

Chapitre II : Réseaux de Neurones Artificielles et Perceptron Multicouches

1. Introduction

Les réseaux de neurones sont utilisés pour résoudre des problèmes complexes en optimisation non linéaire ainsi que de classifications. Les Techniques utilisées sont basées sur le fonctionnement des neurones du cerveau, d'où l'inspiration biologique.

1.1. Inspiration biologique

Le neurone du cerveau peut être considéré comme un automate comprenant :

- la cellule centrale (ou soma) contenant le noyau. Il fait quelques microns de diamètre il effectue les transformations biochimiques nécessaires à la synthèse des enzymes et des autres molécules qui assure la vie du neurone.
- les dendrites recevant les signaux d'autres neurones par l'intermédiaire des synapses. Celles-ci sont de quelques dixièmes de microns de diamètre et d'une longueur de dizaines de microns
- l'axone qui achemine l'information, en sortie de la cellule vers les autres neurones. Il est généralement plus long que les dendrites, il communique avec d'autres neurones par son extrémité. Les connexions entre deux neurones se font en des endroits appelés synapses ou ils sont séparés par un petit espace synaptique de l'ordre d'un centième de microns.
- L'ensemble des signaux d'entrée (excité et inhibé) est moyenné. Si cette moyenne est suffisamment importante sur un laps de temps très court, la cellule délivre un signal électrique aux cellules suivantes par l'intermédiaire de son axone.
- Les impulsions délivrées (ou potentiel d'action) sont du type tout ou rien, c.a.d., que les neurones communiquent entre eux en langage binaire.

Une interprétation "mathématique" du neurone biologique peut être donnée par la Figure 1.1. Le neurone reçoit un nombre variable d'entrées en provenance de neurones amonts. A chacune de ces entrées est associée un poids w représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals.

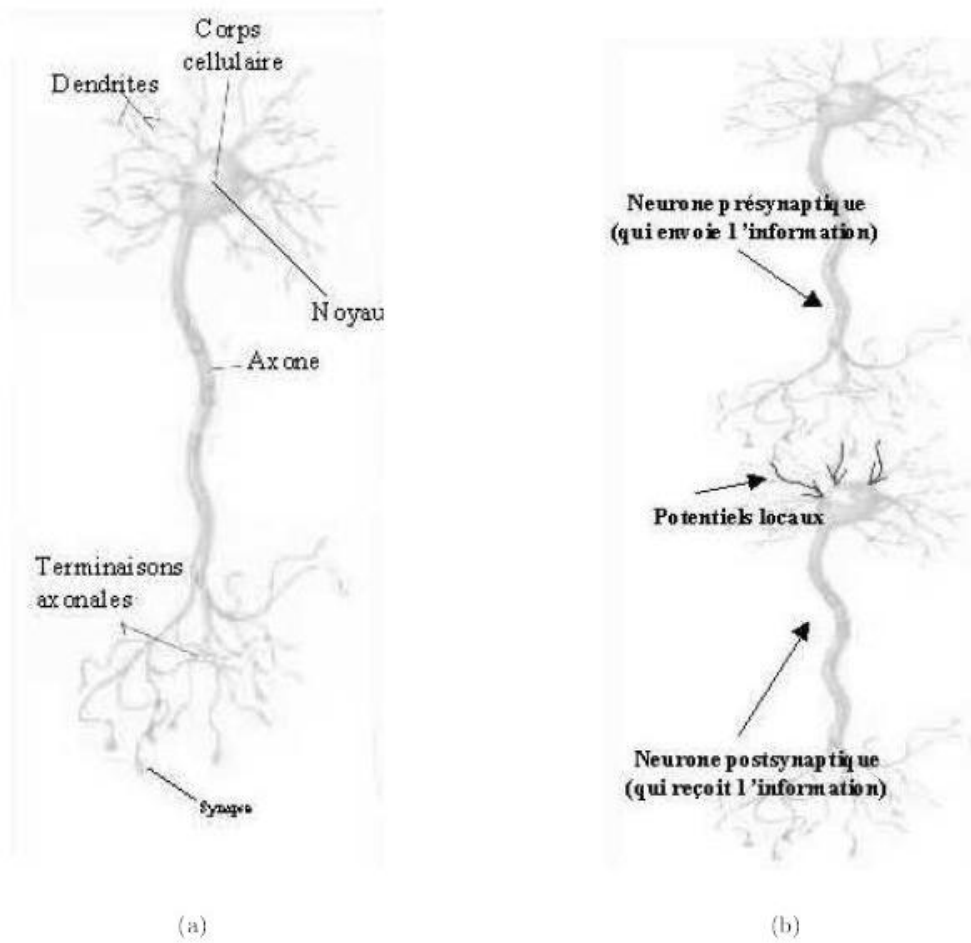


Figure 1.1 – Neurones biologiques

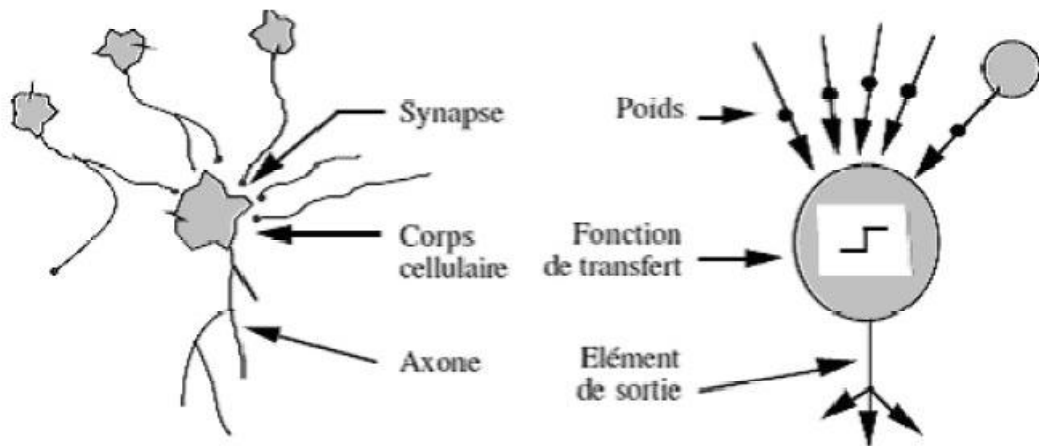


Figure 1.2 – Correspondance entre neurone biologique et neurone artificiel

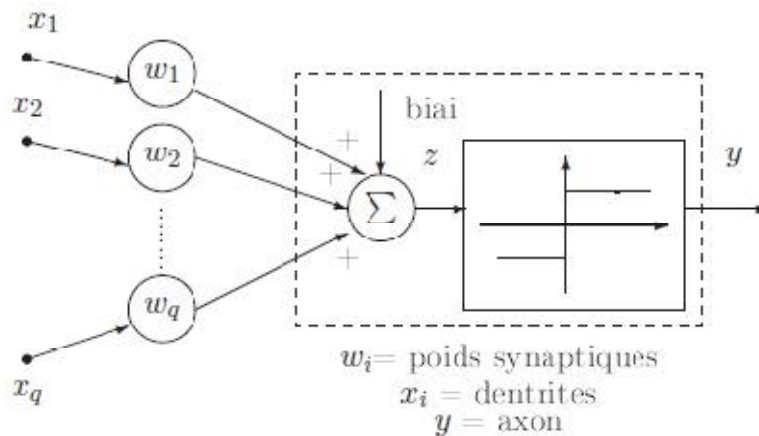


Figure 1.3 – Neurone artificiel

Les caractéristiques des neurones biologiques sont données par:

- Un cerveau humain contient en moyenne entre 10^{10} et 10^{12} neurones.
- Les neurones sont interconnectés les un aux autres dans une forme spatiale très complexe pour former le système nerveux.
- Le noyau a un diamètre de quelques microns.
- Les dendrites ont une longueur dépassant rarement les 10 microns et reçoivent les signaux d'entrées.
- L'axone forme la sortie du neurone, il a une longueur comprise entre 1mm et peut atteindre 1m de long. Il peut connecter le neurone à un ou plusieurs autres neurones.
- Un neurone peut être connecté à une centaine voire plusieurs milliers d'autres neurones.
- Les signaux sont de nature électrochimique.
- La vitesse de propagation varie de 5 à 125 m/s
- Un délai de 1 ms est nécessaire pour que le signal traverse les synapses.
- Il est observé que grâce à une activité cérébrale intense, un développement métabolique s'effectue au niveau synaptique, ceci engendre un élargissement de la surface synaptique, et donc une augmentation des poids.
- Si la somme des signaux reçus par un neurone excède une certaine limite, ce dernier est activé.

- Les neurones peuvent être active (fire) sur une large bande fréquentielle mais toujours a la même amplitude.
- Après être activé (fire) un axons est hors service pour une période de 10ms.
- Les information sont codées en fréquence pour les signaux émis.

1.2. Historique des RNA

ORIGINES

1890 : W.James , célèbre psychologue américain introduit le concept de mémoire associative, et propose ce qui deviendra une loi de fonctionnement pour l'apprentissage sur les RNA connue plus tard sous le nom de loi de Hebb. 1943 : J.Mc.Culloch et W.Pitts sont les premiers à montrer que des RNA formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes.

1949 : D.Hebb physiologiste américain explique le conditionnement chez l'animal par les propri étés des neurones eux-même. Ainsi un conditionnement de type pavlovien tel que nourrir tous les jours à la même heure un chien, entraîne chez cet animal la sécrétion de salive à cette heure précise même en l'absence de nourriture. La loi de modi_cation des propriétés des connexions entre neurones qu'il propose explique en partie ce type de résultats expérimentaux.

1957 : F.Rosenblatt développe le modèle du perceptron. Il construit le premier neuro ordinateur basé sur ce modèle et l'applique au domaine de la reconnaissance de formes.

1960 : B .Widrow un automaticien, développe le modèle Adaline (Adaptative LINear Element) dans sa structure, le modèle ressemble au perceptron. Cependant la loi d'apprentissage est différente. B.Widrow a créé dès cette époque la " Memistor Corporation ". il est aujourd'hui le président de l'International Neural Network Society (INNS).

LES ANNEES CALMES

1969 : M.Minsky et S.Papert publient un ouvrage qui met en exergue les limitations théoriques du perceptron. Limitation alors connue notamment concernant l'impossibilité de traiter par ce modèle des problèmes non linéaires.

1967-1982 : Toutes les recherches se poursuivent, mais déguisées sous le couvert de divers domaines. De grands noms travaillent durant cette période tels S.Grossberg, T.Kohonen. Albus devellopa son adaptatif "Cerbral Model Articulation Controller", qui se résume a une sorte de table de correspondances, imitant sa perception de la mémoire humaine (Albus 1972).

ENTHOUSIAME RETROUVE

1982 : J.J.Hopffield physicien reconnu à qui on doit le renouveau d'intérêt pour les RNA. Notons qu'à cette date, l'IA est l'objet d'une certaine désillusion, elle n'a pas répondu à toutes les attentes et s'est même heurtée à de sérieuses limitations. Aussi, bien que les limitations du perceptron mises en avant par M.Minsky ne soient pas levées par le modèle d'Hopffield.

1983 : la machine de Boltzmann est le premier modèle connu à traiter de manière satisfaisante les limitations recensées dans le cas du perceptron.

1985 : la rétro propagation du gradient apparaît. Elle a été découverte par trois groupes de chercheurs indépendants. De nos jours, le réseau multicouches et la rétro propagation du gradient reste le modèle le plus étudié et le plus productif au niveau d'applications [12]. Une architecture similaire au MLP, est le Probabilistic Neural Network (PNN), qui diffère de ce dernier par sa fonction d'activation exponentiel et les formes de ses connections. Durant cette période plusieurs idées originales tel que : les réseaux ART ("Adaptative Résonance Théorie") sont des réseaux à apprentissage par compétition. Le problème majeur qui se pose dans ce type de réseaux est le dilemme " stabilité / plasticité ".

1987 : La confirmation des réseaux de Hopfield comme les réseaux à mémoires associative, très utiles pour les problèmes de classification (clustering). Plus tard, Kosko étendra ce concept pour développer son "Bi-directional Associative Memory" (BAM) (Kosko 1985).

1988 : Broomhead et Lowe (1988) s'inspirant des fonctions radiales, développèrent les réseaux RBF (Radial Basis Function). Il ouvrirent la voie à Chen et Billings (1992) et leurs Functional Link Network (FLN) ou une transformation nonlinéaire des entrées permet une réduction de la puissance de calcul ainsi qu'une convergence plus rapide.

1997 : Implantation des RNA sous silicium (VLSI) ce qui lève beaucoup de limitation de vitesse et offre des possibilités pas encore envisagées jusque là.

1.3. Types des RNA

Les quatre types de RNA les plus utilisés sont énumérés ci-dessous. A noter que chacun de ses réseaux fera l'objet d'une étude plus approfondie dans les chapitres suivants.

1. Perceptron à une ou plusieurs couches (PMC), ou le single or Multi-Layered Perceptrons (MLP)
 - Le type de réseaux le plus utilisé
 - Bon pour la classification, mais très bon pour l'approximation de fonctions
 - Bénéficie d'un nombre important d'applications documentées et de packages commercialisés.
2. Réseaux à base de fonction radial ou Radial-basis function (RBF) Approximateurs locaux de fonctions.
 - Réseau à une seule couche
 - Apprentissage facile ?
3. Réseaux de Hopfield
 - Réseaux à couche unique, avec un retour des sorties sur les entrées.
 - Très liée aux mémoires associatifs
 - Forme continue et discrète
4. Les cartes de Kohonen ou Kohonen Self-organising Feature Maps (SOFM).

- Modèle simplifié de la notion de caractéristiques liés a une région donnée du cerveau. Apprentissage par compétition Principalement utilisé pour la classification.
- Un réseau similaire au SOFM est le CNN (Cellular Neural Network), qui est une sorte tableau comportant des systemes dynamiques, appelés cellules, qui sont connectées localement.
- Chaque cellule n'inter acte qu'avec les cellules adjacentes.

Le tableau suivant résume, en général, la correspondance entre le type de réseaux et les applications les plus appropriées à ce dernier. Une étude plus détaillée et plus compléter sur les applications des RNA en milieu industriel peut être trouvée dans (Magali et al., 2003).

Caractéristiques fonctionnelles	Type de RNA
Reconnaissance de formes	MLP, Hopefield, Kohonen, PNN
Mémoires associatives	Hopefield, MLP récurrents, Kohonen
Aproximation de fonctions	MLP, RBF
Optimisation	Hopefield, ART, CNN
Classi_cation et clustering	MLP, Kohonen, RBF, ART, PNN
Mémoires associatives	Hope_eld, MLP recurents, Kohonen
Reconnaissance de formes	MLP, Hopefield, Kohonen, PNN
Modélisation et control	MLP, MLP recurent, FLN
Traitement d'images	CNN, Hopefield

1.4. Perspectives avec les RNA

Les RNA sont :

- Une copie artificielle des réseaux de neurones biologique.
- Une correspondance non linéaire d'entrées et de sorties.
- Un système adaptatif non linéaire.
- Des approximateurs universels.

Les RNA peuvent :

- Implémenter n'importe quelle fonction non linéaire (Théoreme de Cybenko, 1980) jusqu'à un certain degré de fiabilité.
- Implémente des fonctions dynamiques.
- Avoir n'importe quel nombre d'entrées et de sortie.
- Les différent type de RNA sont distinguées par :
- La topologie du réseau.
- Le nœud, ou le type de neurone.
- La règle d'apprentissage associé au réseau.

1.5. Apprentissage pour les RNA

On distingue deux approches d'apprentissage, Figure 1.4, l'apprentissage supervisé et l'apprentissage non-supervisé.

1.5.1 Approche dite supervisée

On dispose d'observateur, c'est-à-dire un ensemble de données connues (les exemples) :

en discrimination, des exemples de données provenant de chaque groupe (classe) :
des mesures biologiques (tension artérielle, numération sanguine, présence/absence de symptômes divers, etc.) pour des personnes malades et pour des personnes saines
des lettres et chiffres tracés par différentes personnes en régression,
des exemples de la relation fonctionnelle :
le cours d'une action sur une période donnée
la consommation électrique d'une région sur une période donnée, associée à des mesures météorologiques sur la même période point commun : des observations et une cible (la classe, le cours de l'action dans le futur, etc.)

Le but principal est de modéliser la relation entre les observations et l'information cible, pour :

1. Estimer la valeur de la cible pour de nouvelles observations :

- diagnostiquer un patient
- reconnaître un code postal
- prédire le cours d'une action
- prédire la consommation électrique d'une région

2. Comprendre la relation observations/cibles :

- Déterminer les symptômes importants
- Comprendre les facteurs déterminants d'une forte consommation électrique
- Les réseaux de neurones sont très efficaces pour le premier point, moins pour le second.

1.5.2 Approche dite non supervisée

- On ne dispose pas d'observateur, c'est-à-dire que les exemples ne sont pas déjà organisés :
 - pas de classe, pas de cible
 - but général : découvrir de la régularité, sous forme de classes, c'est-à-dire de groupes d'exemples qui se ressemblent
- but opérationnel :
 - groupes de consommateurs pour cibler une campagne marketing
 - groupes d'individus pour l'analyse sociologique, économique, etc.
- motivations :
 - comprendre les données en les simplifiant (analyse sur chaque groupe) décrire des groupes

1.6 Avantages et inconvénients des RNA

Les avantages :

Une tolérance aux incertitudes très élevée.

- Etant une multiple copie d'une unité simple (le neurone), donc facilement extensible.
- Facilité d'application car ne nécessitant pas une compréhension approfondie.
- Un choix de réseaux, types, architecture et fonctions d'activations diverses.
- Possède la capacité de généraliser (interpoler) dans le cadre des données utilisés pour l'apprentissage.

Les inconvénients :

- Facilité d'applications donnant lieu a de nombreuses implémentations et choix pas toujours Justifiés.
- Malgré une solide base théorique, le choix du réseau incombe souvent a l'utilisateur car il n'existe pas de guide prouvé pour toute utilisation.
- La nature nonlineaire des RNAs peut piéger l'utilisateur dans un minimum local.

1.7 Applications des RNA

Les RNAs ont trouvé applications dans une multitude de domaines, aussi divers que différents. Un aperçu des domaines ou les RNA ont été d'un grand apport sont :

1. Signal processing
 - Egalisation non linéaire
 - L'élimination de bruit et de l'écho.
 - Reconnaissance de signaux radar ou sonar
2. Automatique
 - Identification/modélisation des systèmes non linéaire.
 - Commande basé modèle (commande prédictive, IMC, etc).
3. Compression de données
4. Systèmes de classifications ou de diagnostics
5. Mémoires adressable

2. Neurone simple et problème de classification

Dans ce chapitre, nous allons, tout d'abord, présenter les types de neurones simples ainsi que quelques règles d'apprentissage de base, soutenus par des exemples. Une fois cette introduction effectuée, un aperçu sur la classification en utilisant les neurones présentés sera introduite avec quelques exemples de modélisation des fonctions logiques AND et OR.

2.1 Type de neurones simples

2.1.1 Le neurone de Mc Culloch-Pitt

Le neurone de Mc Culloch-Pitt est donné dans la Figure 2.1, il se caractérise principalement par une fonction d'activation de type limitation 0 ou 1 ($s=1$ ou 0).