

Introduction à l'Intelligence Artificielle

<https://sleek-think.ovh/enseignement>

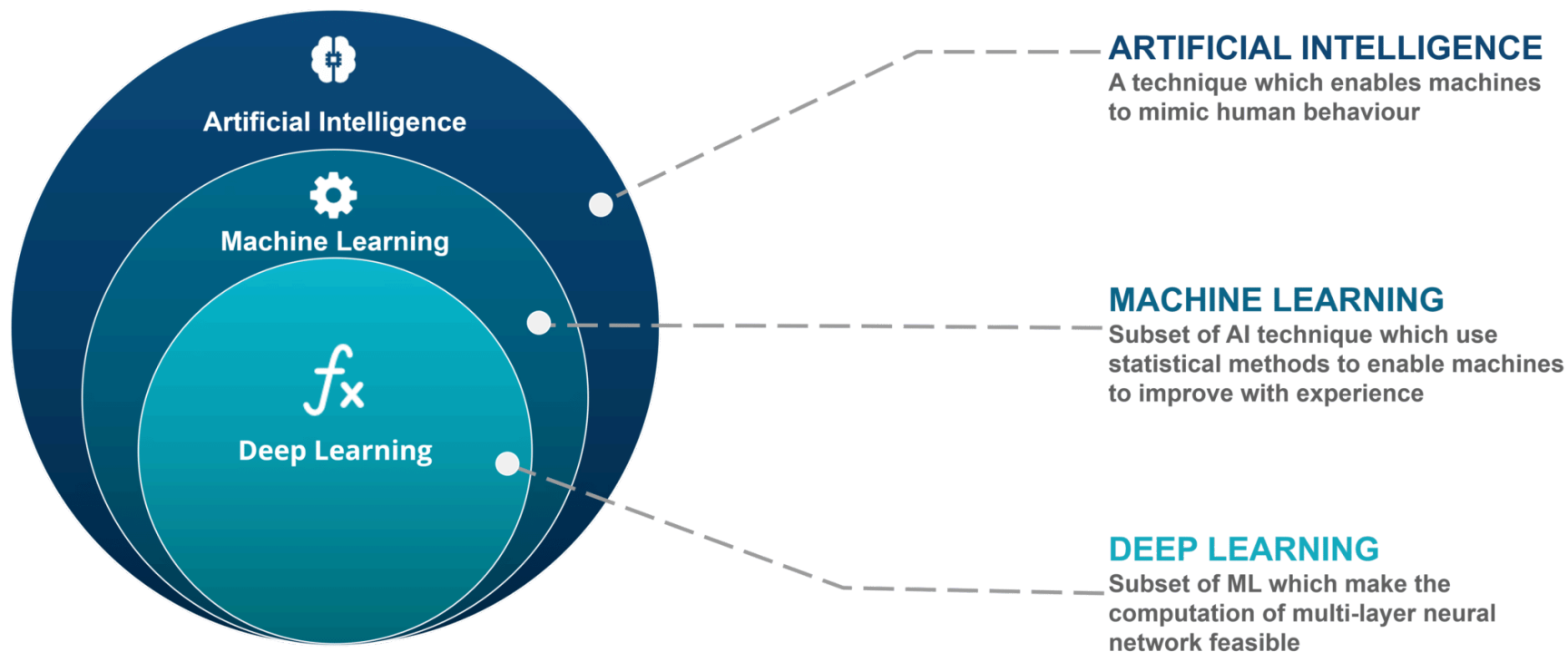
Dr. Jehan-Antoine Vayssade

Introduction à l'Intelligence Artificielle

- Définitions et concepts fondamentaux
- Histoire de l'IA
- Grandes familles de méthodes
- Applications majeures

L'IA ?

L'intelligence artificielle désigne un ensemble de techniques et de systèmes visant à **simuler** certaines capacités cognitives ou de raisonnement. En réalité, c'est l'humain qui insuffle son intelligence dans ces systèmes en définissant des règles de décision et en utilisant des outils mathématiques.



L'IA la plus simple

Prenons l'exemple d'une image en couleur (RGB) :

- **R** : Composante Rouge
- **G** : Composante Verte
- **B** : Composante Bleue

Une règle simple pour détecter la végétation :

Si $2G - R - B > 0$ alors Végétation, sinon Non-Végétation

Cette équation simple fonctionne car :

- La végétation réfléchit beaucoup le vert (G élevé)
- Elle absorbe le rouge et le bleu (R et B faibles)

Cette "IA" repose sur une règle de décision formalisée par l'homme pour résoudre un problème précis. L'optimisation est difficile et nécessite beaucoup d'essai/erreur.

Le Machine Learning

Cette fois-ci, nous définissons une **équation générique**, mais nous déléguons **l'apprentissage des paramètres** à un algorithme. Nous allons utiliser des outils mathématiques et des jeux de données pour ajuster ces paramètres.

Si $\alpha G + \beta R + \gamma B + b > 0$, alors **Végétation**, sinon **Non-Végétation**

- Les coefficients α, β, γ et le biais b sont appelés des **paramètres** du modèle.
- On peut utiliser la **méthode des moindres carrés** pour résoudre ce problème.
- Concrètement, on cherche à minimiser la somme des carrés des différences :

$$\text{Erreur} = \sum_i (\text{valeur prédite}_i - \text{valeur réelle}_i)^2$$

Le modèle ajuste alors les paramètres α, β, γ , et b pour que les **prédictions soient le plus proche possible des données réelles**.

Les Réseaux de Neurones

Deleger l'apprentissage de l'équation et des paramètres !

Le Deep Learning est une branche du Machine Learning basée sur les réseaux de neurones artificiels. Ces réseaux s'inspirent du fonctionnement du cerveau humain :

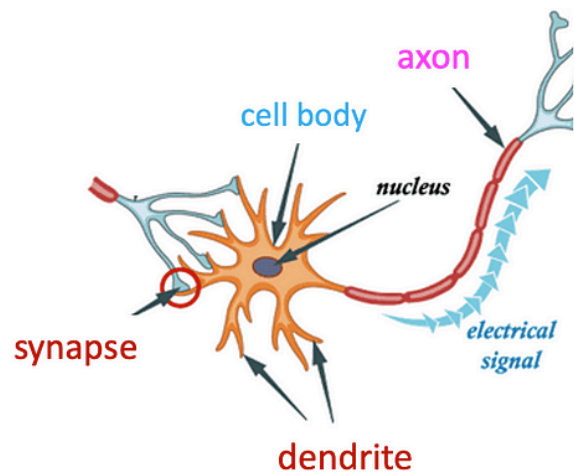
- **Neurone artificiel** : Unité de base qui reçoit des entrées, les pondère, et produit une sortie
- **Réseau de neurones** : Ensemble de neurones connectés en couches
- **Deep Learning** : Utilisation de réseaux avec plusieurs couches cachées

Histoire et Théorie

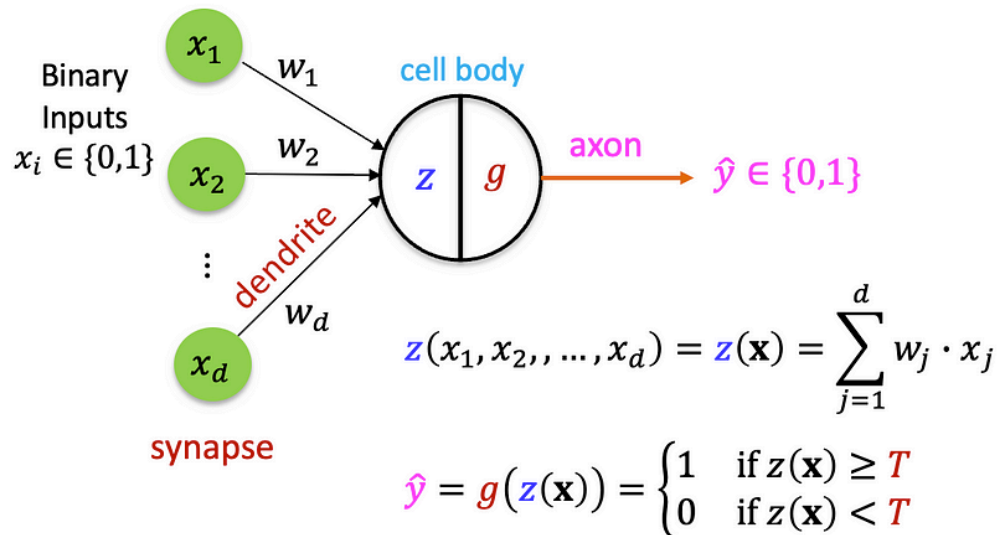
- **1943** : Warren McCulloch et Walter Pitts proposent le premier modèle mathématique du neurone
- **1958** : Frank Rosenblatt invente le Perceptron, premier neurone artificiel
- **1986** : Geoffrey Hinton et ses collègues développent la rétropropagation du gradient
- **2012** : AlexNet révolutionne la vision par ordinateur (modèle convolutif)

Le neurone (theorique 1943)

A Biologic Neuron



A McCulloch-Pitts Neuron



Le neurone (pratique 1958)

Le perceptron est le premier modèle de neurone artificiel, créé par Frank Rosenblatt en 1958.

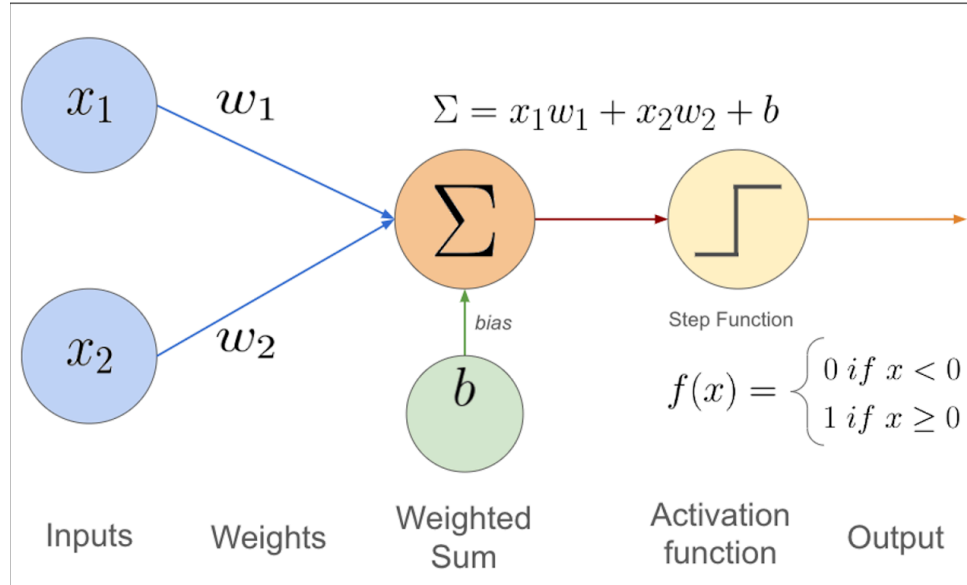
- Une seule couche de neurones
- Une fonction d'activation à seuil (Heaviside $\{0, 1\}$)
- Capable de séparer linéairement deux classes
- Correspond explicitement à notre cas initial (vegetation)

Le perceptron apprend en ajustant ses poids selon la règle : $w_{new} = w_{old} + \eta(y - \hat{y})x$

Où :

- η est le taux d'apprentissage
- y est la sortie désirée
- \hat{y} est la sortie prédite
- x est l'entrée

Le neurone

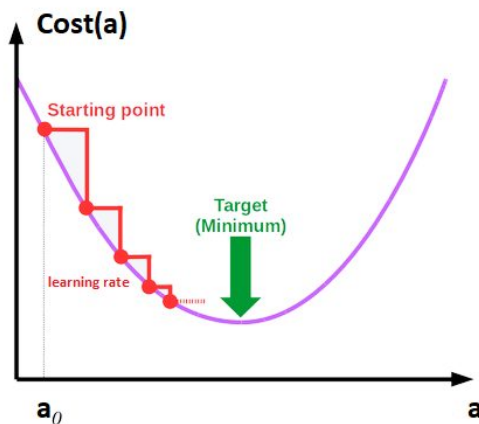


- C'est un hyper-plan separateur ! Dans le cas 2D ici, c'est une droite ...
- Analyse discriminante linéaire 1936 ... Descartes (1637) Euler (1736) Möbius (1865) ...
- 2000 av. J.-C La méthode de la fausse position
- <https://www.numworks.com/fr/blog/interview-marianne-michel/>

La descente de gradient

La descente de gradient est un algorithme d'optimisation.

1. On part d'un point initial (paramètres aléatoires)
2. On calcule le gradient de la fonction de coût (exemple: mse)
3. On se déplace dans la direction opposée au gradient
4. On répète jusqu'à convergence

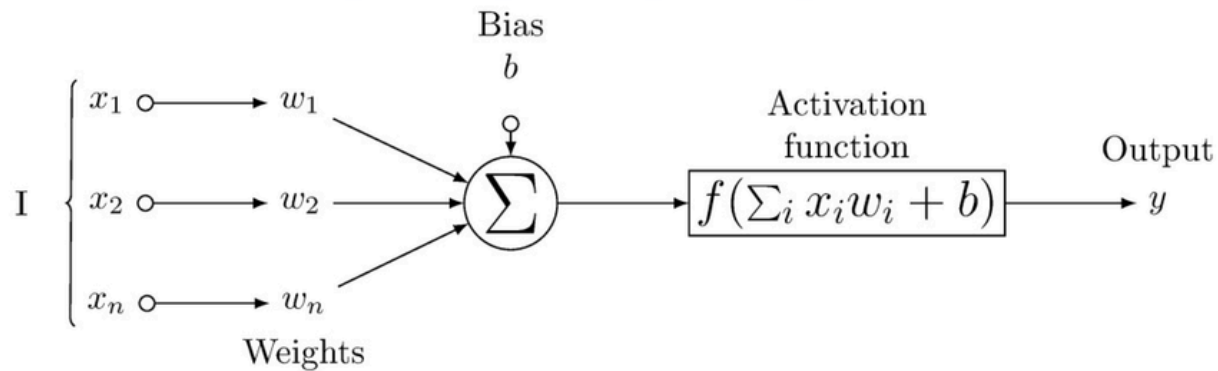
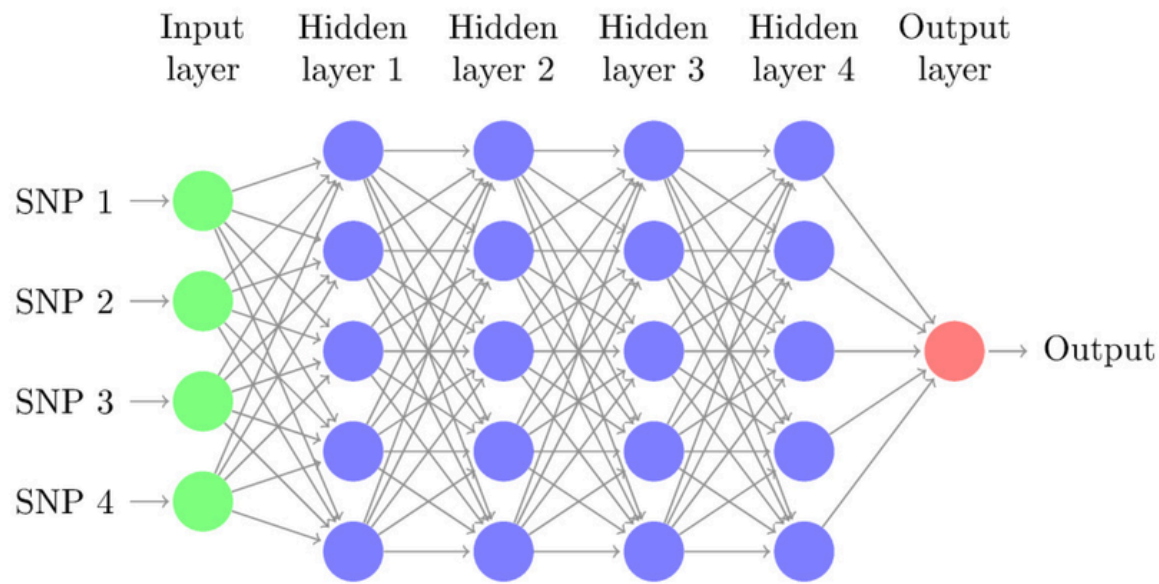


Le réseau de neurones (Multi-Layer Perceptron)

Un réseau de neurones est composé de plusieurs neurones interconnectés organisés en couches :

- **Couche d'entrée** : Reçoit les données brutes
- **Couches cachées** : Traitent l'information de manière séquentielle
- **Couche de sortie** : Produit le résultat final

Chaque connexion entre les neurones a un poids qui est ajusté pendant l'apprentissage. Le réseau apprend en modifiant ces poids pour minimiser l'erreur entre la sortie prédite et la sortie souhaitée.

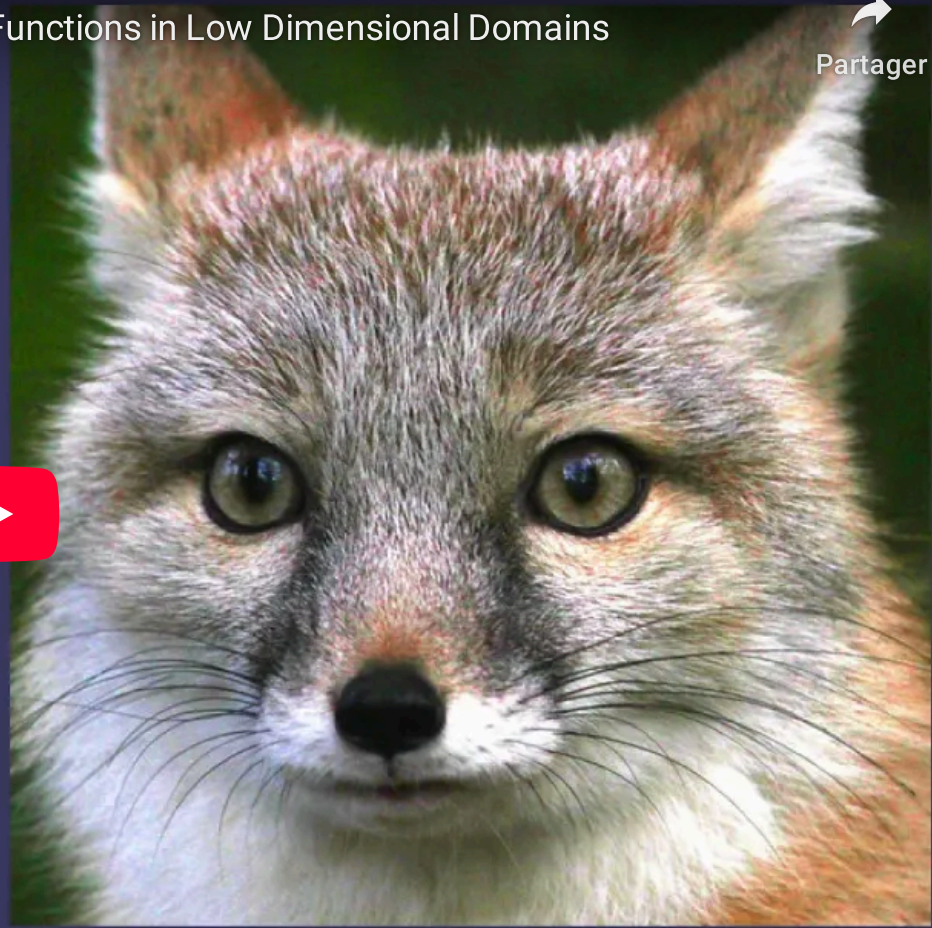
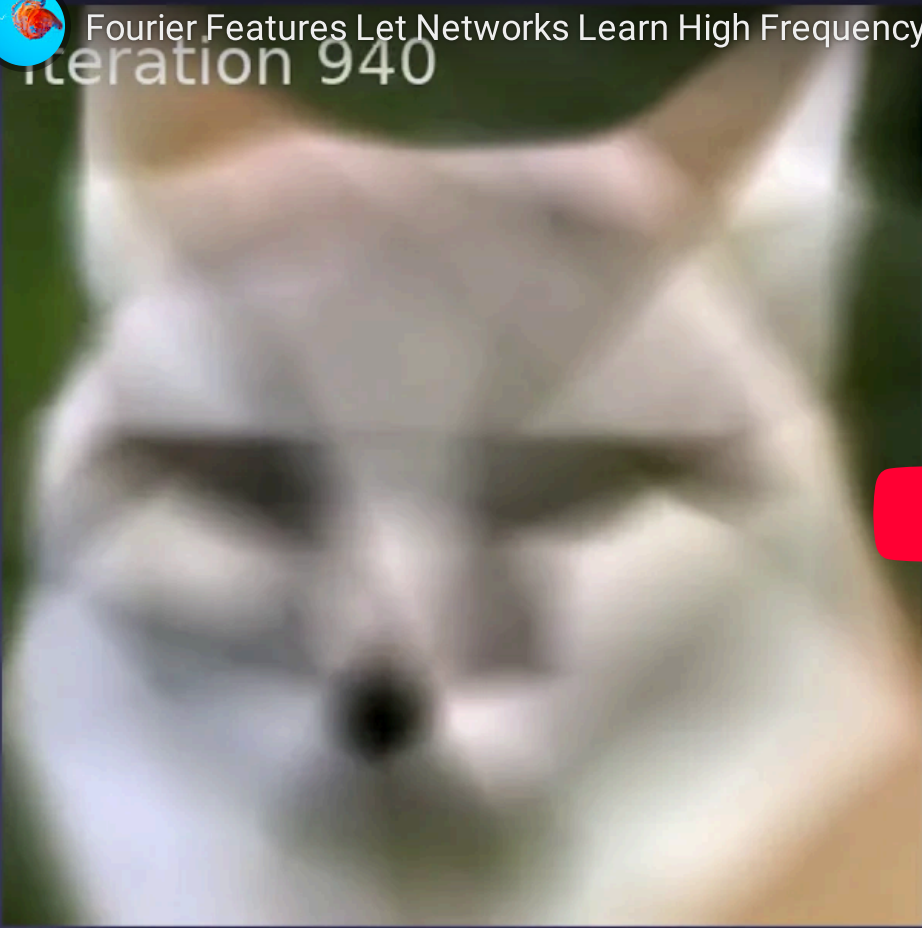




Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains

Iteration 940

Partager



MLP output

Supervision image

Depuis 2013, le Deep Learning prend de plus en plus de place :

- Apprentissage automatique des caractéristiques (features) sans intervention humaine
- Architecture flexible et adaptable à différents types de données
- Excellentes précision sur des tâches complexes
- Automatisation de tâches qui nécessitaient auparavant une expertise humaine
 - DeepIndices

Ce qui ce dit (principalement faux)

- Nécessite de grandes quantités de données d'entraînement
- Coût de calcul important (besoin de GPUs/TPUs depuis 2013)
- Sensibilité aux données bruitées ou biaisées
- Amélioration continue des performances avec plus de données (discutable)
- Forte capacité d'abstraction et de "généralisation" (fortement lié aux données)
- Difficulté d'interprétation des décisions (discutable)
- Complexité des modèles (boîte noire)

Pourquoi a t'on besoin d'une fonction d'activation ?

Exercice, supposon une variable x defini dans l'interval $[-10, 10]$

Supposon 3 neurones avec les coeficients suivants :

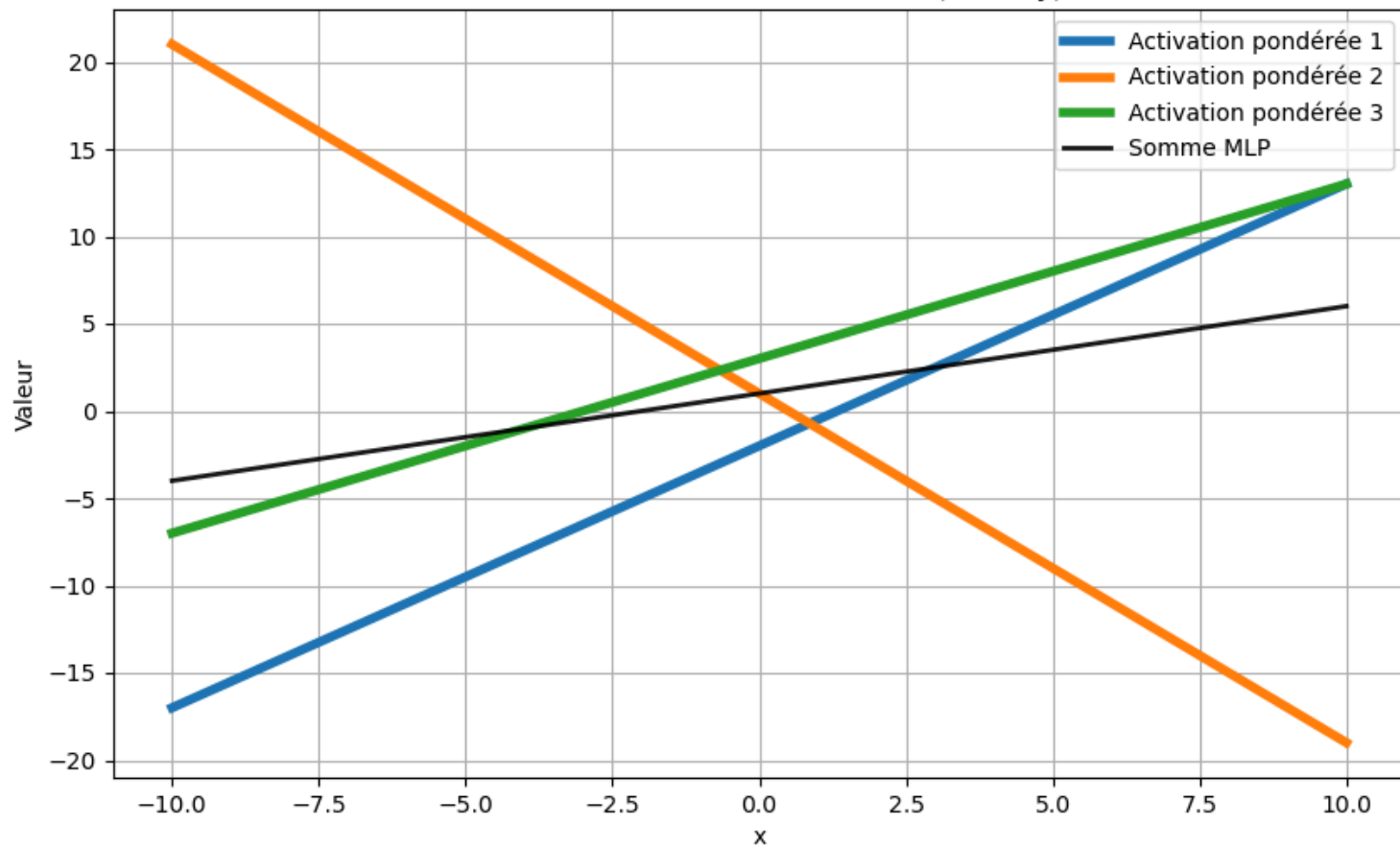
weights = $[1.5, -2, 1]$

biases = $[-2, 1, 3]$

Quel est la sortie, sans fonction d'activation, apres simplification ?

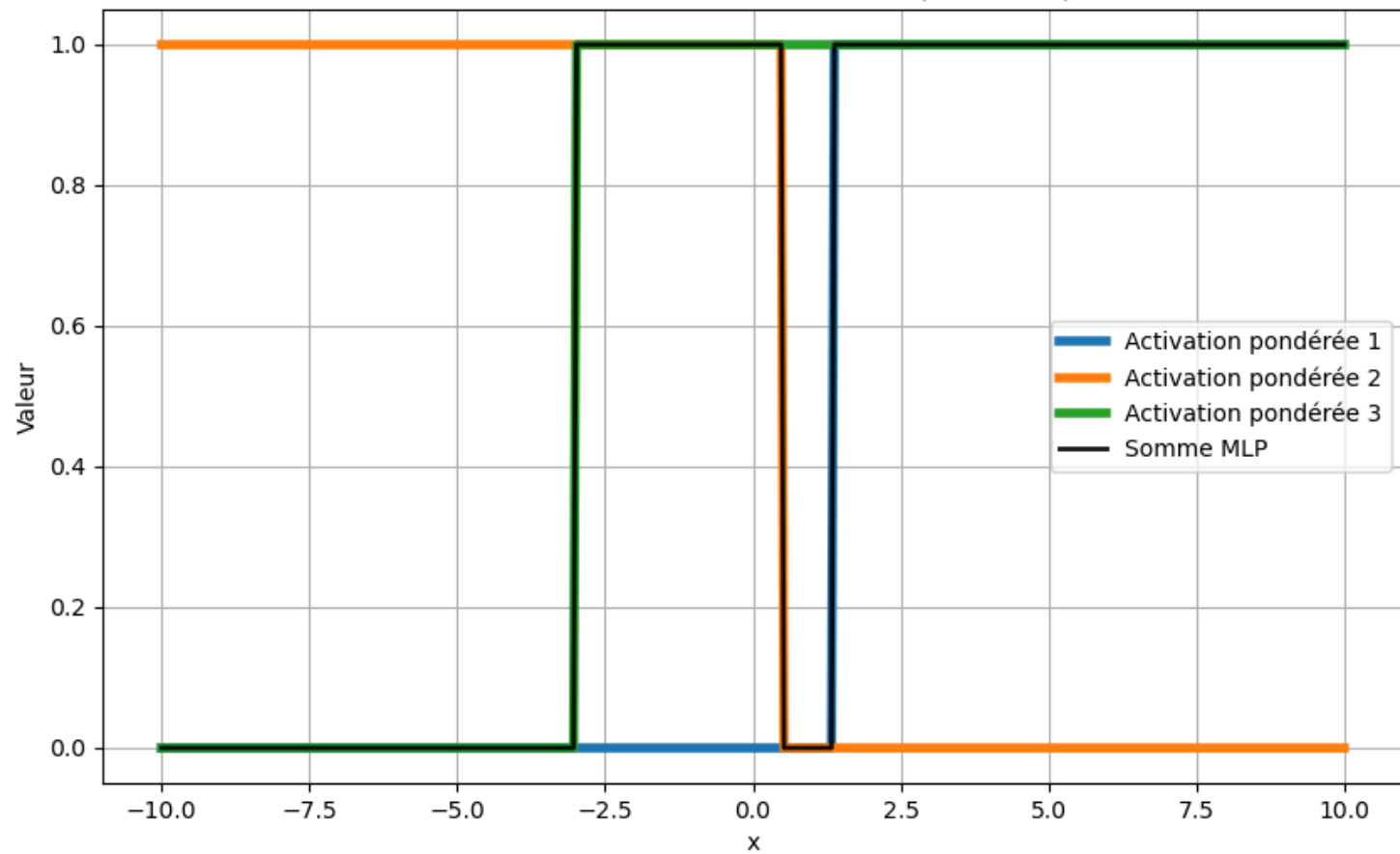
5 min

Somme des fonctions d'activation (identity)



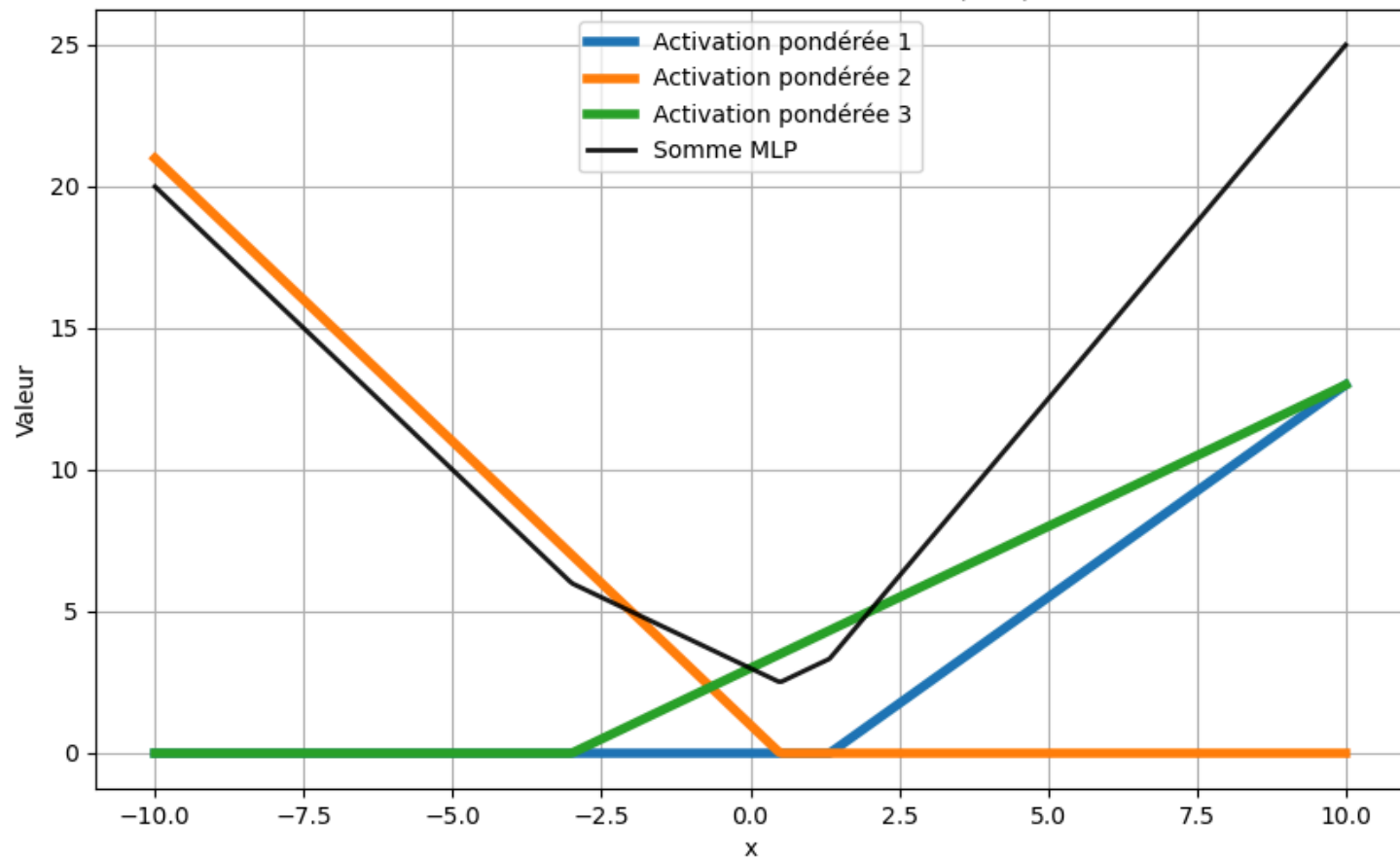
Supposon la fonction heaviside `1 if x > 0 else 0`

Somme des fonctions d'activation (heaviside)



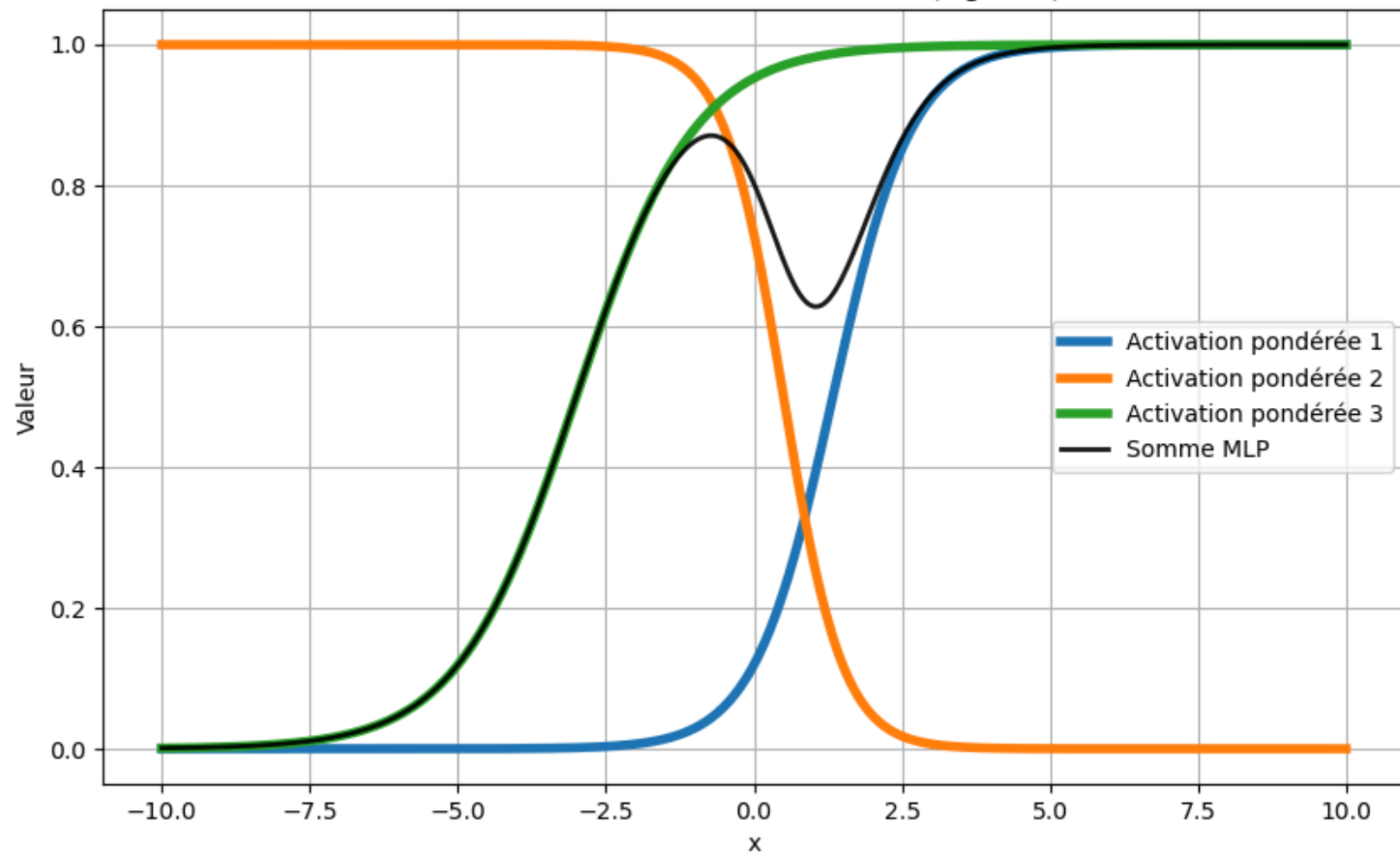
Supposon la fonction relu `x if x > 0 else 0`

Somme des fonctions d'activation (relu)



Supposon la fonction sigmoid $1/(1+\exp(-x))$

Somme des fonctions d'activation (sigmoid)





Tinker With a **Neural Network** in Your Browser.

Don't Worry, You Can't Break It. We Promise.



Epoch
000,000

Learning rate

0.03

Activation

Tanh

Regularization

None

Regularization rate

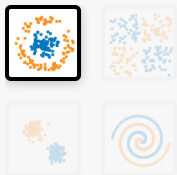
0

Problem type

Classification

DATA

Which dataset do you want to use?



Ratio of training to test data: 50%

FEATURES

Which properties do you want to feed in?



+

-

2 HIDDEN LAYERS

+

-

4 neurons

+

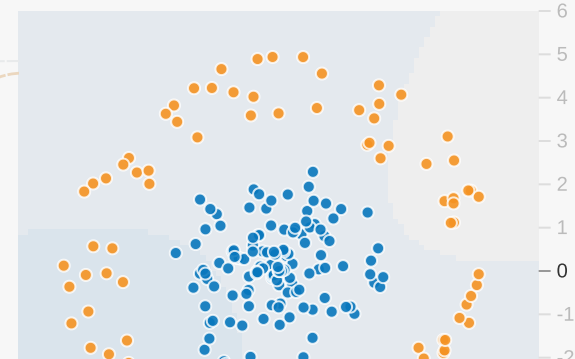
-

2 neurons

OUTPUT

Test loss 0.506

Training loss 0.503



The outputs are mixed with varying weights, shown by

Définition de l'IA

IA Faible (ANI)

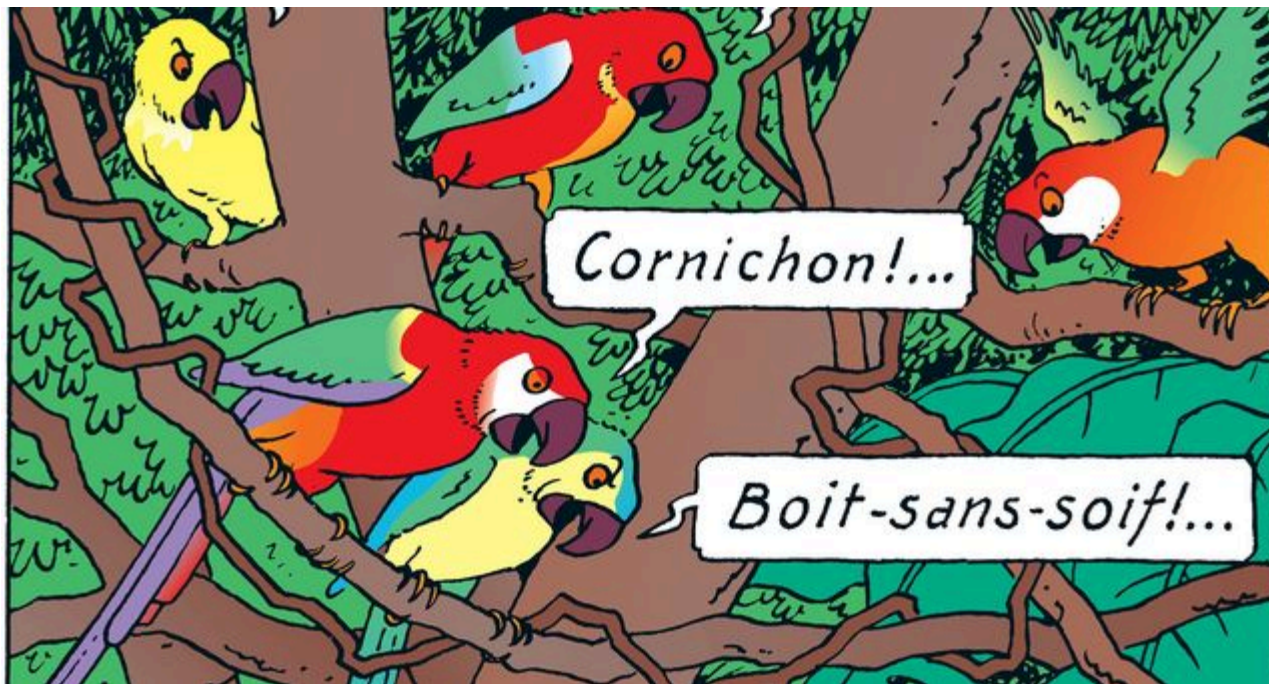
- Systèmes spécialisés, conçus pour accomplir des tâches précises
- Haute performance et précision dans des domaines spécifiques
- Exemples : reconnaissance faciale, traduction automatique, recommandation de contenu

IA Forte (AGI)

- Théorique : une intelligence artificielle capable de s'adapter à des tâches variées, comme un humain
- Capacité à raisonner de manière abstraite, à apprendre n'importe quelle tâche, mais sans preuve d'une conscience artificielle formelle.
- Pas encore réalisée à ce jour

Les modèles de langage de grande taille (LLM), comme GPT, se situent à mi-chemin entre l'IA faible et l'IA forte. Ils sont conçus pour traiter et générer du texte. Ils se trouvent à l'interface entre la compression d'information et l'utilisation de modèles probabilistes, extrayant des motifs à partir de vastes ensembles de données sans véritable compréhension sémantique du contenu. D'où des soucis d'allucinations.

L'IA : Un Perroquet Sophistiqué !



Applications Majeures

Vision par Ordinateur

- Reconnaissance d'images
- Détection d'objets
- Analyse vidéo
- Applications médicales

Industrie 4.0

- Maintenance prédictive
- Contrôle qualité
- Optimisation de production
- Robotique industrielle

Traitement du Langage

- Traduction automatique
- Analyse de sentiment
- Chatbots et assistants virtuels
- Génération de texte

Santé

- Diagnostic assisté
- Découverte de médicaments
- Analyse génomique
- Modélisation de maladies