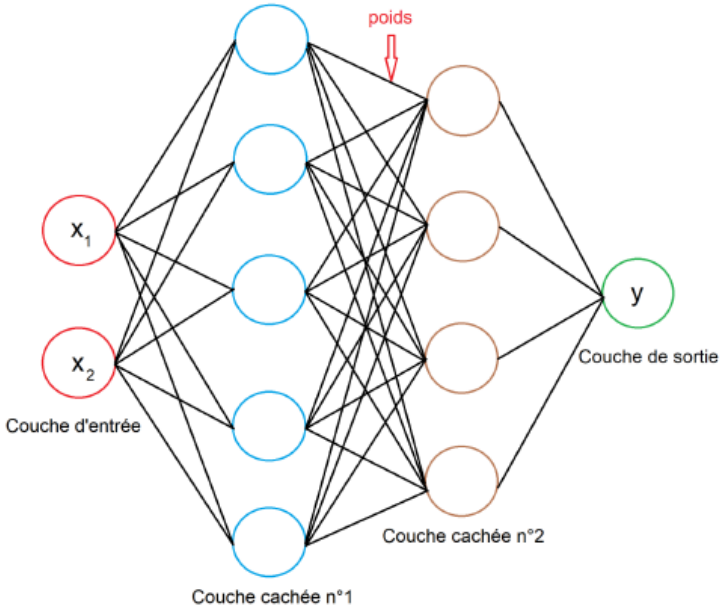
Une IA représentée par un réseau de neurones artificiels



Inspiré de ibm.com

Les réseaux de neurones artificiels sont constitués de couches : une couche d'entrée où entrent les informations (p. ex. une question à poser à l'IA), plusieurs couches cachées où les informations sont traitées et une couche de sortie où sortent les informations (la réponse). Les couches sont constituées de neurones artificiels (cercles) et de connexions (lignes). Les neurones sont des carrefours où se rejoignent les informations. Cette structure peut soit être représentée par un programme informatique, soit exister physiquement sous la forme de circuits électroniques.

Principe :

Imaginez une rivière qui est capable de passer par-dessus un barrage (un seuil) si la quantité d'eau est suffisamment grande (son poids).

Quand vous posez une question à l'IA, certaines neurones avec certaines connexions sont activées, d'autres pas. La rivière passe ou pas. En effet, chaque neurone se connecte à d'autres mais à certaines conditions que voici :

Pour qu'une information circule dans une connexion elle doit posséder une certaine importance, un certain **poids**. Un nombre représente ce poids. On pourrait penser que si le poids d'une information est grand, alors cette information passe d'un neurone à l'autre. C'est vrai à condition de passer un barrage qu'il y a dans le neurone : le **seuil**. C'est un nombre qui représente une limite en-dessous de laquelle l'information ne passe pas. C'est comme un filtre. Si une information arrive dans un neurone avec une valeur supérieure au seuil, alors cette information traverse le neurone (le barrage) et va activer un nœud de la couche suivante. Sinon, aucune donnée n'est transmise à la couche suivante. Ainsi, il ne suffit pas qu'une information soit importante pour traverser un neurone, encore faut-il que cette importance, représentée par un nombre, soit supérieure au seuil (barrage).

Exemple :

Cas simple avec **un seul neurone pour la couche cachée**. On pose cette question à l'IA : « **Dois-je faire du surf aujourd'hui ?** » On lui précise que trois facteurs influencent votre prise de décision :

1. Les vagues sont-elles bonnes ? (Oui donne 1, Non donne 0)
2. La route pour y aller est-elle libre, pas de trafic ? (Oui : 1, Non : 0)
3. Y a-t-il eu récemment une attaque de requin ? (Oui : 0 car mauvais pour mon surf, Non : 1)

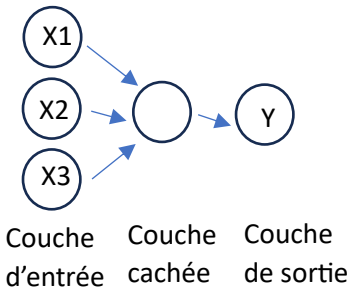
Donc les 1 sont bien pour aller surfer.

Supposons que vous pouvez répondre aux 3 questions ci-dessus (conditions initiales). Ce sont nos trois entrées :

X1 = 1, car les vagues sont bonnes

X2 = 0, car la foule est là, malheureusement

X3 = 1, car il n'y a pas eu d'attaque de requin récemment, tant mieux.



A présent, l'IA peut-elle décider si vous devez aller surfer ? Non, car vous n'avez pas encore défini l'importance de ces trois facteurs : vagues, route, requins.

Alors, attribuons des poids pour déterminer leur importance. Un poids important pour un facteur signifie que ce facteur a grande importance dans la décision.

Poids1 = 5, c'est un grand poids pour les vagues. Comme il n'y a pas des vagues tous les jours, pour une fois qu'il y en a, c'est important d'aller surfer.

Poids2 = 2, c'est un petit poids pour la foule car vous êtes habitué à la foule, pas grave s'il y aura du trafic.

Poids3 = 4, c'est un grand poids pour les requins car vous avez très peur des requins.

Et maintenant, l'IA peut-elle décider si vous devez aller surfer ? A ce stade oui, mais nous allons encore introduire votre seuil de tolérance par rapport à la réponse attendue. Si vous voulez diminuer au maximum les risques d'avoir un problème, alors vous n'êtes pas tolérant, votre seuil de tolérance est élevé (un grand nombre). Si cela ne vous dérange pas de prendre un certain risque cela signifie que votre seuil de tolérance est bas (un petit nombre).

Quel serait votre seuil de tolérance ? Observons les poids ci-dessus (grand poids : 5, petit poids : 2). Comparativement, disons que votre seuil de tolérance est de **3**. Vous êtes moyennement tolérant.
Maintenant, multiplions les entrées par leur poids et additionnons-les (combinaison linéaire) : Décision = (1*5) + (0*2) + (1*4) = 9 ce qui est plus grand que 0. Mais la valeur du neurone de sortie doit être soit 0 : non, je ne vais pas surfer, soit 1 : oui, je vais surfer. Comme 9 est plus grand que zéro, cela signifie que la valeur du neurone peut être ramené à 1. Donc, oui vous pourriez aller surfer.
Votre seuil de tolérance de 3, change-t-il ce résultat ? On peut le considérer comme un frein qui dissuade d'aller surfer donc comme un nombre à soustraire au résultat :
9 -3 = 6 donc plus grand que 0, vous allez toujours surfer. Mais si vous ne vouliez ne prendre aucun risque avec, disons, un seuil de 9, alors 9 – 9 = 0 pas de surf ! Votre seuil de tolérance est un **biais** sur le résultat.
On pourrait par la suite poser cette question plus complexe à l'IA : « Si je vais surfer, après, puis-je aller manger au restaurant Bora-Bora, sans réservation, avec des amis, dont je connais l'emploi du temps … »
Il nous faudrait alors prendre la décision *surf oui*, et la faire entrer dans une deuxième couche de neurones qui tient compte des contraintes, des poids et des seuils du restaurant, de l'emploi du temps des amis … C'est le même principe mais avec des calculs plus long.

<u>Du réseau de neurones à l'IA</u>
La plupart des Intelligences Artificielles reposent sur ce type de réseau de neurones, avec de très nombreuses couches et de très nombreux neurones d'entrée (par exemple, les valeurs des pixels d'une image).
Renseigner les poids et les seuils à la main serait alors impossible en un temps « humain ». C'est là que l' « apprentissage automatique » entre en jeu. A force d'essais-feedback-corrections, les poids et les seuls sont ajustés jusqu'à arriver à un résultat satisfaisant sur des données dites « d'entrainement ». S'il y a assez de seuils et poids (« paramètres ») et que leur ajustement est fait suffisamment longtemps, le réseau de neurones donnera des résultats qui restent pertinents avec de nouvelles données.