3. LES RÉSEAUX DE NEURONES

Résumé

Dans ce chapitre de présentation des réseaux de neurones formels, nous commençons par donner quelques définitions relatives aux réseaux de neurones non bouclés (ou statiques). En effet, c'est ce type de réseaux que nous avons utilisé dans le cadre de la résolution de problèmes de classification présentés dans ce mémoire. Nous présentons l'architecture de réseaux non bouclés la plus générale (les réseaux complètement connectés), puis une autre disposition dite à couches, notamment les réseaux à une seule couche cachée. Ensuite, nous justifions l'utilisation de cette dernière architecture en énonçant et commentant la propriété fondamentale de tels réseaux de neurones.

3.1 Définitions

Un réseau de neurones est une fonction paramétrée qui est la composition d'opérateurs mathématiques simples appelés *neurones formels* (ou plus simplement *neurones*) pour les distinguer des neurones biologiques. Dans ce paragraphe, nous présentons les définitions relatives aux neurones et les différentes architectures de réseaux de neurones.

3.1.1 Les neurones

Un neurone est une fonction algébrique non linéaire, paramétrée, à valeurs bornées, de variables réelles appelées *entrées*.

Par souci de commodité, on commet fréquemment un abus de langage en désignant par le vocable "neurone linéaire" une fonction linéaire (et plus généralement affine) qui n'est pas bornée.

On a pris l'habitude de représenter un neurone formel comme indiqué sur la figure 3.1.

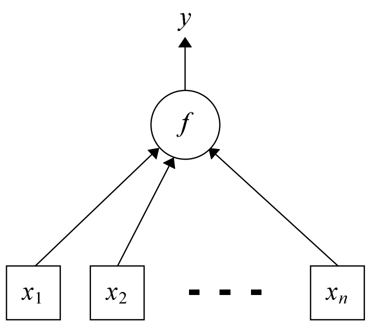

Figure 3.1 : Un neurone réalise une fonction non linéaire bornée $y = f(x_1, \ldots x_n ; c_1, \ldots c_p)$ où les $\{x_i\}$ sont les entrées et les $\{c_j\}$ sont des paramètres

Les paramètres dont dépend la valeur de y peuvent intervenir de deux manières :

- ils peuvent intervenir dans la fonction f elle-même,

- ils peuvent intervenir dans l'argument de la fonction f.

Les réseaux d'ondelettes ou de fonctions radiales entrent dans la première catégorie : les paramètres ajustables sont le centre et la dilatation (pour les ondelettes), ou le centre et la largeur (pour les fonctions radiales).

Dans ce travail, nous avons toujours utilisé des neurones (ou fonctions) qui appartiennent à la seconde catégorie : l'argument de la fonction f est une combinaison linéaire des entrées du neurone (à laquelle on ajoute un terme constant, le *"biais"*). La combinaison linéaire est appelée *potentiel* ; les coefficients de pondération $\{c_j\}$ sont fréquemment appelés *"poids synaptiques"* (ou plus simplement *poids*) en référence à l'origine "biologique" des réseaux de neurones.

Le potentiel d'un neurone est donc calculé de la façon suivante :

$$v = c_0 + \sum_{i=1}^{n} c_i x_i \quad \text{: potentiel du neurone}$$

Le biais c_0 peut être envisagé comme le coefficient de pondération de l'entrée n°0, qui prend toujours la valeur 1 :

$$v = \sum_{i=0}^{n} c_i x_i \quad \text{avec } x_0 = 1$$

La valeur de la sortie du neurone est donc :

$$y = f(v) = f\left(c_0 + \sum_{i=1}^{n} c_i x_i\right) \text{: sortie du neurone}$$

La fonction f est appelée *"fonction d'activation"* ; la fonction sigmoïde (ou tangente hyperbolique) est la plus utilisée :

$$y = \mathrm{th}(v)$$

Dans le présent mémoire, un neurone qui possède :

- une fonction d'activation sigmoïdale,
- et un potentiel défini par la somme pondérée des entrées,

est représenté comme indiqué sur la figure 3.2a :

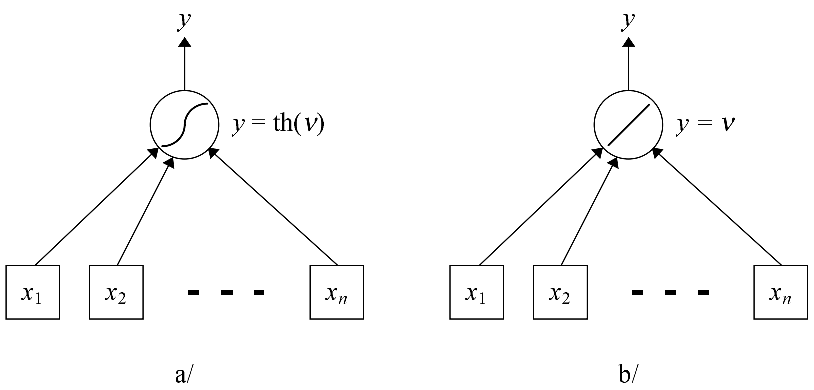

Figure 3.2 : Symboles de neurones à fonction d'activation sigmoïde et linéaire

La figure 3.2b représente un neurone linéaire.

3.1.2 Les réseaux de neurones non bouclés

Les fonctions non linéaires réalisées par les neurones décrits ci-dessus peuvent être combinées en un réseau de neurones. Dans un tel réseau, les entrées d'un neurone peuvent être soit les entrées du réseau, soit les sorties d'autres neurones.

Les valeurs des poids associés aux variables d'entrée des neurones sont en général déterminées par apprentissage (voir chapitre suivant) ; certaines d'entre elles peuvent être fixées à l'avance si une étude préalable du problème le recommande.

Il existe deux types d'architectures de réseaux de neurones :

- les réseaux non bouclés (ou statiques)
- les réseaux bouclés (ou dynamiques).

Les réseaux de neurones bouclés sont utilisés pour la modélisation dynamique de processus non linéaires et pour leur commande. Notre travail ne se situe pas dans ce domaine ; nous ne présenterons donc que la famille des réseaux de neurones non bouclés.

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonctions algébriques de ses entrées par composition des fonctions réalisées par chacun de ses neurones.

Dans un tel réseau, le flux de l'information circule des entrées vers les sorties sans "retour en arrière". Ainsi, si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les connexions entre ceux-ci, le graphe d'un réseau non bouclé est acyclique.

Tout neurone dont la sortie est une sortie du réseau est appelé "neurone de sortie". Les autres, qui effectuent des calculs intermédiaires, sont des "neurones cachés".

Nous présentons deux types de réseaux de neurones : les réseaux complètement connectés et les réseaux à couches. Le réseau de neurones à une couche cachée et une sortie linéaire est un cas particulier de ce dernier type.

3.1.2.1 Les réseaux de neurones complètement connectés

La figure 3.3 représente le réseau de neurones non bouclé le plus général possible : le réseau complètement connecté. Sur la figure, nous ne représentons pas les coefficients qui correspondent au "biais" (il suffit d'ajouter une entrée à valeur constante égale à 1 et portant le numéro 0).

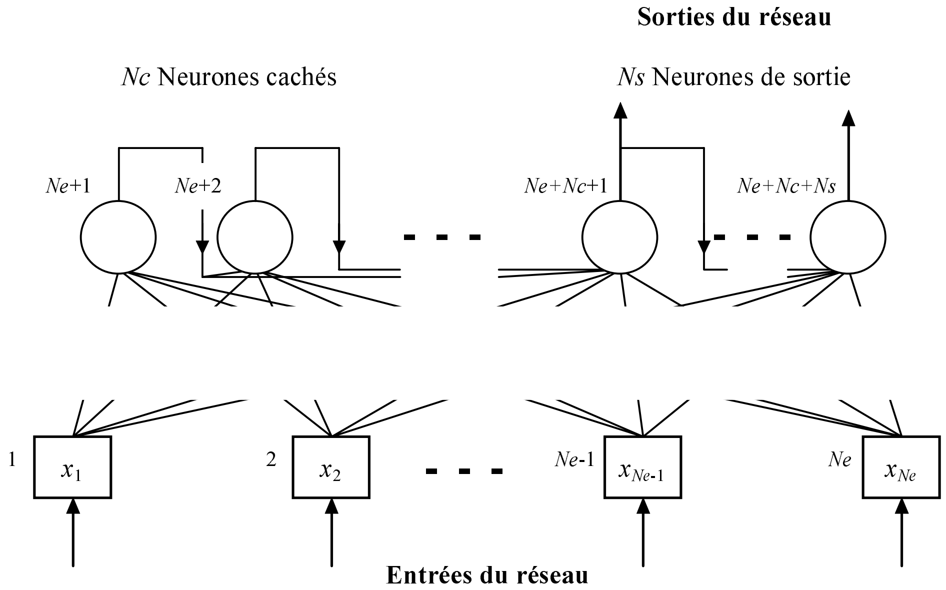

Figure 3.3 : Réseau de neurones non bouclé complètement connecté

Dans un réseau complètement connecté, les entrées puis les neurones (cachés et de sortie) sont numérotés (en italique sur le dessin), et, pour chaque neurone :

- ses entrées sont toutes les entrées du réseau ainsi que les sorties des neurones de numéro inférieur,
- sa sortie est connectée aux entrées de tous les neurones de numéro supérieur.

Un réseau de neurones non bouclé, complètement connecté, possède un nombre maximal de coefficients possible compte tenu du nombre de neurones qui le constituent, car les connexions de "retour en arrière" sont interdites.

3.1.2.2 Les réseaux de neurones à couches

Dans une architecture de réseaux à couches, les neurones cachés sont organisés en couches, les neurones d'une même couche n'étant pas connectés entre eux. De plus, les connexions entre deux couches de neurones non consécutives sont éliminées. La figure 3.4 représente un réseau à une couche cachée :

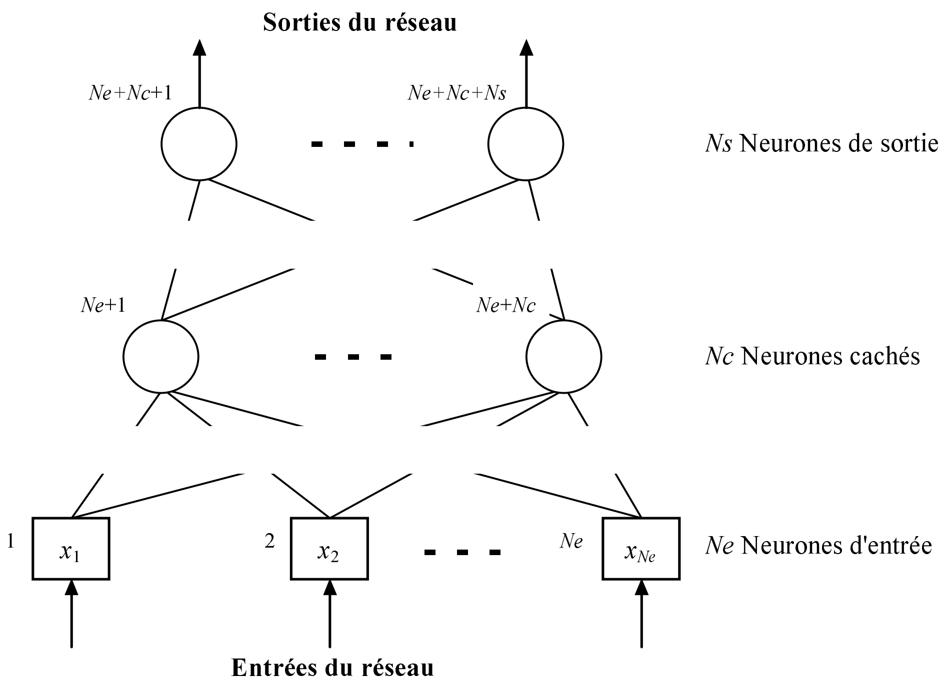

Figure 3.4 : Réseau de neurones non bouclé à une couche cachée

Les réseaux de neurones disposés suivant cette architecture sont aussi appelés *"perceptrons multicouche"* (ou MLP pour Multi-Layer Perceptrons). On trouve dans [Dreyfus 97] une perspective sur l'histoire et l'état de l'art des Perceptrons.

Une dernière architecture de réseau est très fréquemment utilisée, car elle possède des propriétés mathématiques intéressantes : les réseaux de neurones à une couche cachée et une sortie linéaire.

3.1.2.3 Les réseaux de neurones à une couche cachée et une sortie linéaire

La figure 3.5 représente un réseau de neurones à une couche cachée (N_c neurones cachés) et une sortie linéaire.

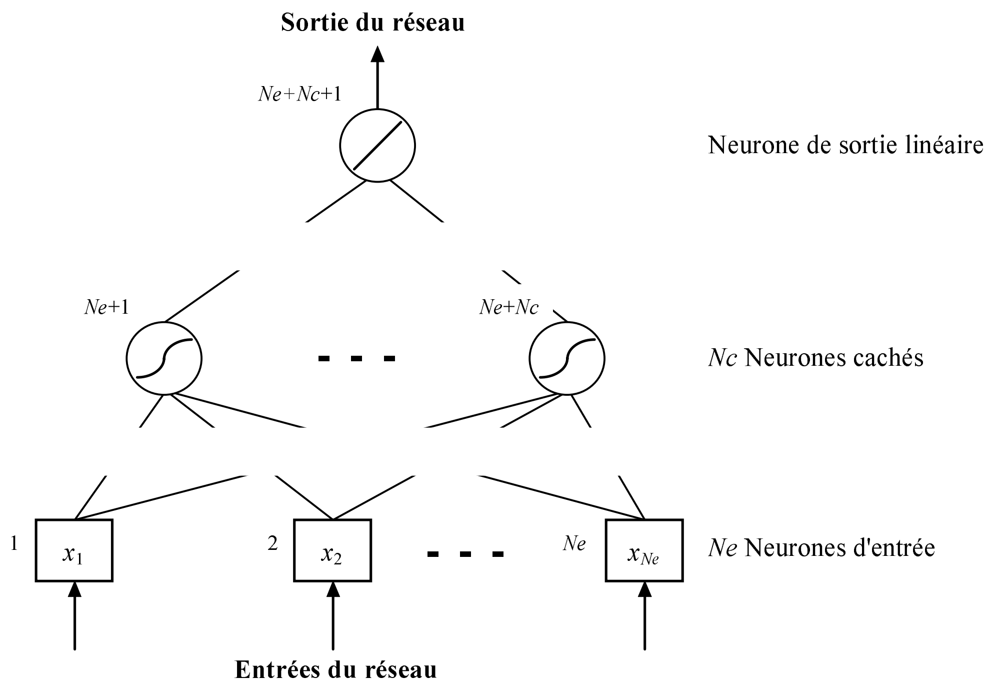

Figure 3.5 : Réseau de neurones non bouclé à une couche cachée et une sortie linéaire

Remarque : Pour des raisons que nous expliquerons au paragraphe 3.2, on utilise, *pour la classification*, un neurone de sortie dont la fonction d'activation, comprise entre 0 et 1, est définie par :

$$y = \frac{1 + th(v)}{2}$$

3.1.3 En résumé

Pour résoudre les problèmes de classification, nous utilisons la famille des réseaux de neurones non bouclés ; dans cette famille, nous avons choisi de mettre en œuvre des réseaux de neurones à une couche cachée. Nous justifions à présent ce choix en présentant la propriété fondamentale de ce type de réseaux.

3.2 Propriété fondamentale des réseaux de neurones

La propriété fondamentale des réseaux de neurones est *l'approximation parcimonieuse*. Cette expression traduit deux propriétés distinctes : d'une part, les réseaux de neurones sont des approximateurs universels, et, d'autre part, une approximation à l'aide de réseau de neurones nécessite, en général, moins de paramètres ajustables que les approximateurs usuels.

3.2.1 Les réseaux de neurones sont des approximateurs universels

La propriété d'approximation universelle [Cybenko 89 et Funahashi 89] peut s'énoncer de la façon suivante :

> Toute fonction bornée suffisamment régulière peut être approchée uniformément, avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini, possédant tous la même fonction d'activation, et un neurone de sortie linéaire.

Cette propriété est vraie pour les neurones présentés précédemment : neurones à fonction d'activation sigmoïdale, fonctions radiales, et ondelettes.

C'est cette propriété qui justifie notre choix de l'architecture de réseaux de neurones à une couche cachée. De plus, le seul degré de liberté qui subsiste pour la détermination de l'architecture du réseau est alors le nombre de neurones cachés, ce qui simplifie l'optimisation de l'architecture de réseaux, comme nous le verrons plus loin.

3.2.2 La parcimonie

Lorsque l'on veut modéliser un processus à partir des données, on cherche toujours à obtenir les résultats les plus satisfaisants possibles avec un nombre minimum de paramètres ajustables. Dans cette optique, [Hornik 94] a montré que :

> Si le résultat de l'approximation (c'est-à-dire la sortie du réseau de neurones) est une fonction non linéaire des paramètres ajustables, elle est plus parcimonieuse que si elle est une fonction linéaire de ces paramètres. De plus, pour des réseaux de neurones à fonction d'activation sigmoïdale, l'erreur commise dans l'approximation varie comme l'inverse du nombre de neurones cachés, et elle est indépendante du nombre de variables de la fonction à approcher. Par conséquent, pour une précision donnée, donc pour un nombre de neurones cachés donné, le nombre de paramètres du réseau est proportionnel au nombre de variables de la fonction à approcher.

Ces résultats s'appliquent aux réseaux de neurones à fonction d'activation sigmoïdale, puisque la sortie de ces neurones n'est pas linéaire par rapport à leurs coefficients. Ainsi, l'avantage des réseaux de neurones par rapport aux approximateurs usuels (tels que les polynômes) est d'autant plus sensible que le nombre de variables de la fonction à approcher est grand : pour des problèmes faisant intervenir une ou deux variables, on pourra utiliser indifféremment des réseaux de neurones, des polynômes, des réseaux d'ondelettes, etc. En revanche, pour des problèmes présentant trois variables ou plus, il est généralement avantageux d'utiliser des réseaux de neurones.

Bien entendu, cette propriété est démontrée de manière générale, et peut se révéler inexacte pour un problème particulier. Elle constitue néanmoins une justification fondamentale de l'utilisation des réseaux de neurones, et elle est avérée dans la très grande majorité des problèmes pratiques.

Rappelons que ces résultats concernent l'utilisation de réseaux de neurones pour l'approximation uniforme de fonctions connues ; il est pourtant rare que les réseaux de neurones soient mis en œuvre dans ce cadre. Nous allons montrer dans le paragraphe suivant que la technique des réseaux de neurones est généralement utilisée comme une méthode de *modélisation statistique*.

3.2.3 De l'approximation de fonction à la modélisation statistique

Les problèmes que l'on rencontre en pratique ne sont que très rarement des problèmes d'approximation de fonction *connue*. Dans la très grande majorité des cas, on cherche à établir un modèle à partir de mesures, ou, en d'autres termes, à trouver la fonction qui passe "au plus près" (en un sens qui sera précisé plus loin) d'un nombre fini de points expérimentaux, généralement entachés de bruit. On cherche alors à approcher la *fonction de régression* du processus considéré, c'est-à-dire la fonction que l'on obtiendrait en calculant la moyenne d'une infinité de mesures effectuées en chaque point du domaine de validité du modèle. Le nombre de points de ce domaine étant lui-même infini, la connaissance de la fonction de régression nécessiterait donc une infinité de mesure en un nombre infini de points.

Les réseaux de neurones, en raison de leur propriété fondamentale, sont de bons candidats pour réaliser une approximation de la fonction de régression. C'est ce qui justifie l'utilisation *réelle* des réseaux de neurones : la recherche d'une *approximation* de la fonction de régression à partir d'un nombre *fini* de points.

L'utilisation des réseaux de neurones entre donc complètement dans le cadre de méthodes statistiques : les méthodes de recherche d'une approximation de la fonction de régression. De telles méthodes ont été très largement développées pour les fonctions de régression *linéaires*. L'apport des réseaux de neurones réside donc dans leur capacité à approcher des fonctions *non linéaires*.

3.2.4 En résumé

En raison des propriétés fondamentales que nous venons de mentionner, les réseaux de neurones sont capables d'intervenir dans la résolution de nombreux problèmes de modélisation et de classification à partir de mesures. Ainsi, il peut être avantageux de les mettre en œuvre pour toute application nécessitant de trouver, par des méthodes statistiques, une relation non linéaire entre des données numériques.

Il va de soi que les méthodes "neuronales", ont des limitations, et que leur mise en œuvre nécessite quelques précautions de bon sens qui découlent directement des paragraphes précédents :

- Tout d'abord, nous avons vu que l'apport des réseaux de neurones réside dans leur capacité à réaliser des approximations de fonctions de régression non linéaires ; avant d'utiliser des réseaux de neurones dans une application, il faut donc s'assurer de la nécessité d'un modèle non linéaire. En effet, la mise en œuvre d'un modèle linéaire est généralement plus simple que celle d'un réseau de neurones.
- D'autre part, l'utilisation des réseaux de neurones (et plus généralement, des méthodes statistiques) nécessite un échantillon représentatif de la population

étudiée. Le chapitre 4 (Apprentissage des réseaux de neurones) montrera l'importance de la taille de l'échantillon.

Notons enfin qu'il peut arriver que d'autres approximateurs donnent, pour un problème particulier, de meilleurs résultats (résultats plus précis avec le même nombre de paramètres ajustables, ou résultats aussi précis avec moins de paramètres ajustables) que les réseaux de neurones. Il est donc toujours possible, si l'on dispose de temps pour cela, de tester ces approximateurs.

3.3 Conclusion

La propriété d'approximation parcimonieuse fait des réseaux de neurones d'excellents outils pour la résolution des problèmes de modélisation et de classification. Pour obtenir de bons résultats, il faut s'assurer d'avoir bien posé le problème (voir les deux conditions du paragraphe précédent). De multiples expériences, dont celles qui sont décrites dans le présent mémoire, montrent que, si l'on pose bien le problème, les réseaux de neurones fournissent toujours d'excellentes solutions.