

## Introduction

L'industrie moderne a besoin d'obtenir les meilleurs résultats pour la commande des systèmes travaillant toujours aux limites des contraintes. Les scientifiques proposent plusieurs méthodes de commande comme éventuelle solution, l'une de ces méthodes est l'utilisation de la technique de réseaux de neurones dans le processus de contrôle, soit dans la conception du modèle soit dans la conception du contrôleur.[21]

Nous allons présenter dans ce chapitre les principes de base des réseaux de neurones artificiels et leur apprentissage ainsi que ses utilisations pour la modélisation et pour la prédiction de la tension de contournement des isolateurs.

### II.1 Historique des RNA

L'origine de l'inspiration des réseaux de neurones artificiels remonte à 1890 où W. James, célèbre psychologue américain, introduit le concept de mémoire associative. Il propose ce qui deviendra une loi de fonctionnement pour l'apprentissage des réseaux de neurones, connue plus tard sous le nom de loi de Hebb. Quelques années plus tard, en 1949, J. Mc Culloch et W. Pitts[1] donnent leurs noms à une modélisation du neurone biologique (un neurone automatique comportement binaire). Ce sont les premiers à montrer que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes. C'est ensuite que D. Hebb, physiologiste américain, présente en 1949 les propriétés des neurones par le conditionnement chez l'animal. 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 modification des propriétés des connexions entre neurones qu'il propose, explique en partie ce type de résultats expérimentaux. Les premiers succès de cette discipline remontent à 1957, lorsque 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 des formes. Notons qu'à cette époque les moyens à sa disposition étaient limités et c'était une prouesse technologique que de réussir à faire fonctionner correctement cette machine plus de quelques minutes. C'est alors qu'en 1960, l'automaticien Widrow développe le modèle Adaline (Adaptative Linear Element). Dans sa structure, le modèle ressemble au Perceptron, cependant la loi d'apprentissage est différente. Celle-ci est à l'origine de l'algorithme de rétropropagation de gradient très utilisé aujourd'hui. en 1969 : M.Minsky et S.Papert publient un ouvrage qui met en exergue les limitations théoriques avec les Perceptrons Multi Couches. M. Minsky et S. Papert publient ensuite en 1969 un ouvrage qui met en évidence les limitations théoriques du Perceptron. Ces limitations concernent l'impossibilité de traiter des problèmes non

linéaires en utilisant ce modèle. Quelques années d'ombre se sont ensuite succédées de 1967 à 1982.

Le renouveau de cette discipline reprend en 1982 grâce à J. J. Hopfield, un physicien reconnu. Il présente une théorie du fonctionnement et des possibilités des réseaux de neurones. Il faut remarquer la présentation anticonformiste de son article. Alors que les auteurs s'acharnent jusqu'alors à proposer une structure et une loi d'apprentissage, puis à étudier les propriétés émergentes, J. J. Hopfield fixe préalablement le comportement à atteindre par son modèle et construit, à partir de là la structure et la loi d'apprentissage correspondant au résultat escompté. Ce modèle est aujourd'hui encore très utilisé pour des problèmes d'optimisation. On peut citer encore la machine de Boltzmann en 1983 qui était le premier modèle connu, apte à traiter de manière satisfaisante les limitations recensées dans le cas du Perceptron. Mais l'utilisation pratique s'avère difficile, la convergence de l'algorithme étant extrêmement longue (les temps de calcul sont considérables). C'est ensuite qu'en 1985 la rétro-propagation de gradient apparaît. C'est un algorithme d'apprentissage adapté au Perceptron Multi Couches. Sa découverte est réalisée par trois groupes de chercheurs indépendants. Dès cette découverte, nous avons la possibilité de réaliser une fonction non linéaire d'entrée/sortie sur un réseau, en décomposant cette fonction en une suite d'étapes linéairement séparables. Enfin, en 1989 Moody et Darken exploitent quelques résultats de l'interpolation multi variables pour proposer le Réseau à Fonctions de base Radiales (*RFR*), connu sous l'appellation anglophone *Radial Basis Function network (RBF)*. Ce type de réseau se distingue des autres types de réseaux de neurones par sa représentation locale.[22]

## II.2 Modèle Biologique

Les neurones, au nombre d'une centaine de milliards, sont les cellules de base du système nerveux central. Chaque neurone reçoit des influx nerveux à travers ses dendrites (récepteurs), les intègre pour en former un nouvel influx nerveux qu'il transmet à un neurone voisin par le biais de son axone (émetteur) [23] comme le montre la Fig II.1 :

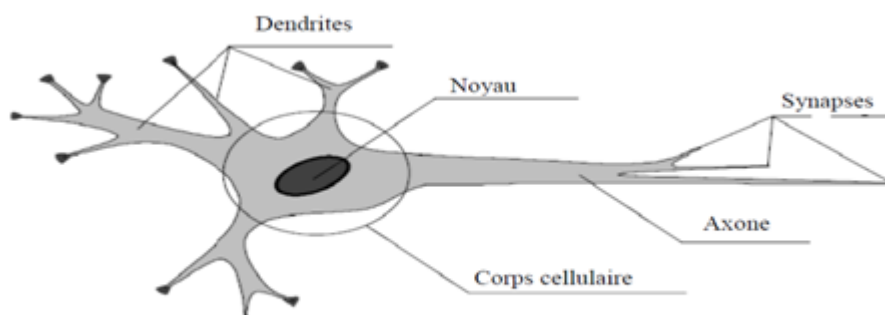


Figure II.1 : Neurone biologique

A partir de ces concepts on va donner une description des fondamentales du réseau de neurones.

### **II.2.1 Neurone**

Est une cellule vivante, qui peut prendre des formes variables (pyramidale, sphérique ou étoilée) sa forme est définie par une membrane qui sépare l'intérieur du neurone à l'extérieur. Ces neurones sont considérés comme l'élément de base qui constituent les unités élémentaires de traitement dans le cerveau. [24]

### **II.2.2 Structure de neurone**

#### **1- corps cellulaire (soma)**

Il est composé d'un noyau qui effectue les transformations biochimiques essentielles à la vie de neurone et se ramifie pour former les dendrites, qui établit la liaison avec d'autres cellules.

#### **2-Axone**

C'est le prolongement unique, qui diffuse le signal du neurone vers d'autres cellules donc son rôle est représenté dans la communication avec d'autre neurone. Il peut diviser à son extrémité pour entrer en contact avec un grand nombre d'autre cellules.

#### **3-Synapse**

C'est un élément de jonction qui assure le contact du cytoplasme (membrane) d'un neurone et les membranes de ses voisins, il joue un rôle essentiel dans la transmission des signaux [25].

### **II.3 Définition de réseaux de neurones artificiels**

Un réseau de neurones (RN) est un système d'opérateurs non linéaires interconnectés, recevant des signaux de l'extérieur par ses entrées, et délivrant des signaux de sortie, ces réseaux de neurones sont une métaphore des structures cérébrales, de traitement parallèle, de distribution d'information et comportent plusieurs éléments de traitement appelé neurone.

Chaque neurone fonctionne indépendamment des autres de telle sorte que l'ensemble est un système parallèle fortement interconnecté. L'information détenue par le réseau de neurone est distribuée à travers l'ensemble des constituants et non localisée dans une partie de mémoire sous la forme d'un symbole.

Le réseau de neurone ne se programme jamais pour réaliser une ou telle tâche, il est entraîné sur des données acquises, grâce au mécanisme d'apprentissage qui agit sur les constituants du réseau afin de réaliser au mieux la tâche souhaitée. [24]

#### II.4 Structure des réseaux de neurones artificiels

En général on peut dire que les réseaux de neurones peuvent être caractérisés comme des modèles de calculs apprenant, généralisant et organisant des données.

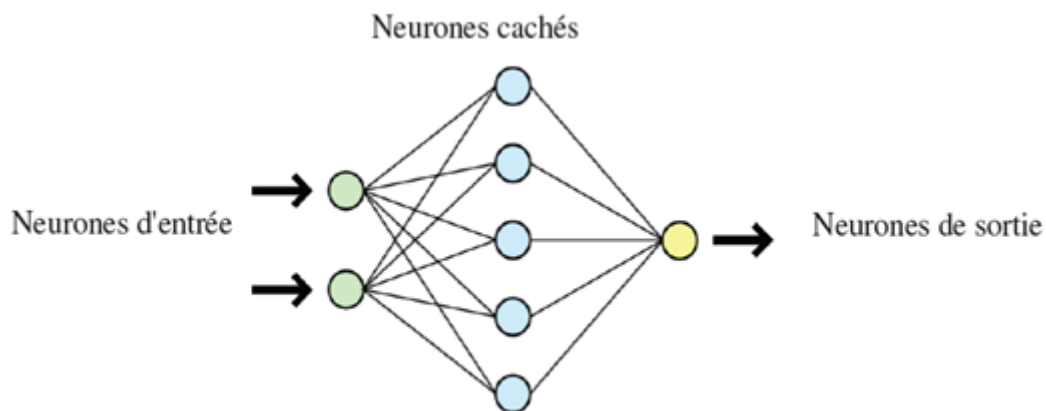


Figure II.2 Schéma simple d'un réseau de neurone [26]

Un réseau de neurone artificiel contient un grand nombre d'unités, les neurones, qui communiquent entre eux en s'envoyant des signaux à travers de liens, appelées connexions synaptiques. En général le système de neurone possède trois types des neurones :

- les neurones d'entrée qui reçoivent les données ;
- les neurones de sortie qui envoient les données par la sortie du système ;
- les neurones cachés, dont les signaux d'entrée et de sortie demeurent dans le système. [26]

$U_i$  représente la somme pondérée des entrées du neurone, elle est donnée par :

$$U_i = \sum_j w_{ij} x_j + b_i$$

où:  $x_j$  représente l'entrée  $j$  connectée au neurone  $i$ .  $b_i$  le seuil interne du neurone.  $w_{ij}$  désigne le poids de la connexion reliant l'entrée  $j$  au neurone.  $O_i = g(U_i)$  est la sortie du neurone et  $g$  sa fonction d'activation.

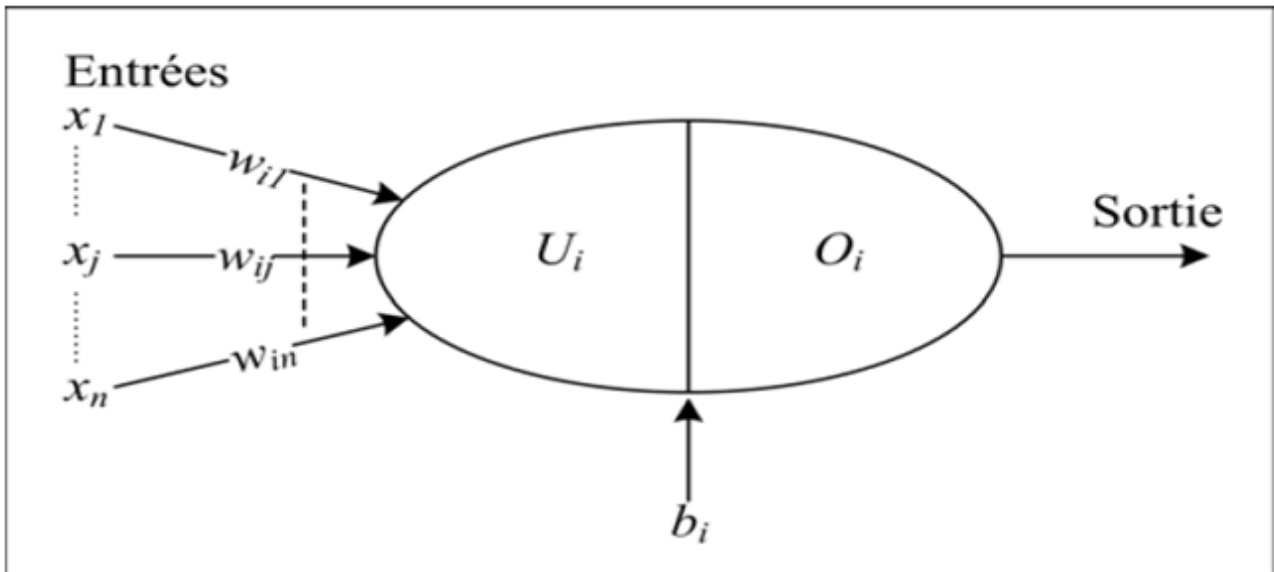


Figure II.3 Modèle de base d'un neurone formel[27]

La fonction d'activation de chaque neurone détermine ses propres caractéristiques. Par conséquent, le type du neurone est caractérisé par sa fonction d'activation. Conformément au neurone biologique, les fonctions d'activation sont généralement croissantes et continues. Les fonctions les plus utilisées sont la fonction linéaire et la fonction sigmoïde.

Leur choix revêt une importance capitale et dépend souvent du type de l'application et du domaine de variation des variables d'entrée/sortie. Un réseau de neurones est un système informatique qui a des caractéristiques semblables aux réseaux de neurones biologiques. Il est constitué de plusieurs unités (neurones) organisées sous forme de niveaux différents appelés couches du réseau. Les neurones appartenant à la même couche possèdent les mêmes caractéristiques et utilisent le même type de fonction d'activation. Entre deux couches voisines les connexions se font par l'intermédiaire de poids qui jouent le rôle des synapses. L'information est portée par la valeur de ses poids, tandis que la structure du réseau de neurones ne sert qu'à traiter l'information et l'acheminer vers la sortie. La structure ou la topologie d'un réseau de neurones est la manière dont les neurones sont connectés. Les structures résultantes peuvent être très variées mais elles sont souvent réparties en deux grandes familles à savoir : les réseaux de neurones non bouclés et les réseaux de neurones bouclés [27].

Sur la figure suivante on peut voir la structure d'un neurone artificiel. Le neurone calcule la somme de ses entrées puis cette valeur passe à travers la fonction d'activation pour produire sa sortie.

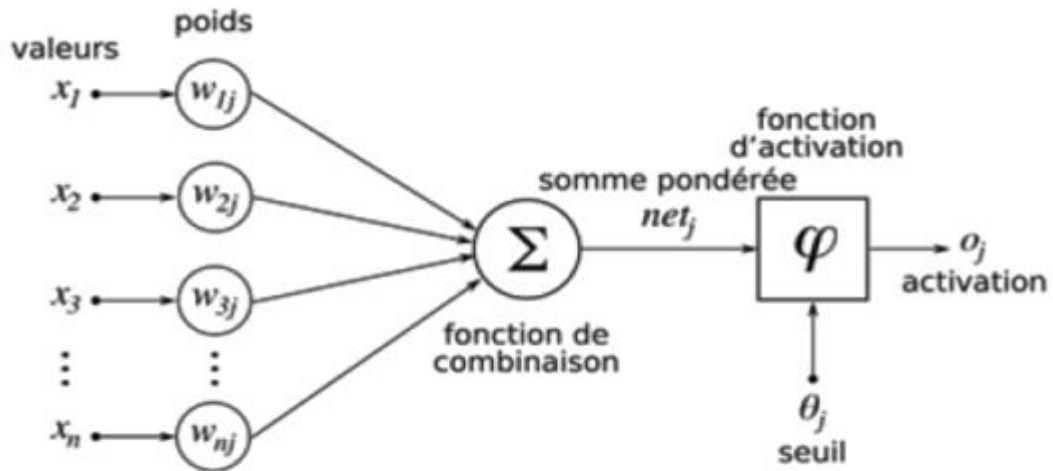


Figure II.4 Structure d'un neurone artificiel [26].

On peut comparer la correspondance entre les propriétés respectives de neurones biologiques et neurones artificiels comme le montre la Figure II.5. [23]

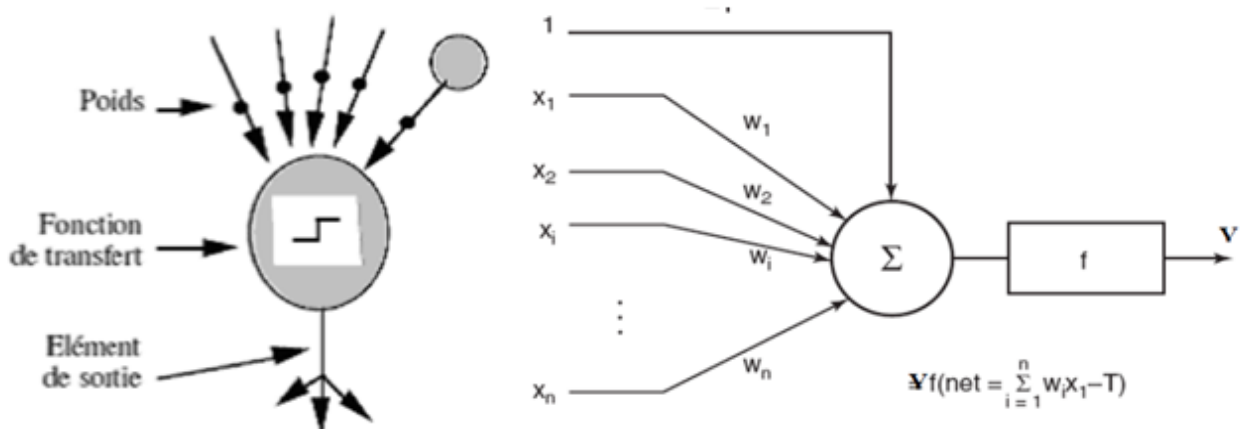


Figure II.5: Correspondance entre neurones biologiques et neurones artificiels[23]

**II.5 Architectures de réseaux de neurones Artificiels:**

L'organisation des réseaux de neurones peut se faire à partir de deux grandes catégories d'architecture: les réseaux statiques et les réseaux dynamiques.

**II.5.1. Réseaux statique (non bouclés)**

Un réseau est non bouclé, ou statique, si son graphe ne possède pas de cycle. Il réalise donc, de manière générale, une relation algébrique non linéaire entre ses entrées et ses sorties [28].

C'est à dire qu'il est obligatoire de représenter un réseau graphiquement par un ensemble de neurones connectés entre eux et que l'information circule des entrées vers les sorties sans retour (un seul sens), c'est à dire que la sortie courante n'a aucune influence sur les entrées futures.[29]

On peut dire que l'utilisation dominante des réseaux non bouclés est l'organisation des réseaux en plusieurs couches de neurones appelés réseaux multicouches ou perceptron multicouches. Un réseau multicouche comporte une couche d'entrée, une ou plusieurs couches cachées, et une couche de sortie.

Dans un tel réseau, la sortie de chaque neurone de la couche (L) est reliée à l'entrée de chaque neurone de la couche (L+1). L'architecture des réseaux multicouches est représentée dans la Figure II.6 [30]

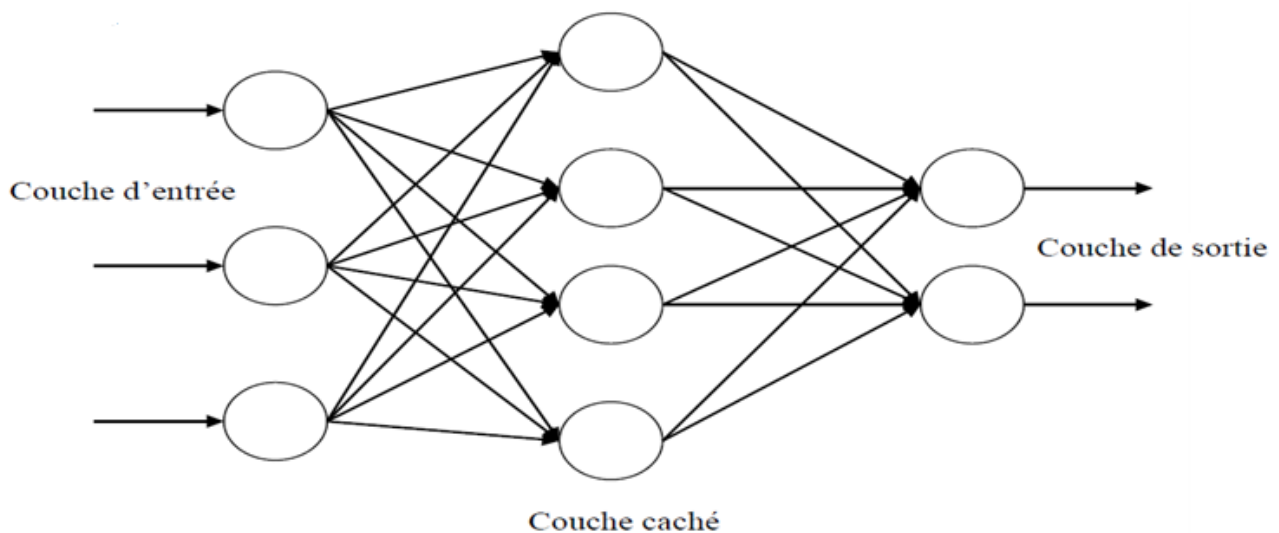


Figure II.6 Réseau multicouche[30]

### II.5.1. Réseaux dynamique (bouclés)

Les réseaux de neurone bouclés sont utilisés pour la modélisation dynamique de processus non linéaires et pour leur commande. [31] [32] Ces réseaux récurrents sont aux contraire des réseaux statiques, ils peuvent avoir une topologie de connexion quelconque (autrement dit un réseau totalement connecté) comprenant des boucles qui aux entrées la valeur d'une ou plusieurs sorties. L'état général (global) dépend aussi de ses états précédents comme il est indiqué sur la Figure II.7[33].

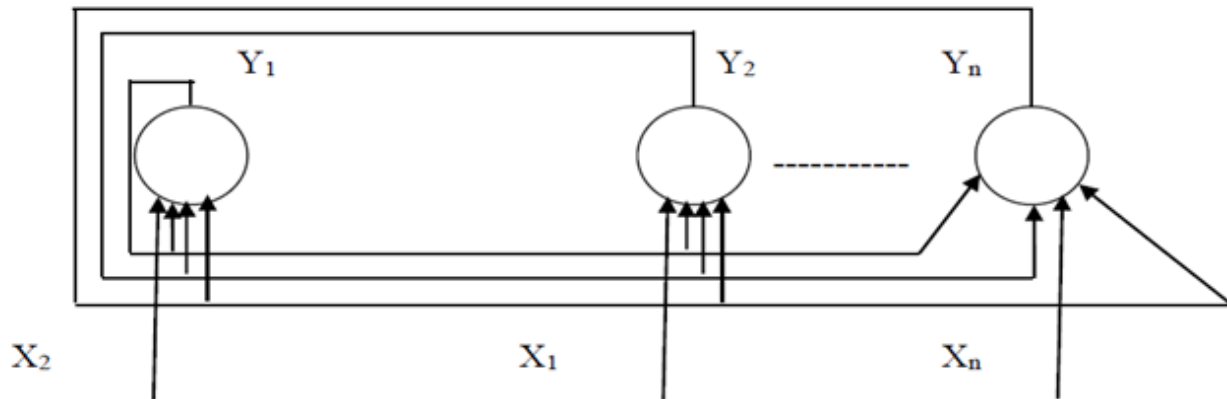


Figure II.7 Exemple de réseaux de neurone dynamique (Modèle de Hopfield)

## II.6 Types des réseaux de neurones artificiels (RNA):

Pour concevoir un réseau de neurones, nous devons établir des connexions entre les neurones. Nous avons quatre types principaux de connexion: directe, récurrente, latérale et à délais. Tous les réseaux de neurones utilisent la connexion directe pour acheminer l'information de l'entrée vers la sortie. La connexion récurrente permet d'acheminer l'information de la sortie des neurones des couches supérieures vers les entrées des neurones précédents. Les réseaux de neurones qui doivent choisir un neurone gagnant utilisent la connexion latérale pour établir une relation entre les neurones de sortie et la maintenir. Finalement, les problèmes temporels sont résolus par les modèles de réseaux dynamiques avec des connexions à délais [34]. Les connexions entre les neurones peuvent être complètes ou partiellement complètes. Une connexion est complète lorsque les neurones d'une couche inférieure sont reliés à ceux de la supérieure et elle est locale lorsque les deux couches de neurones ne sont pas complètement reliées. Une couche est définie comme un ensemble de neurones situé au niveau d'un réseau de neurones.

### II.6.1 Perceptron:

Le perceptron est la forme la plus simple d'un réseau de neurones, il modélise la perception visuelle. Il comprend trois principaux éléments: la rétine, les cellules d'association et les cellules de décision. La fonction d'activation utilisée dans ce réseau est de type tout ou rien (0 ou 1). L'apprentissage du perceptron peut se faire avec plusieurs méthodes déjà utilisées, il n'y a qu'une seule couche de poids modifiables entre les cellules d'association et les cellules de décision. Le perceptron est limité dans ses applications. Premièrement, il ne peut être applicable que dans la classification dont les variables sont linéairement séparables et deuxièmement la sortie ne peut être que 0 ou 1.



**II.6.2 Perceptron multicouche:**

Cette classe est la plus importante des réseaux de neurones car elle représente la généralisation du perceptron monocouche avec une fonction d'activation de type sigmoïde et une ou plusieurs couches cachées. Le vecteur d'entrée se propage dans le réseau de couche en couche jusqu'à la sortie, l'entraînement de celui-ci se fait avec l'algorithme par la rétro-propagation de l'erreur [34][35]. Ce réseau est caractérisé par son modèle du neurone traitant les non-linéarités. Il peut comporter une ou plusieurs couches cachées et un plus grand nombre de connexions permettant de résoudre la majorité des problèmes. Même avec les avantages des couches cachées et la performance de l'algorithme d'apprentissage, il reste plusieurs problèmes non réglés comme le choix du nombre de couches, le nombre de neurones par couche et le problème des minimums locaux où le réseau peut converger.

**II.6.3 Réseau de neurones linéaire:**

Cette classe de réseaux diffère du perceptron car elle possède un neurone dont la fonction d'activation est linéaire. L'une des règles d'apprentissage permet d'effectuer une descente de gradient de l'erreur sur une mesure d'erreur quadratique [36]. Les domaines d'application comprennent la commande, le contrôle, et le traitement du signal. L'avantage de ce réseau est qu'il converge sur un seul minimum si la solution existe, sinon l'ajout de couches n'a aucun effet. Parmi ses inconvénients, il est limité à une couche de sortie et ne peut résoudre que les problèmes dont la relation entrées/sorties est linéaire.

**II.6.4 Réseau RBF:**

Les réseaux RBF ("*Radial Basis Function*") sont des réseaux à couches qui ont comme origine une technique d'interpolation nommée la méthode d'interpolation à base de fonction radiale BFR. Ce réseau comporte une seule couche cachée dont la fonction d'activation est appelée fonction-noyau ou gaussienne et une couche de sortie avec une fonction d'activation linéaire. La méthode RBF est particulière par ses réponses utiles pour un domaine de valeurs restreint. La réponse de la fonction-noyau est maximale au noyau et décroît généralement de façon monotone avec la distance qui existe entre le vecteur d'entrée et le centre de la fonction-noyau. Afin d'approximer un comportement donné, les fonctions-noyau sont assemblées pour couvrir par leurs champs récepteurs l'ensemble des données d'entrée. Ces fonctions sont ensuite pondérées et la somme de leurs valeurs est calculée pour produire la valeur de sortie. Les réseaux RBF sont capables de calculs puissants. L'apprentissage est plus rapide et plus simple mais demande beaucoup de neurones par rapport aux réseaux multicouches. De plus, ils s'avèrent davantage insensibles à la destruction de leurs poids. Leur domaine d'application est vaste: le traitement d'image, la reconnaissance de forme, et surtout dans les problèmes de classification.

### **II.6.5 Réseau Hopfield:**

En 1982 John Hopfield a publié un article intitulé « Neural network and physical system with émergent collective computation habilités », dans lequel, il introduit l'architecture du réseau et il a décrit comment les capacités de calcul peuvent être construites, il illustre une mémoire associative, qui peut être implantée avec son réseau [37]. Selon lui, le système recherche des états stables, attracteurs dans son espace d'états. Les états voisins, tendent à se rapprocher d'un état stable, ce qui autorise la correction des erreurs et la capacité à compléter des informations manquantes.

Ce réseau est basé sur les principes de la physique statistique et il est fondamentalement une mémoire adressable par son contenu. Les neurones sont tous interconnectés. Plusieurs domaines d'application sont possibles, en particulier: les mémoires associatives et l'économie. La principale limitation est qu'il n'y a pas de couches cachées.

### **II.7. Domaine d'application des réseaux de neurones (RNA)**

Les RNA aujourd'hui ont des applications dans des domaines variés:

#### **II.7.1 L'application au traitement du signal :**

Dans ce domaine les applications réalisées traitent essentiellement de la reconnaissance de signatures radar ou sonar cette application a été construite à partir de signaux sonars et des traits caractéristiques de ces signaux déterminés par les experts comme utiles pour identifier une cible en environnement bruité. Le réseau agit comme un filtre pour éliminer le bruit.

#### **II.7.2 L'application à la planification :**

L'optimisation est souvent sollicitée pour la planification des actions ou des tâches. A cet effet, l'utilisation du modèle de Hopfield paraît très utile pour la résolution des problèmes d'optimisation.

#### **II.7.3 L'application au contrôle :**

Grâce à leurs propriétés (parallélisme de traitement, capacité d'adaptation, et de généralisation, etc...). Les réseaux de neurones sont appliqués au contrôle intelligent. Par exemple plusieurs tentatives ont été faites pour appliquer les réseaux de neurones ; Ces applications peuvent être classifiées en plusieurs méthodes, telles que : le contrôle supervisé. Le contrôle inverse et le contrôle neuronal adaptatif.

### II.7.4 L'application au diagnostic :

Les réseaux de neurones sont bien adaptés à la résolution des problèmes de diagnostic, utilisant la classification automatique des signaux et des formes. Dans ce contexte on distingue plusieurs applications des réseaux de neurones pour le diagnostic des défaillances et en particulier, pour le diagnostic des pannes des machines électriques.[21]

## II.8 Modélisation générale

On peut modéliser un réseau de neurone par des élémentaires qu'il s'agit de :

**II.8.1. Nature de ses entrées :** qu'ils peuvent être binaire (0 ou 1) ou réelles appartenant souvent à l'intervalle bornée  $[a, b]$ .

**II.8.2. Fonction des entrées:** elle peut définir le pré traitement effectuée sur les entrées.

**II.8.3 Fonction d'activation (seuil) :** Qui représente la limite à partir de la quelle le neurone s'activera, la sortie de neurone dépend d'une fonction de transfert. On peut identifier trois principaux types de fonctions les plus connues: binaire à seuil, rampe avec saturation et la sigmoïde[21]

### II.8.3.1. Fonction binaire à seuil:

La figure II.8 présente la fonction d'activation , le seuil introduit une non-linéarité dans le comportement du neurone, c'est le modèle tout ou rien



Figure II.8: Fonction d'activation binaire à seuil.

### II.8.3.2. Fonction à rampe avec saturation:

Cette fonction représente un compromis entre la fonction linéaire et la fonction seuil: entre ses deux bornes, elle confère au neurone une combinaison linéaire de l'entrée. À la limite, la fonction linéaire est équivalente à la fonction seuil (Figure II.9).

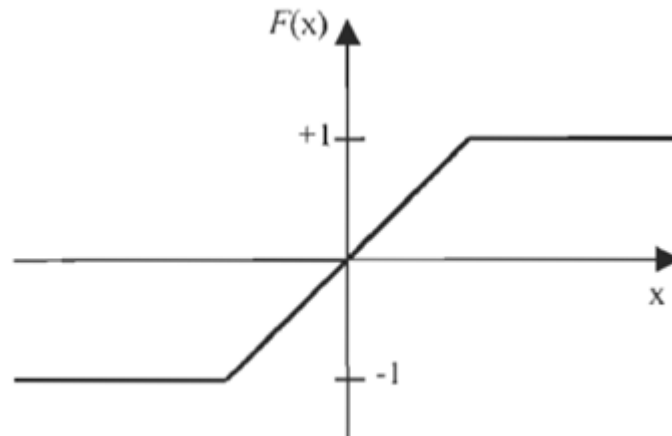


Figure II.9: Fonction d'activation à rampe avec saturation.

### II.8.3.3. Fonction sigmoïde:

La fonction sigmoïde est une fonction continue qui maintient la sortie dans l'intervalle  $[-1,1]$  (Figure II.10). Son avantage principal est l'existence de sa dérivée en tout point. Elle est employée en général dans le perceptron multicouche

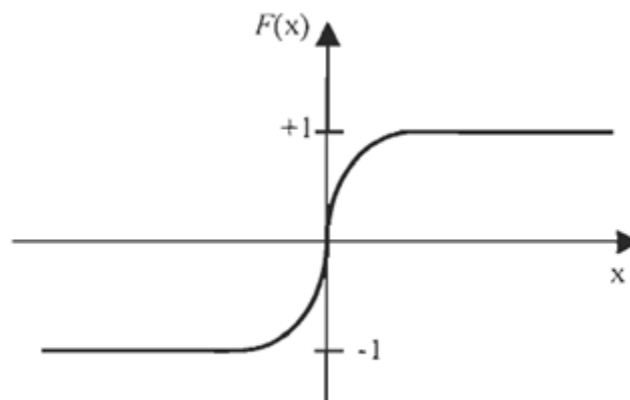


Figure II.10: Fonction d'activation sigmoïde.

**II.8.4 Fonction de sortie :** Cette fonction calcule la sortie du neurone en fonction de son état d'activation.

## II.9 Apprentissage dans les réseaux de neurones

L'apprentissage dans le contexte des réseaux de neurones est le processus de modification des poids de connexions (y compris les biais) ou plus rarement du nombre de couches et de neurones, afin d'adapter le traitement effectué par le réseau à une tâche particulière[27].

On distingue trois familles d'apprentissage :

**II.9.1 apprentissage supervisé:**

Dans ce cas, un superviseur (ou expert humain) fournit une valeur ou un vecteur de sortie (appelé cible ou sortie désirée), que le réseau de neurones doit associer à un vecteur d'entrée. L'apprentissage consiste dans ce cas à ajuster les paramètres du réseau afin de minimiser l'erreur entre la sortie désirée et la sortie réelle du réseau.

**II.9.2 apprentissage semi-supervisé ou apprentissage par renforcement:**

Ce mode d'apprentissage, suppose qu'un comportement de référence n'est pas possible, mais en revanche, il est possible d'obtenir des indications qualitatives (vrai, faux, ...) sur les performances du réseau.

**II.9.3. Apprentissage non supervisé:**

Dans cet apprentissage, les données ne contiennent pas d'informations sur la sortie désirée. Il n'y a pas de superviseur. La tâche du réseau consiste, par exemple dans ce cas, à créer des regroupements de données selon des propriétés communes (classification).

**II.10 Règles d'apprentissage**

L'apprentissage d'un Réseau de Neurones Artificiels se base sur des règles d'apprentissage. Nous citons ci-après les règles les plus répandues [38]:

**II.10.1. Règle de Hebb:**

Elle permet de modifier la valeur des poids synaptiques en fonction de l'activité des unités qui les relient. Le but principal est le suivant : si deux unités s'activent en même temps la connexion qui les lie est renforcée (c'est une connexion excitatrice) si non elle est affaiblie (c'est une connexion inhibitrice).

**II.10.2. Règle delta**

Qui calcule la différence entre la valeur de la sortie et la valeur désirée pour ajuster les poids synaptiques. Elle emploie une fonction d'erreur, nommée « la moindre carré moyen », basée sur les différences utilisées pour l'ajustement des poids.

### **II.10.3. Règle d'apprentissage**

Compétitif qui ne concerne qu'un seul neurone. Son principe est de regrouper les données en catégorie. Les neurones similaires vont donc être rangés dans une même classe en se basant sur des corrélations des données et seront représentés par un seul neurone. L'architecture d'un tel réseau possède une couche d'entrée et une couche de compétition. Une forme est présentée à l'entrée du réseau. Elle est projetée ensuite, sur chacun des neurones de la couche compétitive. Le neurone gagnant est celui qui possède un vecteur de poids le plus proche de la forme présentée en entrée. Chaque neurone de sortie est connecté aux neurones de la couche d'entrée et aux autres cellules de sortie (c'est une connexion inhibitrice) ou à elle même (c'est une connexion excitatrice). La sortie dépend alors de la compétition entre les connexions inhibitrices et excitatrices.

### **II.10.4. Règle de corrélation en cascade**

Est une technique d'apprentissage, ajoute progressivement des neurones cachés au réseau jusqu'à ce que l'effet bénéfique de ces nouveaux neurones ne soit plus perceptible. Cette règle suit les deux étapes suivantes :

1. On entraîne le système par un apprentissage classique qui s'effectue premièrement dans un petit réseau sans couche cachée.
2. On entraîne par la suite un petit groupe des neurones supplémentaires qui doit diminuer l'erreur résiduelle du réseau. La règle d'apprentissage utilisée modifie les poids de ces neurones. Le neurone qui réussit le mieux est retenu, et intégré au réseau. L'étape 1 est relancée encore une fois, pour permettre au réseau de s'adapter à la nouvelle ressource.

### **II.10.5. Règle de correction d'erreurs**

Qui est caractérisée par les étapes suivantes :

1. On commence avec des valeurs des poids de connexions qui sont pris au hasard.
2. On introduit un vecteur d'entrée de l'ensemble des échantillons pour l'apprentissage.
3. Si la sortie ou la réponse n'est pas correcte, on modifie toutes les connexions pour atteindre la bonne réponse

### **II.10.6. Règle de rétro-propagation**

Inventée par Rumelhart, Hinton et Williams en 1986. Elle s'utilise pour ajuster les poids de la couche d'entrée à la couche cachée. Cette règle peut aussi être considérée comme une généralisation de la règle delta pour des fonctions d'activation non linéaire et pour des réseaux multicouches. Les poids dans le réseau de neurones sont au préalable initialisés avec des valeurs aléatoires. On considère ensuite un ensemble de données qui vont servir comme un échantillon d'apprentissage. Chaque échantillon possède ses valeurs cibles qui sont celles que le réseau de neurones doit atteindre lorsqu'on lui présente le même échantillon.

### **II.11 Détermination des entrées/sorties du réseau de neurones**

Pour toute conception de modèle, la sélection des entrées doit prendre en compte deux points essentiels :

\_ Premièrement, la dimension intrinsèque du vecteur des entrées doit être aussi petite que possible, en d'autre terme, la représentation des entrées doit être la plus compacte possible, tout en conservant pour l'essentiel la même quantité d'information, et en gardant à l'esprit que les différentes entrées doivent être indépendantes.

\_ En second lieu, toutes les informations présentées dans les entrées doivent être pertinentes pour la grandeur que l'on cherche à modéliser : elles doivent donc avoir une influence réelle sur la valeur de la sortie [27].

### **II.12 Elaboration de la structure du réseau**

La structure du réseau dépend étroitement du type des échantillons. Il faut d'abord choisir le type de réseau : un perceptron standard, un réseau de Hopfield, un réseau à décalage temporel, un réseau de Kohonen, un ARTMAP etc... Par exemple, dans le cas du perceptron multicouches, il faudra aussi bien choisir le nombre de couches cachées que le nombre de neurones dans cette couche.

#### **II.12.1 Choix du type de réseau de neurones**

Comme il s'agit d'une modélisation statique type boîte noire pour approcher une fonction bornée et continue, la meilleure approche neuronale envisageable est celle des MLP à une seule couche cachée.

**II.12.2 Nombre de couches cachées :**

Mis à part les couches d'entrée et de sortie, il faut décider du nombre de couches intermédiaires ou cachées. Sans couche cachée, le réseau n'offre que de faibles possibilités d'adaptation. Néanmoins, il a été démontré qu'un Perceptron Multicouches avec une seule couche cachée pourvue d'un nombre suffisant de neurones, peut approximer n'importe quelle fonction avec la précision souhaitée [27].

**II.12.3 Choix du nombre de neurones dans la couche cachée**

Il n'existe pas de règle permettant de prévoir à l'avance le nombre de neurones dans la couche cachée. On pourrait penser qu'en plaçant un grand nombre de neurones dans la couche cachée, on obtiendrait systématiquement de meilleurs résultats. Mais il convient de prendre garde d'une part du temps de calcul, d'autre part du risque du sur apprentissage ou du sous apprentissage lorsque leur nombre est insuffisant.

**II.12.4. Choix de la fonction de transfert**

La fonction de transfert neuronale, que l'on appelle aussi fonction d'activation, est la fonction mathématique appliquée à la sortie de chaque neurone. Là encore, il n'existe pas de méthode permettant de prévoir a priori quelle sera la fonction qui donnera les meilleurs résultats. Mais pour les MLP, il faut des fonctions de type sigmoïde pour les neurones de la couche cachée et généralement une fonction linéaire pour le neurone de la couche de sortie.

**II.13. Avantages et limites**

Les Réseaux de Neurones Artificiels ont plusieurs avantages dont on peut citer [39,40]:

- 1- L'apprentissage
- 2- Le parallélisme massif l'architecture d'un réseau de neurone permet le traitement parallèle et rapide des informations.
- 3- Tolérance d'incertitude très élevée.
- 4- Etant une multiple copie d'unités simples (les neurones), ils sont donc facilement extensibles
- 5- facilité d'utilisation car il ne nécessite pas une compréhension approfondie.



6- Choix de types d'architecture et de fonction d'activation de réseaux diverses.

Bien que les réseaux de neurones soient capables d'effectuer beaucoup de tâches [40]

7- Difficulté de choix de L'architecture et le paramétrage du réseau

8- L'importance de L'état initial du réseau car il a une très grande influence sur la fiabilité et le temps de calcul et de réponse du réseau, donc un mauvais choix des poids initiaux peut conduire à un blocage d'apprentissage

### **Conclusion**

Dans ce chapitre, nous avons décrit quelques notions de bases sur la théorie des réseaux de neurones, les différentes formes des fonctions d'appartenance, les structure de neurones biologiques et neurones artificiels et les types des réseaux de neurones artificiels pour démontrer comment créé un réseau de neurone; nous avons cité les principaux opérateurs de réseaux de neurones et aussi les domaines d'application des réseaux de neurones.

Dans le chapitre suivant nous détaillerons l'utilisation de réseaux de neurones artificiels pour la prédiction de la tension contournement et de courant de fuite d'un isolateur de Haute Tension, ce qui prouvera l'utilité de cette nouvelle technique dans le domaine de la Haute Tension.