Réseau de neurones artificiels

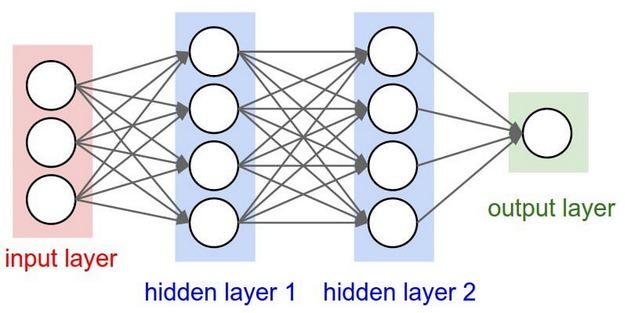
Grant Sanderson - But what is a Neural Network? | Deep learning, chapter 1 <https://www.youtube.com/watch?v=aircAruvnKk>

Andrew NG - What is a Neural Network? (C1W1L02) <https://www.youtube.com/watch?v=n1l-9lIMW7E>

**Réseau de neurones artificiels**

Un réseau de neurones artificiels est un système inspiré du fonctionnement du cerveau humain. Le concept de réseau de neurones artificiels a été inventé en 1943 par les neurologues Warren McCullough, et Walter Pitts.

En Deep Learning, le principe est de fournir un très grand nombre d’exemple à un réseau de neurones pour qu’il puisse s’adapter au problème par lui même.



Un réseau de neurones est composé d’un grand nombre d’éléments (neurones) interconnectés qui fonctionnent en parallèle et rangés par couches. La première couche, appelée Input Layer, reçoit les informations d’entrées. La couche de neurones de sortie est appelée l’Output Layer. Les couches intermédiaires sont appelées Hidden Layer. Le nombre de Hidden Layer est très variable, potentiellement de zéro à l’infini.

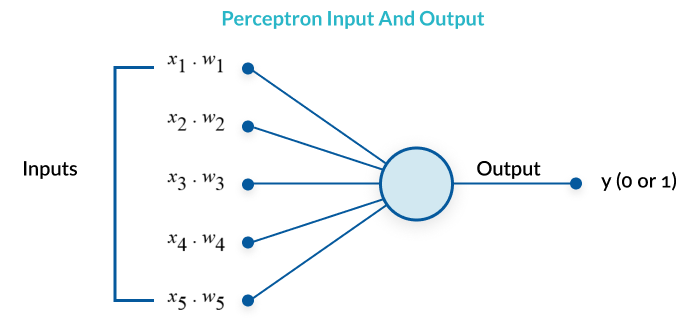
A la première entrée de donnée, le modèle risque de faire beaucoup d’erreur, car il aura essayé de prédire la bonne valeur au hasard. Le modèle va donc apprendre par lui-même en ajustant les poids, de telle sorte à ce que cela donne un meilleur résultat.

En comparant le résultat avec celui qu’on attendait de lui, on est capable de déterminer le coût. Plus ce dernier est élevé, plus le réseau est éloigné du résultat. Cet indice nous permet donc de savoir quel poids a le plus participé à l’erreur, et on pourra l’ajuster en fonction.

Andrej Karapathy - CS231n Winter 2016: Lecture 5: Neural Networks Part 2<https://www.youtube.com/watch?v=gYpoJMlgyXA> 6:04

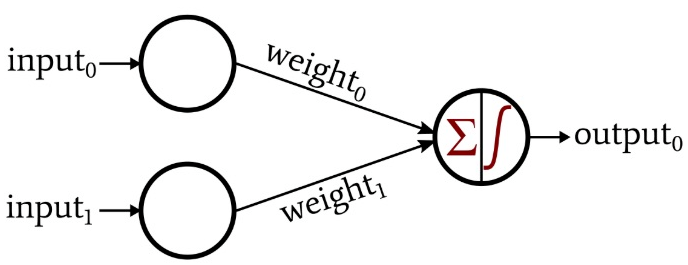
**Perceptron**

Un perceptron est un algorithme d'apprentissage supervisé de classifieurs binaires, [introduit par Franck Rosenblatt en 1957](http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.335.3398&rep=rep1&type=pdf), c’est le plus ancien algorithme de Machine Learning. Cet algorithme permet d’apprendre et résoudre des problèmes complexes. Il permet de diviser un ensemble de valeurs en entrée, en deux (“oui” ou “non”).

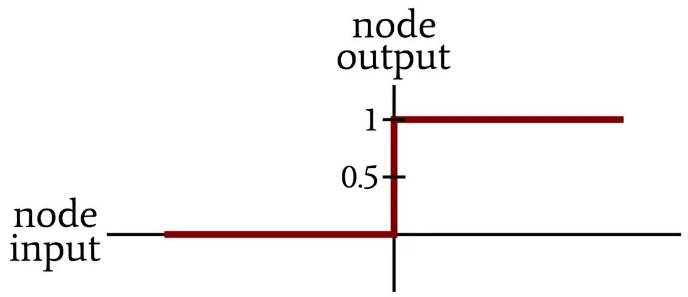


<https://missinglink.ai/guides/neural-network-concepts/perceptrons-and-multi-layer-perceptrons-the-artificial-neuron-at-the-core-of-deep-learning/>

Le perceptron aura plusieurs entrées qui auront chacune un poids selon leur importance, et une seule sortie binaire (0 ou 1).



Au niveau du noeud (aussi appelé neurone), les données qui arrivent seront multipliées par leur poids respectif, puis sommées pour ensuite les soumettre à une fonction “d’activation”. Celle-ci va générer une décision de classification. En appliquant une fonction de seuil, la valeur en sortie sera égale à 1 si la valeur en entrée est supérieure à la valeur de seuil, et 0 sinon.



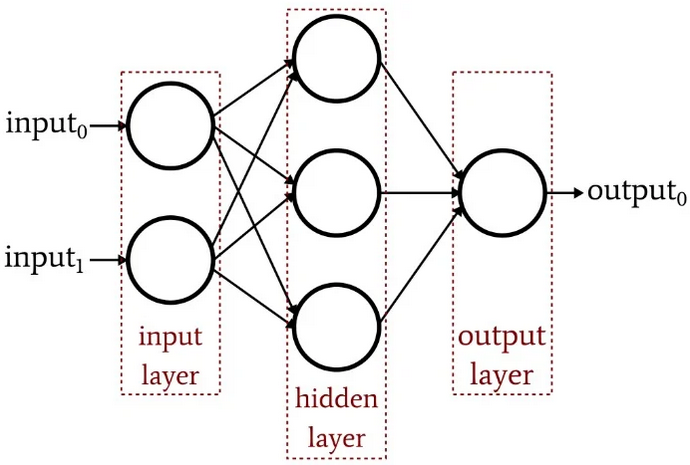
En utilisant cette transformation par le seuil, une non-linéarité est introduite au système.

Cependant un perceptron à une seule couche ne peut classifier des groupes que s’ils sont séparables de façon linéaire.

Par la suite une solution à été découverte : plusieurs perceptrons combinés forment alors un réseau de neurones artificiels, capable de résoudre toute sorte de problème, même non-linéaires.

**Multilayers Perceptron (MLP)**

Un perceptron multicouche (MLP) est un ensemble de perceptrons empilés formant plusieurs couches. Il a été introduit pour la première fois par Rumelhart en 1985, en même temps que Yann Le Cun sous une autre appellation. Ce réseau de neurones est capable de traiter des problèmes non-linéaires.



<https://www.allaboutcircuits.com/technical-articles/how-to-train-a-multilayer-perceptron-neural-network/>

Sur ce MLP à 3 couches, chaque perceptron de la couche d’entrée (input layer) enverra des sorties aux perceptrons de la deuxième couche (hidden layer), qui enverront à leur tour des sorties à la couche de sortie (output layer).

Chaque perceptron envoie plusieurs signaux, avec un signal à chaque perceptron de la couche suivante. Les perceptrons utilisent des poids différents pour chaque signal.

Comme dans cet exemple, un MLP à 3 couches s’appelle un *Non-Deep Neural Network*, et lorsqu’il aura 4 couches ou plus il portera le nom de *Deep Neural Network*.

Le MLP fonctionne alors avec une fonction de step qui aura une sortie binaire, contrairement à d’autres réseaux de neurones qui auront des valeurs réelles entre 0 et 1, ou entre -1 et 1 en sortie.