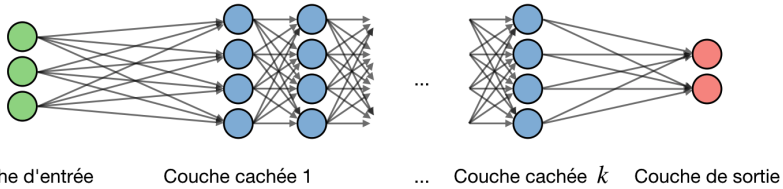


3 Apprentissage profond

3.1 Réseau de neurones

Les réseaux de neurones (en anglais *neural networks*) sont une classe de modèles qui sont construits à l'aide de couches de neurones. Les réseaux de neurones convolutionnels (en anglais *convolutional neural networks*) ainsi que les réseaux de neurones récurrents (en anglais *recurrent neural networks*) font parti des principaux types de réseaux de neurones.

□ **Architecture** – Le vocabulaire autour des architectures des réseaux de neurones est décrit dans la figure ci-dessous :



En notant i la $i^{\text{ème}}$ couche du réseau et j la $j^{\text{ième}}$ unité de couche cachée, on a :

$$z_j^{[i]} = w_j^{[i]T} x + b_j^{[i]}$$

où l'on note w, b, z le coefficient, le biais ainsi que la variable sortie respectivement.

□ **Fonction d'activation** – Les fonctions d'activation sont utilisées à la fin d'une unité de couche cachée pour introduire des complexités non linéaires au modèle. En voici les plus fréquentes :

Sigmoïde	Tanh	ReLU	Leaky ReLU
$g(z) = \frac{1}{1 + e^{-z}}$	$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	$g(z) = \max(0, z)$	$g(z) = \max(\epsilon z, z)$ with $\epsilon \ll 1$

□ **Cross-entropy loss** – Dans le contexte des réseaux de neurones, la fonction objectif de cross-entropie $L(z, y)$ est communément utilisée et est définie de la manière suivante :

$$L(z, y) = - \left[y \log(z) + (1 - y) \log(1 - z) \right]$$

□ **Taux d'apprentissage** – Le taux d'apprentissage (appelé en anglais *learning rate*), souvent noté α ou parfois η , indique la vitesse à laquelle les coefficients évoluent. Cette quantité peut être fixe ou variable. L'une des méthodes les plus populaires à l'heure actuelle s'appelle Adam, qui a un taux d'apprentissage qui s'adapte au file du temps.