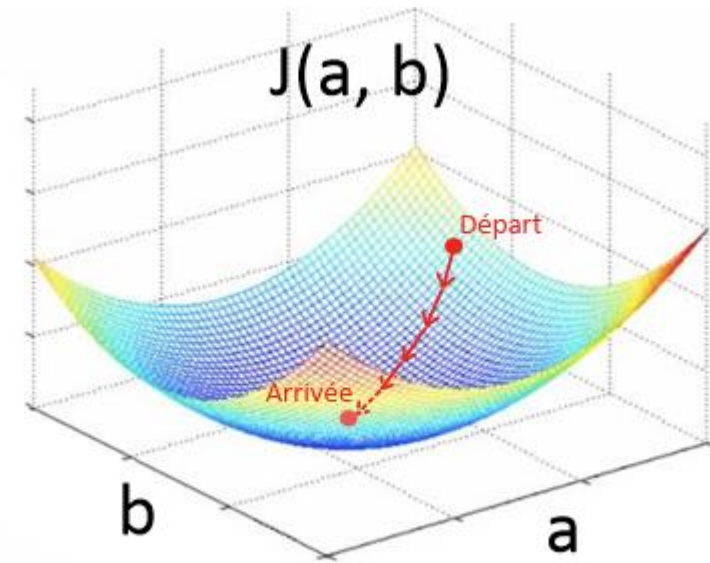
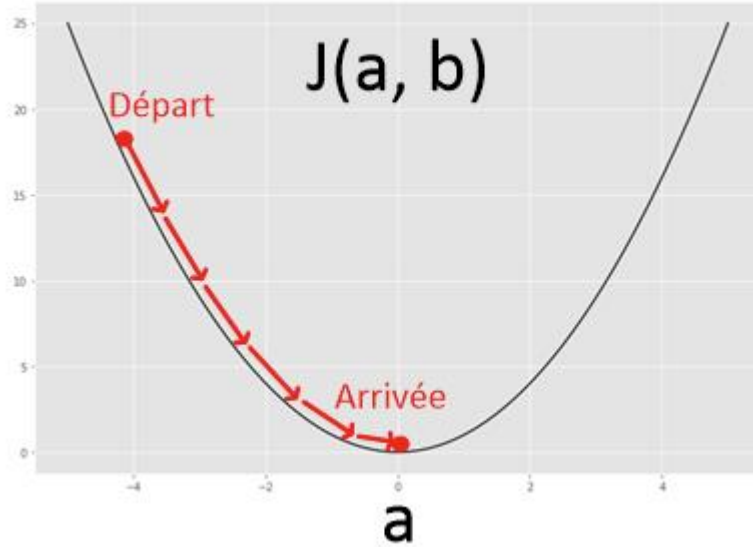
1. On initialise les poids  $W$  de manière aléatoire
2. On calcule quel est la modification de coût lorsqu'on fait varier  $W$  d'une toute petite valeur dans un sens.

$$W^{\text{étape suivante}} = W - \eta \nabla_W \text{MSE}(W)$$

Si l'erreur diminue alors on avance dans le bon sens et on conserve  $W^{\text{étape suivante}}$  pour calculer la valeur d'après. Sinon on repart de  $W$  et on change de sens.

3. On réitère l'étape précédente jusqu'à atteindre les critères de convergences : Nombre d'itération ou différence entre valeurs successive très petite.

# Apprentissage d'un neurones



# Apprentissage d'un neurones

Les **hyperparamètres** sont les paramètres propres au programme d'apprentissage. Ils n'ont rien à voir avec les mesures du terrain, mais interviennent dans le processus d'apprentissage :

- Le nombre d'itérations : s'il est trop bas, l'algorithme ne peut pas converger et nous allons trouver une valeur de  $a$  mauvaise. S'il est trop important, le temps nécessaire trop long.
- Le pas déterminé grâce au taux d'apprentissage  $\eta$  : Lorsque nous sommes loin de la valeur optimale, nous avons intérêt à choisir un pas important pour aller plus vite.  
Quand nous sommes proches de la valeur optimale, il nous faut un pas plus faible pour converger. Un pas trop important nous ferait rebondir autour de la solution optimale sans jamais pouvoir la trouver.

# Conclusion

Pour développer une IA comprendre comment un algorithme fonctionne n'est pas forcément indispensable mais fortement recommandé si on veut être un bon développeur d'IA.

Il faut savoir choisir quel est le modèle le plus adapté au problème que l'on souhaite résoudre et savoir le paramétrer correctement.

Pour bien comprendre leur fonctionnement il convient d'avoir certaine base en mathématiques notamment dans les domaines :

- De l'algèbre linéaire, comprendre les vecteurs et les matrices.
- Du calcul différentiel, comprendre la notion de dérivé et dérivée partielle.
- Des statistiques et probabilités, variance, biais, fonction de densité, ...

# Pour aller plus loin

- Compromis biais variance
- Métrique de performance
- Test et validation
- Apprentissage supervisé et non supervisé
- Fonction d'activation
- Difficulté liée à l'apprentissage
  - Données insuffisante
  - Données non représentative
  - Surapprentissage
  - ...
- ...

# Quiz

## 1. Un neurone informatique :

- a. Possède plusieurs entrées et une sortie
- b. Possède des états
- c. Associe un poids à chaque entrée
- d. Calcule sa sortie en fonction de ses entrées et de ses poids
- e. Possède un niveau d'erreur
- f. Je ne sais pas

# Quiz

2. Pour calculer sa valeur de sortie, un neurone utilise :

- a. Une fonction polynôme du premier degré
- b. Une fonction polynôme du second degré
- c. Des fonctions trigonométriques
- d. Je ne sais pas

# Quiz

3. En IA, une image est considérée comme :

- a. Un hyperparamètre
- b. Une série de valeurs
- c. Les poids du réseau
- d. Je ne sais pas



# Quiz

4. CIFAR 10 est :

- a. Un algorithme d'apprentissage
- b. Une base de données d'images
- c. Un célèbre blog d'IA
- d. Le paramètre réglant la rapidité de convergence
- e. Je ne sais pas

# Quiz

5. La fonction d'erreur permet :

- a. D'estimer la qualité de la prédiction d'un échantillon
- b. D'améliorer les poids du réseau
- c. De calculer la première couche
- d. Je ne sais pas

# Quiz

6. L'algorithme du minimax utilise :

- a. Des simulations de partie
- b. Des réseaux de neurones
- c. Des bases de données
- d. Je ne sais pas

# Solution du quiz

1. a, c et d
2. a
3. b
4. b
5. a et b
6. a