

الجمهورية الجزائرية الديمقراطية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE

جامعة محمد بوضياف بالمسيلة

Université Mohamed Boudiaf - M'sila

كلية التكنولوجيا

FACULTE DE TECHNOLOGIE

قسم الإلكترونيك

DEPARTEMENT D'ELECTRONIQUE

**Mémoire**

N° d'ordre .....

Présentée pour l'obtention du diplôme de

**Master**

Filière : Electronique

Spécialité : Système Embarque

**THEME**

**Classification de la maladie d'Alzheimer à  
l'aide de l'apprentissage statistique**

Présenté par

**OUANAS Houdham et KOUIDRI Yacine**

Devant le jury composé de :

**Nom & Prénom**

**Grade**

**Qualité**

DJERIOUI Mohamed

MCB

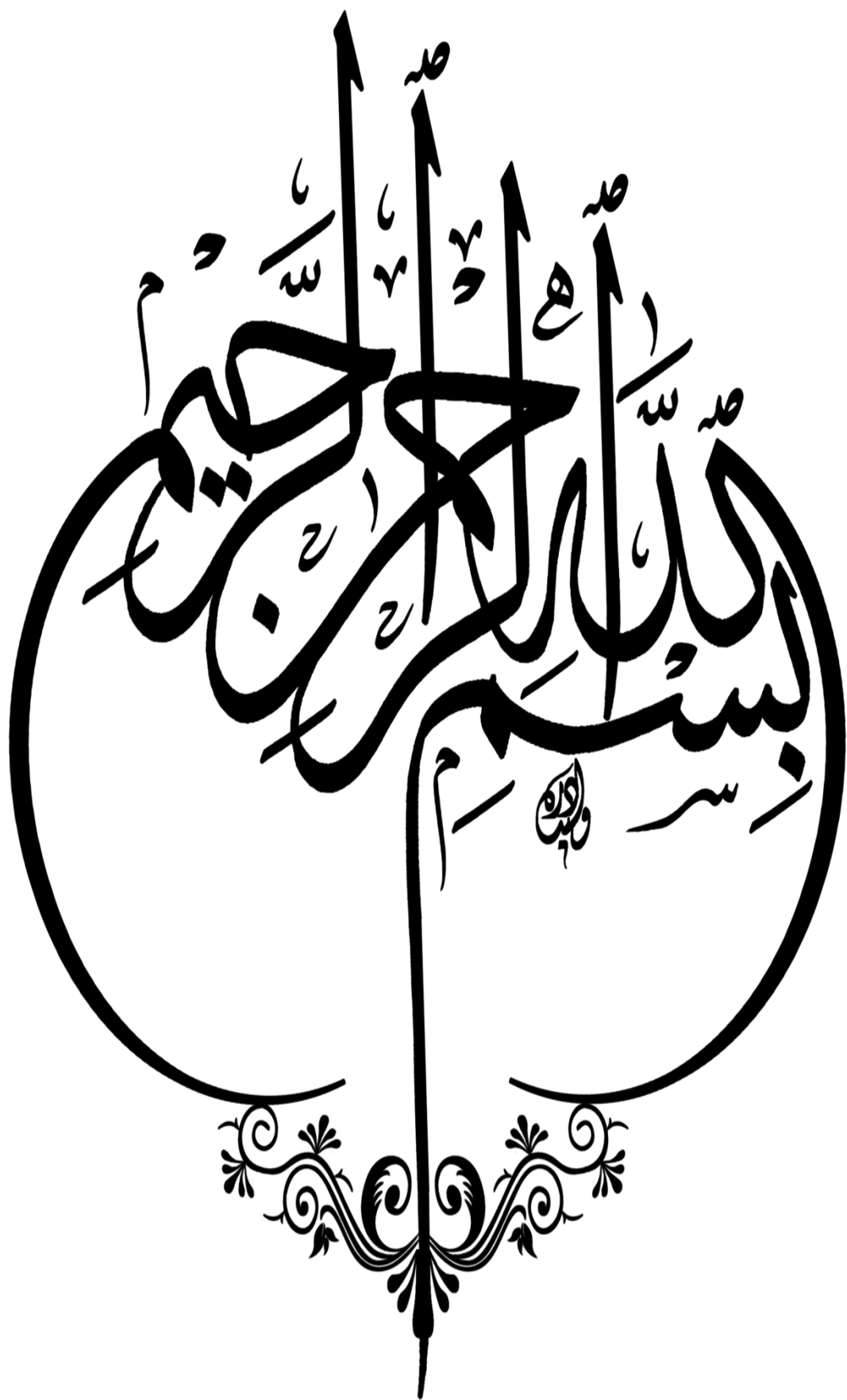
Encadreur

LADJAL Mohamed

MCA

Co-Encadreur

**Année universitaire : 2020/2021**





## *Remerciements*

*Au début, nous remercions et remercions Dieu  
Tout-Puissant, qui, grâce à Lui, nous a permis au  
cours de ces années d'étude et de patience de  
pouvoir accomplir cet humble travail.*

*Ensuite, nous tenons à remercier grandement  
l'Encadreur et Co-Encadreur*

*DR.M.DJERIOUI et DR.M.LADJAL*

*pour leur aide précieuse et leurs précieux  
conseils.*

*Nous remercions également tous les enseignants  
du département d'électronique de l'Université de  
M'Sila, en particulier les membres du jury pour  
notre travail.*

*Merci à tous*



## *Dédicaces*

*Nous dédions ce modeste travail :*

*KOUIDRI Yacine*

*OUANAS Houdham*

*Nos chers parents*

*Surtout ma mère.*

*Nos chers frères et sœurs.*

*Toute la famille OUANAS.*

*Toute la famille KOUIDRI.*

*Toute la famille DJOUDI.*

*Tous les enseignants qui m'ont aidé.*

*Tous nos amis.*

*Tous les collègues de notre promotion*

*(2020/2021).*

*Enfin à tous ceux qui nous ont*

*encouragés et soutenus*

## ***Abréviations***

**RNA** : les réseaux de neurones artificiels.

**ELM** : Les Machines d'apprentissage extrême.

**PMC** : Perceptron Multi Couches.

**RN** : Un réseau de neurones.

**MLP**: multi-layer perceptron.

***f***: est la fonction d'activation.

**$S_M$** : La couche de sortie.

**$R^n$** : muni de la distance euclidienne.

**IRM** : Imagerie par Résonance Magnétique

**T** : Taux de classification.

**NCC** : Nombre des Couches Caché

**NNCC** : Nombre Neurone des Couches Caché

**Test\_app**: le test d'apprentissage.

# *Liste des Tableaux*

Tableau.I.1 :Les différentes étapes du diagnostic de la maladie d'Alzheimer.....	9
Tableau.II.1: la relation entre le neurone biologique, le neurone formel et le RNA .....	21
Tableau.II.2: Correspondance RNA - domaines d'application.....	27
Tableau.III.1: Informations d'attributs de la base Alzheime.....	43
Tableau.III.2: Résultats de test de modèle MLP.....	47
Tableau. III.3. : Résultats de test de modèle ELM.....	49
Tableau.III.4 : Temps d'entraînement des MLP et ELM avec deux méthodes de normalisation. .....	49
Tableau. III.5 : Tableau comparatif des modèles (MLP et ELM).....	50

## *Liste des figures*

Figure 1.1: Définition et chiffres de la d'Alzheimer.....	6
Figure 1.2: Image radiographique d'une personne qui n'est pas malade.....	9
Figure 1.3: Image radiographique d'une personne malade.....	9
Figure .II.1 : Neurone biologique et ses principaux composants.....	19
Figure. II.2 :Modèle de base d'un neurone formel.....	21
Figure II.3 : Différents types de fonctions d'activations pour le neurone.....	23
Figure .II.4: Réseau de neurones : architecture "feedforward" MLP à deux couches cachées.....	24
Figure.II.5: Structure générale d'un réseau de neurone artificiel(RNA).....	26
Figure .II.6 : Structure d'une couche de neurones.....	28
Figure.II.7: Un perceptron multicouche.....	29
Figure .II.8 : Réseau à connexions locales.....	30
Figure .II.9: Réseau à connexions récurrentes.....	30
Figure .II.10 : Les réseaux de neurones multicouches.....	31
Figure .II.11: Réseau de neurones à apprentissage machine extrême (ELM).....	39
Figure. III.1 : Architecture du réseau MLP.....	46
Figure.III.2 : Architecture du réseau ELM.....	48

# *Sommaire*

Introduction générale.....	01
----------------------------	----

## *Chapitre I Généralité sur Alzheimer*

Introduction.....	4
1. HISTORIQUE.....	4
2. Qu'est-ce que la maladie d'Alzheimer.....	5
3. Définition de la maladie d'Alzheimer.....	5
4. La maladie d'Alzheimer et ses manifestations.....	6
5. Les tests pour détecter Alzheimer – diagnostic et bilan.....	7
5.1. Les étapes du diagnostic.....	7
5.1.1. Le bilan neuropsychologique.....	7
5.1.2. L'Imagerie par Résonance Magnétique (IRM).....	8
5.1.3. La Tomographie par émission de positons (TEP).....	8
6. L'examen diagnostic doit être entrepris dès l'apparition de plusieurs signes.....	9
7. Les causes de la maladie d'Alzheimer.....	10
8. Les différentes formes de la maladie d'Alzheimer.....	11
8.1. Forme sporadique.....	11
8.2. Forme héréditaire.....	11
9. Rapport sur la maladie d'Alzheimer dans le monde.....	12
10. Les Connaissances [10].....	12
11. Comportment [10].....	13
12. Les traitements de la maladie d'Alzheimer.....	13
12.1. Les traitements psychosociaux de la maladie d'Alzheimer.....	13
12.2. Les mesures psychothérapeutiques contre la maladie d'Alzheimer.....	13

12.3. Les mesures de stimulation intellectuelle contre la maladie d'Alzheimer	14
12.4. Les mesures de stimulation émotionnelle contre la maladie d'Alzheimer	14
12.5. Les mesures de communication dans la maladie d'Alzheimer	14
13. Traitements pharmacologiques	15
CONCLUSION	15

## *Chapitre II Les Techniques D'Apprentissages Statistiques*

Introduction	17
1. Historique de RNA	18
2. Neurone biologique	19
3. Neurone	19
4. Structure de neurone	20
5. Fonctionnement	22
6. Modélisation générale	22
7. Réseaux de Neurones Artificiels	23
7.1. Définition de RNA	23
7.2. Domaine d'application des réseaux de neurones (RNA)	26
7.3. Différentes configurations de réseaux	28
7.3.1. Perceptron à une seule couche	28
7.3.2. Réseau multicouche (PMC)	29
8. Algorithme de retro-propagation du gradient (generalized delta rule)	33
9. Formalisation d'apprentissage	33
9.1. Adaptation des poids	34
9.2 Algorithme de la rétro-propagation	37
10. Les applications des méthodes neuronales	38
11. Les Machines d'apprentissage extrême (ELM)	39
11.1. Formulation mathématique de l'ELM	39

11.2. Caractéristiques d'ELM	40
11.3. L'algorithme de l'ELM	41
Conclusion	41

## *Chapitre III SIMULATION*

INTRODUCTION	42
1.BASE DE DONNÉES UTILISÉE	42
2.Normalisation des données :	43
3.1. Normalisation Min et Max	44
3.2. Normalisation standardscaler	44
3.Apprentissage et Test	45
4. La Prédiction	...45
4.1. Prédiction avec MLP	47
4.2. Prédiction avec ELM	47
4.3. Comparaison	49
5. Discussion des résultats	51
5.1. Analyse et évaluation	51
Conclusion	52
Conclusion générale	53

# Introduction générale

Dans le domaine de la médecine, un défi majeur consiste à prélever un échantillon vital ou un signal du patient par l'intermédiaire d'un capteur spécifique ou par une analyse chimique physique lors de l'étude, du traitement et de l'extraction de l'information pertinente à la maladie à diagnostiquer. Comme nous l'avons dit, le problème, c'est où se trouvent les médecins. L'un des plus grands problèmes avec la médecine, c'est le bon diagnostic de la maladie. Les maladies intracalculables, comme la leucémie et le cancer du sein, sont difficiles à diagnostiquer et précoces, et le diagnostic correct se produit rarement tôt parce qu'il est souvent détecté à des stades avancés, ce qui est dû à plusieurs raisons. Dans cette étude, nous avons examiné la maladie d'Alzheimer, qui, à son tour, est classée comme une maladie qui est difficile à diagnostiquer dans les premiers temps, parce que, comme nous le savons, il est devenu une maladie de l'âge dans lequel la médecine est impuissante, malgré l'évolution que nous voyons dans le monde médical. Grâce à la technologie moderne d'étude des données médicales, nous avons pu trouver un moyen de diagnostiquer des maladies autres que les maladies traditionnelles, et ils nous ont donné de bons résultats afin d'économiser du temps et de l'argent au patient.

Nous pouvons supposer que le problème de la reconnaissance de pathologies est vu comme un problème de reconnaissance de formes, où les formes représentent l'ensemble des observations ou des prélèvements liés aux patients. Actuellement, la reconnaissance de formes a été largement étudiée au cours des dernières années dans la littérature scientifique. Cependant, malgré toutes les recherches intensives, les techniques de reconnaissance disponibles dans ce domaine restent au-dessous des capacités visuelles humaines [01].

Les avancées technologiques ont facilité l'acquisition et le recueil de nombreuses données, notamment dans le domaine médical lors d'examen des patients. Ces données peuvent être utilisées comme support de décision médicale, conduisant aux développements d'outils capables de les analyser et de les traiter; dans la littérature, nous trouvons régulièrement la notion d'aide au diagnostic, ces systèmes sont même considérés comme étant essentiels dans beaucoup de disciplines, ces systèmes reposent sur des techniques issues de

## Introduction générale

l'intelligence artificielle mais les problèmes les plus intéressants sont souvent basés sur des données de haute dimension. Ces problèmes désignent les situations où nous disposons peu d'observations alors que le nombre de variables explicatives est très grand. Cette situation est de plus en plus fréquente dans plusieurs les applications [02].

La sélection de variables consiste à choisir parmi l'ensemble global de variables, un sous-ensemble de variables pertinentes pour le problème étudié. Cette problématique peut concerner différentes tâches de fouille de données, elle regroupe des méthodes permettant de sélectionner un sous-ensemble de variables parmi un ensemble de départ, en utilisant divers critères et différentes techniques [03].

Des systèmes de diagnostic médical permanents doivent alors assurer le contrôle des divers paramètres biomédicaux. Les méthodes traditionnelles sont basées sur la connaissance de différents paramètres effectuées en laboratoire, pour décider après sur son état, et chercher les méthodes adéquates pour le soin. L'intérêt donc est de disposer d'un contrôle efficace des différents patients pour une meilleure décision. Les techniques de l'intelligence artificielle (IA) qui servent comme outil de base pour l'aide à la décision.

Leur réponse est plus élaborée et peut être obtenue soit à partir de données brutes venant directement des variables de contrôle. Il est judicieux de supposer que le problème de diagnostic médical peut être vu comme un problème de reconnaissance de formes, où les formes représentent l'ensemble des mesures biomédicaux, et les sorties correspondent aux différents états des patients. Ces informations biomédicales sont ensuite utilisées comme des données d'entrée dans la procédure de diagnostic de la partie amont d'un système d'aide à la décision. La construction de ce type de ce système peut être vue aussi comme un problème d'apprentissage de relations entre variables à partir de données observées. Ils sont spécialement programmés avec des algorithmes dérivés de l'intelligence artificielle. Le problème devient dans ce cas-là, comme un problème de classification. Parmi les techniques d'IA utilisées, on trouve les réseaux de neurones artificiels (RNA), Ces techniques sont de mieux en mieux acceptées dans le médecine en au beaucoup qu'objet de modélisation et de diagnostic.

Les travaux présentés dans le cadre de cette thèse sont destinés à l'étude et à la mise en œuvre de deux modèles d'apprentissage neurone statistiques Les réseaux de neurones artificiels multicouches MLP et Théorie de la technique ELM (Extrême Learning Machine) en

## Introduction générale

médecine, des modèles sont appliqués pour classer les patients. La maladie d'Alzheimer elle a été choisie à l'étude pour laquelle les données sont recueillies. Une étude de simulation a été menée pour valider, évaluer et comparer les performances de ces modèles dans le but de faire un choix critique répondant à la problématique du contrôle et du suivi de la santé des patients.

Le travail réalisé est axé autour de trois chapitres qui sont présentés comme suit :

Le premier chapitre consiste à une introduction au maladie d'Alzheimer, ses manifestations et de ces tests pour détecter Alzheimer. Les étapes et l'examen du diagnostic, est de ce fait décrite. Les causes, Les différentes formes et les traitements de la maladie d'Alzheimer sont présentées.

Dans le deuxième chapitre, nous allons décrit un état de l'art sur les méthodes appliquées à la prédiction. Deux modèles (Neuronal et ELM) fondés sur ce type d'apprentissage ont été exposés, nous décrivons l'architecture et rappelons les propriétés générales des réseaux de neurones artificiels statiques (PMC) à apprentissage supervisé. Enfin de ce chapitre, nous présentons la théorie et l'algorithme de l'ELM.

Enfin, le troisième et dernier chapitre, est dédié à la conception et mise en œuvre d'un système de prédiction à base de techniques d'apprentissage statistique. L'objectif est de valider et d'évaluer les performances de chacune des méthodes présentées. Afin de mener une étude comparative permettant un choix décisif de la méthode la mieux adaptée pour la prédiction et la classification dans la décision de maladie d'Alzheimer, on évaluera pour ces méthodes les paramètres liés au précision. Une discussion des résultats conclura cette étude de simulation pour choisir la technique la mieux adaptée.

Une conclusion générale en fin de cette mémoire est prévue, elle retrace les différentes étapes réalisées et souligne les perspectives envisagées.

# CHAPITRE I

## GENERALITES SUR ALZHEIMER

### Introduction

La maladie d'Alzheimer (MA), est une maladie définie par une démence progressive et la présence abondante dans le cerveau de structures neuropathologiques caractéristiques. Le symptôme le plus précoce est généralement une perte de mémoire, suivie d'un déclin fonctionnel et cognitif supplémentaire, de sorte que les patients deviennent progressivement moins capables d'effectuer même des tâches de base.

La maladie d'Alzheimer est diagnostiquée cliniquement en effectuant des examens physiques et neurologiques et en vérifiant d'autres signes de déficience intellectuelle au moyen de tests neuropsychologiques et cognitifs standard.

Ce premier chapitre se veut introductif, il a pour objectif de donner une idée générale sur la maladie d'Alzheimer. La première partie est consacrée à la définition de la maladie d'Alzheimer ainsi que de ses différentes étapes de diagnostic. Les causes et les différentes formes de la maladie d'Alzheimer sont particulièrement présentées. Sont cités dans ce cadre, les différentes méthodes de traitements de la maladie d'Alzheimer.

### 1. HISTORIQUE

La maladie d'Alzheimer tient son nom du psychiatre et neurologue allemand Aloïs Alzheimer (1864-1915) qui, en 1906, associa les symptômes (déclin progressif des fonctions cognitives) à des lésions cérébrales spécifiques, les plaques amyloïdes et les dégénérescences neurofibrillaires, grâce à l'étude d'une patiente du nom de Augusta Deter. Par la suite, d'autres chercheurs viendront confirmer ses découvertes, et un deuxième cas identique en 1911 viendra valider sa théorie.

Pendant plus d'un demi-siècle, l'étude de la maladie d'Alzheimer est restée en suspens. C'est à partir des années 1970-1980 qu'est apparue une nouvelle dynamique grâce aux progrès de la médecine et en raison de l'augmentation des cas en lien avec le vieillissement de la population. Ainsi, dans les années 1980, ont été déterminées les formations biologiques à l'origine des plaques (la protéine bêta amyloïde) et des dégénérescences neurofibrillaires (la protéine Tau hyperphosphorylée).

Ces découvertes sont toujours déterminantes actuellement dans la recherche d'un traitement pharmacologique contre la maladie d'Alzheimer. Enfin, dans les années 1990, les progrès en génétique ont permis d'identifier plusieurs gènes liés à l'apparition de la maladie

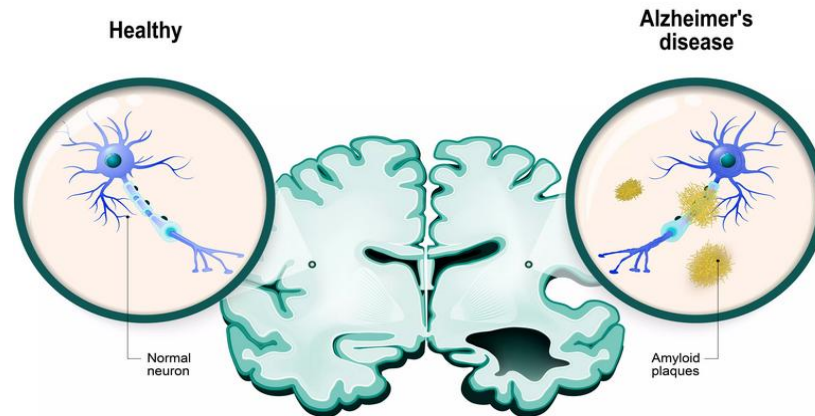
## **2. Qu'est-ce que la maladie d'Alzheimer**

La maladie d'Alzheimer est une maladie du cerveau provoquant une dégénérescence des cellules nerveuses et leur mort progressive. Peu à peu, la maladie s'étend : elle touche de nouvelles zones du cerveau et entraîne la perte de certaines fonctions cérébrales.

La maladie est dite neurone dégénérative, car elle cause une perte progressive des neurones. Sa prévalence augmente avec l'âge.

## **3. Définition de la maladie d'Alzheimer**

La maladie d'Alzheimer est une maladie neurone dégénérative complexe qui provoque un dysfonctionnement des connexions entre les neurones. Une meilleure compréhension de la maladie, l'étude de son évolution et l'identification de ses causes et facteurs de risque sont des éléments indispensables au développement de traitements efficaces. Les symptômes cliniques sont considérés comme étant liés à la perte neuronale qui touche principalement l'hippocampe, siège de la mémoire, et les aires néocorticales donnant à la maladie d'Alzheimer son surnom de « maladie de la mémoire »[05].



**Figure.1.1:**Définition et chiffres de la d'Alzheimer

- D'un point de vue clinique, elle affecte progressivement et insidieusement les fonctions cognitives de l'individu (mémoire, langage, raisonnement, apprentissage, résolution de problèmes, prise de décision, perception, attention...) entraînant une perte d'autonomie [05].
- D'un point de vue physiologique, elle se caractérise par la présence entre les neurones de plaques constituées par une accumulation de la protéine  $\beta$ -amyloïde, et par la présence d'enchevêtrements neurone fibrillaires intracellulaires dans le neurone causés par la protéine Tau anormalement agrégée [05].

La formation de plaques amyloïdes et les enchevêtrements neurone fibrillaires entraînent progressivement le dysfonctionnement des neurones et leur mort fonctionnelle.

#### 4. La maladie d'Alzheimer et ses manifestations

La maladie d'Alzheimer est une maladie neuro-dégénérative qui entraîne peu à peu une atteinte irréversible et progressive du système nerveux central. Ainsi, «chez la personne souffrant de la maladie d'Alzheimer, un grand nombre de cellules nerveuses perdent leurs liens avec les autres, cessent de fonctionner et meurent» Au démarrage de la maladie, c'est dans la région de l'hippocampe que se situent les lésions; des troubles de la mémoire apparaissent. Par la suite, les lésions s'installent dans le cortex associatif temporo-parieto-occipital, ce qui provoque des troubles du comportement. Par définition, le lobe temporal participe aux différentes fonctions du cerveau telles que le goût, l'audition, l'olfaction, la mémoire, le langage et le fonctionnement

automatique des viscères; le lobe pariétal, quant à lui, s'inscrit dans la connaissance du corps, permet d'avoir des repères spatiaux et de contrôler les gestes [11].

Enfin le lobe occipital contient l'aire visuelle qui reçoit et analyse les informations provenant du globe oculaire plus précisément de la rétine. Ces atteintes lésionnelles entraînent chez les personnes porteuses de la maladie d'Alzheimer les manifestations suivantes:

- l'amnésie qui se définit comme la perte de la mémoire. Au début, le symptôme n'est que partiel mais il évolue peu à peu vers un effacement total et irréversible de la mémoire.
- l'apraxie, c'est la perte de la capacité d'effectuer des mouvements coordonnés.
- l'agnosie qui est le trouble de la reconnaissance des informations saisies par les sens, particulièrement par la vue.
- la prosopagnosie c'est l'incapacité de reconnaître les visages connus.
- l'anosognosie: incapacité de reconnaître ses propres difficultés cognitives

## **5. Les tests pour détecter Alzheimer – diagnostic et bilan**

Le diagnostic de la maladie d'Alzheimer repose sur un examen complet des capacités cognitives de la personne (IRM, questionnaires...).

Cela se fait généralement dans le cadre d'une requête mémoire.

Cependant, il n'y a pas de test d'Alzheimer spécifique et il est impossible de détecter la maladie à un stade précoce. Au-delà du simple test de mémoire en ligne gratuit d'Alzheimer, les scientifiques tentent de développer de vrais tests pour la maladie d'Alzheimer.

### **5.1. Les étapes du diagnostic**

#### **5.1.1. Le bilan neuropsychologique**

Le bilan neuropsychologique permet de déterminer les troubles cognitifs du patient à travers une série de tests. Il évalue non seulement la mémoire mais aussi d'autres fonctions cognitives telles que l'orientation dans le temps et dans l'espace, le raisonnement, le langage, la compréhension et l'attention. Ces tests permettent de distinguer des patients atteints d'une maladie d'Alzheimer, même à un stade très

précoce, des personnes saines puis des patients atteints de démence. Cet examen est réalisé par un/une neuropsychologue. Les tests sont adaptés au patient en fonction de son niveau socioculturel et en fonction du stade d'évolution de la maladie [06].

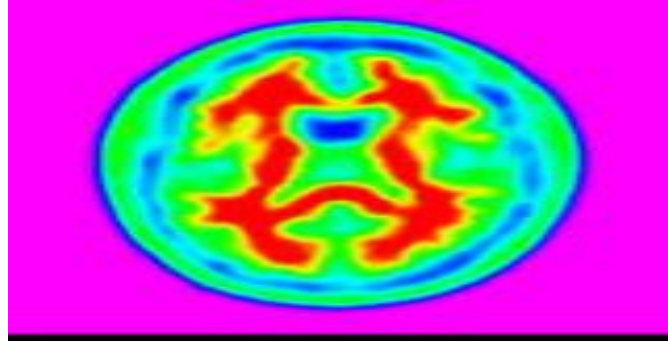
### **5.1.2. L'Imagerie par Résonance Magnétique (IRM)**

L'IRM permet de détecter une atrophie corticale et notamment une atrophie des hippocampes (structure cérébrale impliquée dans la mémoire, dont la taille est souvent diminuée dans la maladie d'Alzheimer). Elle permet par ailleurs d'éliminer d'autres causes de démences telles que la présence de lésions vasculaires, d'une tumeur cérébrale ou d'un hématome. Dans le cas où un examen IRM est contre-indiqué (notamment pour les patients ayant un pacemaker), un scanner cérébral peut être prescrit [06].

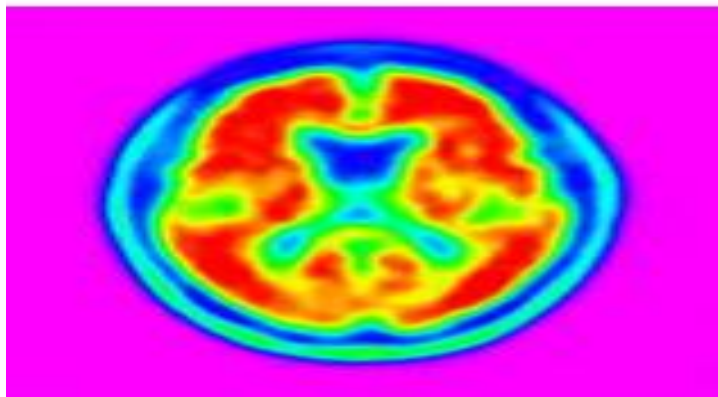
### **5.1.3. La Tomographie par Emission de Positons (TEP)**

La TEP permet d'avoir accès à une imagerie dite fonctionnelle. Elle permet de mettre en évidence un hypo-métabolisme (c'est-à-dire un fonctionnement moins efficace du cerveau) et de visualiser les lésions cérébrales caractéristiques de la maladie, notamment les plaques amyloïdes. Ci-dessous un exemple de visualisation par TEP de la charge en  $\beta$ -amyloïde dans le cerveau de deux hommes de 62 ans : un homme sain (Figure I.2) et un homme atteint de la maladie d'Alzheimer (Figure I.3).[06]

La couleur rouge correspond à la plus forte concentration en protéine  $\beta$ -amyloïde, alors que les couleurs bleu et verte correspondent à pas ou peu de protéine  $\beta$ -amyloïde. On voit nettement que la quantité de zones rouges est beaucoup plus importante chez le malade Alzheimer que chez l'homme sain.



**Figure.1.2:**Image radiographique d'une personne qui n'est pas malade



**Figure.1.3:** Image radiographique d'une malade

**6. L'examen diagnostic doit être entrepris dès l'apparition de plusieurs signes**

Le médecin tentera d'éliminer les autres causes. Le bilan complet peut être effectué dans le cadre d'une consultation mémoire ou chez un neurologue.

**Tableau. I.1:** Les différentes étapes du diagnostic de la maladie d'Alzheimer[07]

Examen	Description
Entretien	Entretien avec un patient ou un aidant : – détails sur les antécédents médicaux, – facteurs de risque, – traitements actuels, – difficultés observées,

Examen Clinique	Bilan de l'état général et cardio-vasculaire. Recherche des éléments suivants : – perte de poids, – hypertension, – confusion, – déficits auditif ou visuel
Évaluation des fonctions cognitive	Examen par un neurologue de : – l'attention et la concentration, – l'orientation, – la mémoire à court et long terme, – du langage, – des fonctions exécutives...
Évaluation des activités de la vie quotidienne	Examen portant sur les conséquences du trouble sur les actes du quotidien
Bilan comportemental et neuropsychologique	Recherche de troubles psychologiques et comportementaux, effectuée par un neuropsychologue
Examens biologiques	Tests de laboratoire et analyses sanguines, pour éliminer une autre affection susceptible d'entraîner des troubles cérébraux similaires
Imageries	Recherche d'autres causes aux troubles observés ou visualisation d'une atrophie de certaines régions du cerveau : – imagerie par résonance magnétique nucléaire (IRM), – tomodensitométrie.

### 7. Les causes de la maladie d'Alzheimer

La cause de la maladie d'Alzheimer est encore inconnue. Cependant, il existe des facteurs de risque qui favorisent l'apparition de la maladie,

- **L'âge et le sexe :** C'est le facteur de risque le plus important, une cause de la maladie d'Alzheimer que personne ne peut changer! En fait, plus vous avancez en âge, plus vous courez le risque de développer cette pathologie : elle est déclarée en général autour de 60-70 ans. En France,

il y a 40% de patients parmi les personnes de plus de 90 ans et 20% dans le groupe d'âge de 80-84 ans. Les femmes sont plus touchées, peut-être en raison de l'espérance de vie plus longue que les hommes [08].

- **Le facteur génétique / les antécédents familiaux** : Dans seulement 5-10% des cas de la maladie d'Alzheimer, la génétique a prouvé son rôle. Bien que les antécédents familiaux ne soient pas une cause de la maladie d'Alzheimer, il est un important "facteur de risque" de recherche. Avoir un être cher affecté augmente le risque de développer la maladie vous-même et l'héritage est toujours à l'étude [08].
- **La santé mentale** : Le stress, la dépression et l'isolement sont les causes de nombreux maux et peut favoriser la maladie d'Alzheimer. En outre, les études Les Américain sont montrées qu'un optimiste, jovial, sûr et serein serait une bonne protection! En effet, les relations sociales, ainsi que les activités physiques et intellectuelles, sont des facteurs protecteurs [08].
- **Les traumatismes crâniens** : Les personnes qui ont été victimes d'un traumatisme crânien ou de blessures à la tête, accompagnées d'une perte de conscience, courent un risque accru de souffrir de la maladie d'Alzheimer. En témoignent des études effectuées sur des boxeurs, afin de déterminer l'origine et les causes de la maladie d'Alzheimer [08].

## 8. Les différentes formes de la maladie d'Alzheimer

### 8.1. Forme sporadique

La forme sporadique (non héréditaire) est la forme la plus courante (90% des cas) ; elle se développe habituellement chez les gens de plus de 65 ans. Bien qu'elle soit non héréditaire, le patrimoine génétique de près de 10% des personnes atteintes de cette forme pourrait jouer un rôle dans le développement de la maladie, notamment lors de la présence de l'allèle  $\epsilon 4$  de l'Apo lipoprotéine E (ApoE) sur le chromosome 19. En effet, l'ApoE4 est identifiée comme l'un des principaux facteurs de risque dans la forme sporadique de la MA [09].

### 8.2. Forme héréditaire

La forme héréditaire est plus rare et représente seulement 10% des patients. C'est une forme familiale dont la transmission suit un mode autosomique dominant. Si un parent est atteint, chaque enfant a une probabilité de 50% d'hériter du gène de la maladie. Ces formes sont plus sévères et

peuvent se déclarer avant l'âge de 65 ans. Elles sont dues à trois mutations génétiques principales. Le premier gène porté par le chromosome 21, est celui du précurseur du peptide amyloïde (APP) qui déclenche la maladie entre 45 et 65 ans. Les deux autres gènes en cause portés respectivement par le chromosome 14 et 1 sont : la préséniline 1 (PS1) qui déclenche la maladie entre 28 et 65 ans et la préséniline 2 (PS2) qui la déclenche entre 40 et 85 ans. (Larner et al., 2006) L'effet de ces mutations semble être de favoriser la production ou l'accumulation d'un peptide neurotoxique (peptide A $\beta$ ) dans le cerveau responsable de la formation des dépôts amyloïdes. (Scheiner et al., 1996) [09].

## 9. Rapport sur la maladie d'Alzheimer dans le monde

En 2019, ADI estime à plus de 50 millions le nombre de personnes atteintes de démence dans le monde, chiffre qui devrait passer à 152 millions d'ici 2050. Quelqu'un développe une démence toutes les trois secondes et le coût annuel actuel de la démence est estimé à 1 billion de dollars américains, chiffre qui devrait doubler d'ici 2030. Dans le Rapport Mondial Alzheimer de 2012 'Surmonter la stigmatisation de la démence', nous avons exploré la nature de la stigmatisation liée à la démence. Le noyau du rapport s'est concentré sur les résultats d'une enquête menée auprès de 2'500 personnes ; des personnes atteintes de démence et des aidants dans un nombre limité de pays à travers le monde. Le rapport a aidé à définir la stigmatisation liée à la démence et a formulé des recommandations pour la combattre, ainsi qu'une série d'études de cas et de pratiques visant à réduire la stigmatisation [10].

## 10. Les Connaissances

- 95% du public pensent qu'ils vont développer une démence au cours de leur vie.
- 78% sont préoccupés qu'ils pourraient développer une démence à un moment donné.
- Une personne sur 4 pense qu'il n'y a rien que l'on puisse faire contre la démence.
- Deux personnes sur 3 pensent que la démence est causée par le vieillissement normal.
- 62% des professionnels de la santé pensent que la démence fait partie du processus normal du vieillissement.

## 11. Comportement

- Plus de 85% des répondants vivant avec une démence ont déclaré que leur opinion n'avait pas été prise au sérieux.
- Environ 40% du grand public pense que les médecins et les infirmières ignorent les personnes atteintes de démence.
- 67% des personnes atteintes de démence en Afrique et 63% en Asie du Sud-Est ont déclaré que des gens avaient plaisanté au sujet de leurs symptômes de démence.
- Entre 35% dans les pays à revenu élevé et 57% dans les pays à revenu faible et intermédiaire ont déclaré avoir été traités injustement dans leurs relations amoureuses et intimes.
- Environ 30% des Européens sont prêts à accueillir une personne atteinte de démence, ce chiffre passant à 62% en Asie du Sud-Est et à 71% dans la région de la Méditerranée orientale.

## 12. Les traitements de la maladie d'Alzheimer

### 12.1. Les traitements psychosociaux de la maladie d'Alzheimer

Les traitements dits « psychosociaux » visent à permettre au patient de conserver une certaine qualité de vie malgré la maladie : activité physique, capacités sensorielles, communication, vie sociale. Ils sont le plus souvent mis en œuvre dans des consultations de jour (« hospitalisations de jour ») ou des centres spécialisés dans l'accueil des personnes souffrant de maladie d'Alzheimer [11].

### 12.2. Les mesures psychothérapeutiques contre la maladie d'Alzheimer

Au début de la maladie d'Alzheimer, les patients peuvent avoir recours à un professionnel de la psychothérapie pour mieux gérer l'annonce du diagnostic et les craintes qu'elle génère. Ensuite, un suivi psychothérapeutique régulier peut être bénéfique pour exprimer ce qui ne peut pas être dit aux proches, pour prendre un peu de recul et pour s'adapter aux contraintes imposées par la progression des symptômes[11]..

La prise en charge psychothérapeutique peut également s'avérer indispensable chez les patients qui présentent des troubles du comportement (par exemple, les patients agressifs ou dépressifs).

### **12.3. Les mesures de stimulation intellectuelle contre la maladie d'Alzheimer**

Les centres d'accueil destinés aux personnes souffrant de maladie d'Alzheimer proposent de nombreuses activités destinées à stimuler les capacités intellectuelles des patients. Ces activités peuvent utiliser la musique, les activités artistiques, la présence d'animaux domestiques, les jeux ou les exercices physiques, etc. Pour les patients en début de maladie, des « ateliers mémoire » sont menés pour les aider à compenser les symptômes de la maladie dans ce domaine.

### **12.4. Les mesures de stimulation émotionnelle contre la maladie d'Alzheimer**

Ce type d'intervention est destiné à prévenir l'isolement des patients et à maintenir leur capacité à échanger sur le plan émotionnel. Lorsque leur capacité à s'exprimer verbalement diminue du fait de la progression de la maladie, les patients ne perdent pas pour autant leur désir et leur capacité à communiquer sur le plan des émotions. Ces mesures peuvent, par exemple, prendre la forme d'échanges autour de vieilles photos, de chansons anciennes, d'objets accumulés pendant leur vie, etc.

### **12.5. Les mesures de communication dans la maladie d'Alzheimer**

Une attitude inadaptée peut embarrasser, choquer ou frustrer une personne atteinte de la maladie d'Alzheimer. Plusieurs conseils sont recommandés pour continuer à communiquer malgré les troubles du langage et de la mémoire : s'assurer que la personne vous voie ou vous entende, se placer en face d'elle, maintenir le contact visuel, parler lentement, laisser du temps au patient pour répondre, etc. Les personnes aidantes doivent être formées à une communication adaptée [11].

### 13. Traitements pharmacologiques

Il n'existe pas à l'heure actuelle de traitement curatif contre la MA et les maladies apparentées; il est toutefois possible de stabiliser ou d'atténuer temporairement la manifestation des symptômes. Il existe deux classes de médicaments pour la thérapie spécifique de la MA et des maladies apparentées : les inhibiteurs de l'acétylcholinestérase (IAChE) et l'antagoniste des récepteurs glutamatergiques NMDA (la mémantine). Quatre médicaments (donépézil, rivastigmine, galantamine et mémantine) sont homologués par Santé Canada pour le traitement symptomatique de la maladie d'Alzheimer, dont un (la rivastigmine) possède également un avis de conformité pour la démence associée à la maladie de Parkinson [12].

Les effets distinctifs des IAChE s'expliquent en partie par les cibles moléculaires, la réversibilité de l'inhibition et les propriétés pharmacocinétiques différentes. En règle générale, il est attendu que les IAChE inhibent de façon réversible l'hydrolyse de l'acétylcholine (ACh) par l'AChE au niveau de la fente synaptique, augmentent la concentration d'ACh et potentialisent la transmission cholinergique. Les IAChE sont reconnus pour avoir des effets cholinomimétiques, c'est-à-dire qu'ils reproduisent les effets de l'ACh et peuvent stimuler la transmission cholinergique dans le système périphérique, notamment dans les systèmes gastro-intestinaux et cardiovasculaire ou encore dans le système nerveux central. La mémantine est un antagoniste non compétitif d'affinité de faible à modérée des récepteurs glutaminergiques NMDA. L'intention thérapeutique sous-jacente à la prise de la mémantine est la protection des neurones contre les dommages potentiels associés à une concentration élevée de glutamate dans le cerveau [12].

### CONCLUSION

Dans ce premier chapitre, nous avons présenté des concepts sur la maladie d'Alzheimer et ce qu'est cette maladie, et nous avons également discuté de la façon de la détecter, de ses étapes de développement, de ses formes, et nous avons également vu les facteurs qui la provoquent, et finalement nous avons parlé de ses traitements, comme il n'y a toujours pas de traitements, et cette maladie reste évolutive et incurable, et le tueur, les médicaments prescrits pour la maladie

d'Alzheimer ne guérit pas mais réduit les symptômes de la vulnérabilité surrénale. L'objet du chapitre suivant va être orienté vers le vaste domaine des Machines Learning (apprentissage statistique), notamment les réseaux de neurones artificielles (RNA) et Extrême Learning Machine (ELM).

# CHAPITRE II

## Les Techniques D'Apprentissages Statistiques

### Introduction

L'informatique est la science du traitement automatique de l'information. Son développement est souvent confondu avec celui des machines de traitement : les ordinateurs. Depuis les débuts (ENIAC 1946) jusqu'à aujourd'hui, les ordinateurs sont devenus de plus en plus puissant.

En apprentissage automatique, domaine qui consiste à utiliser des données pour apprendre une solution aux problèmes que nous voulons confier à la machine, le modèle des Réseaux de Neurones Artificiels (ANN) est un outil précieux. L'apprentissage Automatique est un sous-domaine de l'Intelligence Artificielle, il s'intéresse à la recherche de solutions pour résoudre des tâches que nous voulons confier à la machine, et ce par le biais d'un apprentissage.

Dans le cadre de notre travail, la solution adoptée rappelons-le une fois de plus, est une surveillance et un contrôle de la MA par les techniques de reconnaissance de formes. Dans ce chapitre, nous allons donc pouvoir passer en revue des méthodes d'apprentissage statistique appliquées à la classification et la régression. Après une brève introduction, où nous allons la notion de modèle biologique et neurone formel, nous décrivons définition réseaux de neurones artificiels, modélisation générale, son architecture des réseaux, Les types d'apprentissage des réseaux de neurones et rappelons Les réseaux de neurones les plus utilisés (Perceptron Multi Couches (PMC)), ainsi que les algorithmes de retro-propagation du gradient et formalisation d'apprentissage [13]. Enfin, les aspects théoriques et fondements d'une autre technique appelée ELM sont aussi rappelés et décrits.

## 1. Historique de RNA

\_1890 : W. James, célèbre psychologue américain, introduit le concept de mémoire et propose ce qui deviendra une loi opérationnelle pour l'apprentissage réseaux neuronaux plus tard connu comme la loi de Heb.

- 1943 : J. Mc Culloch et W. Pitts laissent leurs noms à une modélisation du neurone biologique (un neurone au comportement binaire).Ceux sont les premiers à montrer que des réseaux de neurones de forme simple peuvent effectuer des fonctions logiques, arithmétiques et complexes symboliques (tout au moins au niveau théorique).

- 1949 : D. Hebb, physiologiste américain explique le conditionnement chez l'animal par les propriétés des neurones eux-mêmes. Par conséquent, un type de conditionnement de Pavlovian comme,

Nourrir un chien tous les jours à la même heure, provoque chez l'animal la sécrétion de salive à cette époque, même en l'absence de nourriture. Loi modifiant la Loi sur les biens immobiliers connexions neuronales qu'il propose explique en partie ce type de résultats expérimentation.

Les premiers succès : -1957 : F. Rosenblatt développe le modèle du Perceptron. Il construit le premier neurone ordinateur basé sur ce modèle et l'applique au domaine de la reconnaissance de formes. Notons qu'à cet époque les moyens à sa disposition sont limités et c'est une prouesse technologique que de réussir à faire fonctionner correctement cette machine plus de quelques minutes.

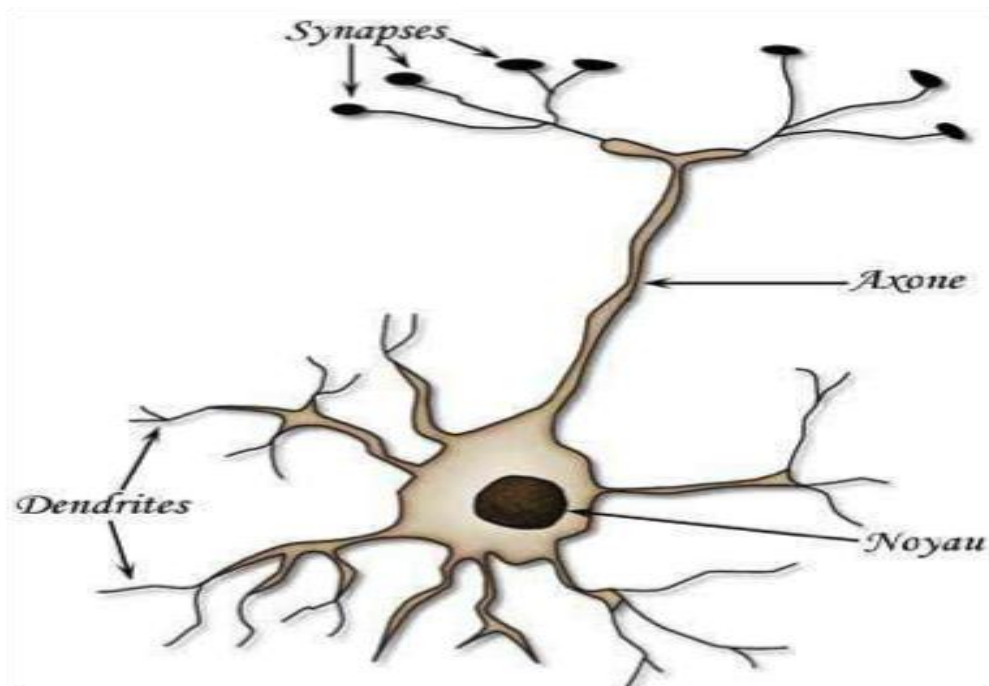
- 1960 : B. Widrow, un automaticien, développe le modèle Adaline (Adaptative Linear Element). Dans sa structure, le modèle ressemble au Perceptron, cependant la loi d'apprentissage est différente. Celle-ci est à l'origine de l'algorithme de rétro propagation de gradient très utilisé aujourd'hui avec les Perceptrons multicouches. Les réseaux de type Adaline restent utilisés de nos jours pour certaines applications particulières. B. Widrow a créé dès cette époque une des premières firmes proposant neuro-ordinateurs et neuro-composants, la "Memistor Corporation". Il est aujourd'hui le président de l'International Neural Network Society (INNS) sur laquelle nous reviendrons au chapitre Informations pratiques.

- 1969 : M. Minsky et S. Papert publient un ouvrage qui met en exergue les limitations théoriques du perceptron. Limitations alors connues, notamment concernant l'impossibilité de

traiter par ce modèle des problèmes non linéaires. Ils étendent implicitement ces limitations à tous modèles de réseaux de neurones artificiels. Leur objectif est atteint, il y a abandon financier des recherches dans le domaine (surtout aux U.S.A.), les chercheurs se tournent principalement vers l'IA et les systèmes à bases de règles.

## 2. Neurone biologique

Le système nerveux repose sur un nombre gigantesque d'unités élémentaires massivement interconnectées : les neurones. Le cerveau humain en possède 100 milliards de neurones et chacun d'entre eux est connecté en moyenne à 103 autres neurones. Le neurone est une cellule vivante. Il en existe de différentes formes (pyramidales, sphériques ou étoilées), mais tous ont une structure commune. Chacun de ces neurones est par ailleurs fort complexe. Essentiellement, il s'agit de tissu vivant et de chimie et sel décompose en trois régions principales (Le corps cellulaire, Les dendrites, L'axone) [14].



**Figure .II.1 :** Neurone biologique et ses principaux composants.

## 3.Neurone

Est une cellule vivante, qui peut prendre des formes variables (pyramidale, sphérique ou étoilée) sa forme est définit par une membrane qui sépare l'intérieur du neurone a l'extérieur. Ces

neurones sont considérés comme l'élément de base qui constitue les unités élémentaires de traitement dans le cerveau [15].

#### **4.Structure de neuron**

- **Le corps cellulaire**

C'est le corps du neurone et comme toutes les autres cellules du corps humain, sauf les hématies (globules rouges), il contient un noyau. Ce corps cellulaire est sous forme sphérique ou pyramidale contenant aussi les autres molécules essentielles à la vie de la cellule [14].

- **Les dendrites**

Ce sont des prolongements de la membrane du neurone qui reçoivent des signaux d'autres neurones. Les dendrites mesurent de 0,2 à 10microns de diamètre et une dizaine de microns de longueur. Il n'y a pas de contact direct entre les membranes des neurones connectés : cette liaison est assurée par une synapse [14].

- **L'axone**

C'est la fibre nerveuse principale, le seul prolongement de la membrane qui se sépare en une chevelure de filaments à destination des dendrites d'autres neurones pour propager le signal. [14].

- **Les synapses**

Elles assurent le passage du signal des terminaisons axonales vers les dendrites, via généralement un processus électrochimique [14].

- **Neurone formel**

Est un processeur très simple imitant grossièrement la structure et le fonctionnement d'un neurone biologique, la première version du neural formel est celle de McCulloch et Pitts qui est largement utilisée [16].

Un neurone formel est un opérateur effectuant la somme pondérée de ses entrées externes puis elle s'active suivant la valeur de cette sommation pondérée puis modulée par une fonction d'activation Figure II.2 [17].

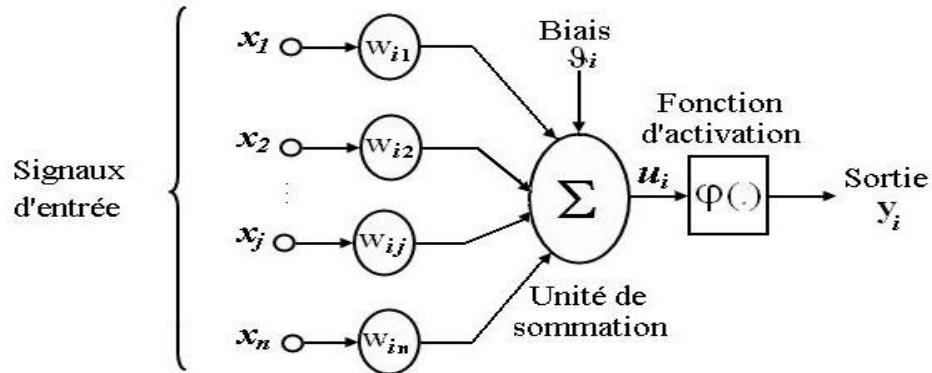


Figure.II.2 : Modèle de base d'un neurone formel.

Le modèle de la figure 2.2 est composé [18]

- ✓ Des entrées du neurone formel  $X_i = 1, 2, \dots, n$
- ✓ Des paramètres de pondération  $W_{ij}$  ;
- ✓ De la fonction d'activation ou de seuillage (non linéaire, sigmoïde, etc....) ;

Et d'une sortie du neurone formel.

Tableau. II.1 : la relation entre le neurone biologique, le neurone formel et le RNA [19]

Neurone biologique	Neurone formel
Dendrite	Signal d'entrée
Axone	Signal de sortie
Synapses	Poids de connexion
Corps cellulaire	Fonction d'activation
système nerveux	réseau neurone Artificiel
Neurone Le corps cellulaire (soma) Axone Synapse	Traitant élément, nœud, neurone artificiel, neurone abstrait, Niveau de l'activation, fonction de l'activation, fonction du transfert, la fonction de la sortie la communication avec d'autre neurone poids multiplicatifs

## 5. Fonctionnement

Avant de pouvoir utiliser les capacités de classification et d'approximation de fonctions d'un réseau de neurones, il faut le construire, ceci se déroule en quatre temps : [20]

- La construction de la structure du réseau (généralement empirique).
- La constitution d'une base de données de vecteurs représentant au mieux le domaine à modéliser. Celle-ci est scindée en deux parties : une partie servant à l'apprentissage du réseau (on parle de base d'apprentissage) et une autre partie aux tests de cet apprentissage (on parle de base de test).
- Le paramétrage du réseau par apprentissage.

Au cours de l'apprentissage, les vecteurs de données de la base d'apprentissage sont présentés séquentiellement et plusieurs fois au réseau. Un algorithme d'apprentissage ajuste le poids du réseau afin que les vecteurs soient correctement appris. L'apprentissage se termine lorsque l'algorithme atteint un état stable.

- La phase de reconnaissance qui consiste à présenter au réseau chacun des vecteurs de la base de test. La sortie correspondante est calculée en propageant les vecteurs à travers le réseau, La réponse du réseau est lue directement sur les unités de sortie et comparée à la réponse attendue.

Une fois que le réseau présente des performances acceptables, il peut être utilisé pour répondre au besoin qui a été à l'origine de sa construction.

## 6. Modélisation générale

On peut modéliser un réseau de neurone par des élémentaires qu'il s'agit de [21] :

- La nature de ses entrées : qu'ils peuvent être binaire (0 ou 1), (-1,1) ou réelles appartenant souvent à intervalle bornée  $[a, b]$ .
- La fonction des entrées : que sa signifie quelle peut définir le pré traitement effectuée sur les entrées.
- 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 dont les principales sont représentés sur la Figure II.3.

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

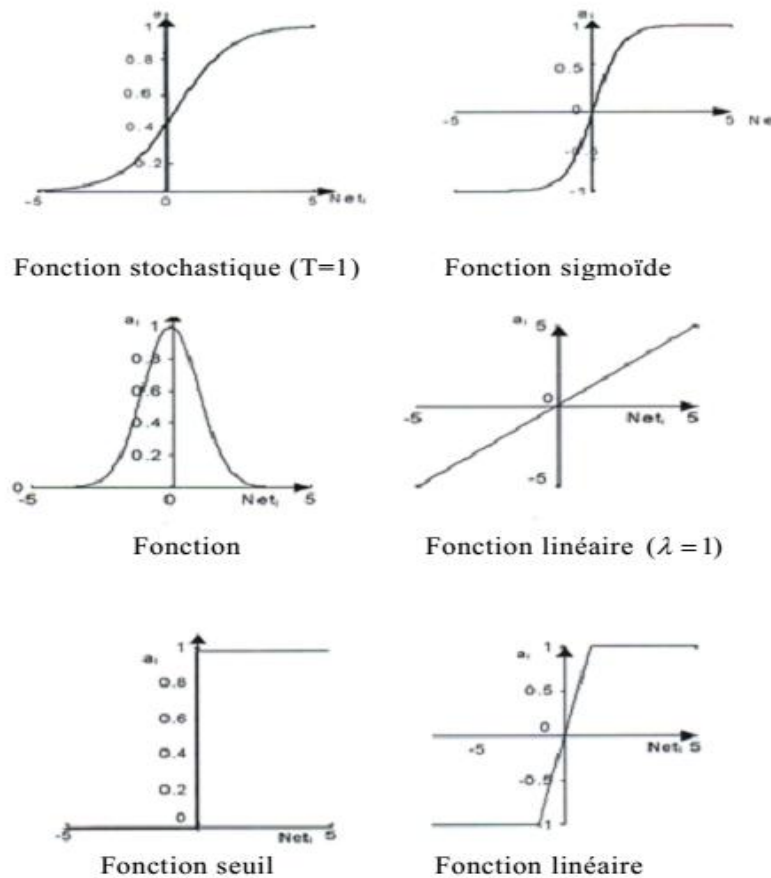


Figure .II.3 : Différents types de fonctions d'activations pour le neurone

## 7. Réseaux de Neurones Artificiels

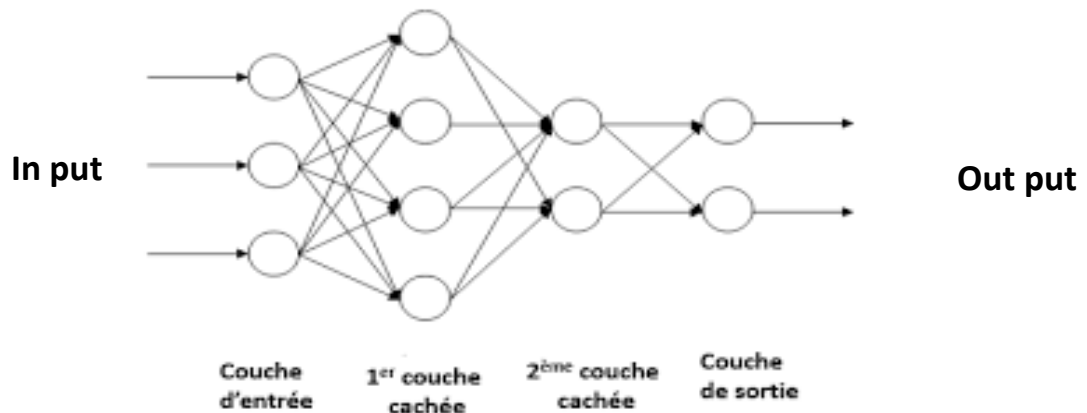
### 7.1. Définition de RNA

Un réseau de neurones est constitué par un certain nombre de neurones interconnectés. Chaque neurone peut recevoir plusieurs signaux intrant des autres neurones, les agrège en se basant sur une fonction d'entrée et génère, enfin, un signal sortant, basé sur une fonction de sortie, appelée fonction d'activation. Le signal sortant alimentera, ensuite, d'autres neurones, suivant la typologie du réseau. La valeur numérique d'un poids synaptique, relatif à une connexion entre deux nœuds, reflète l'ampleur de l'influence d'un neurone sur l'autre neurone et le sens de cette

influence. La résolution d'un problème neuronal revient à déterminer les poids relatifs aux connexions [22].

Dans sa forme typique, un réseau de neurones est composé de trois types de couches : une couche d'input, une couche d'output et des couches cachées. Les neurones de la couche d'input reçoivent les signaux de l'environnement, les neurones de la couche d'output émettent des signaux à l'environnement [22].

Néanmoins, les neurones des couches cachées n'ont aucune interaction directe avec l'environnement, d'où leur appellation. Les connexions entre les couches se font dans une seule direction, de la couche d'input vers la couche d'output, Chaque couche est formée au moins d'un neurone [22].



**Figure .II.4:** Réseau de neurones : architecture "feedforward" MLP à deux couches cachées

Le fonctionnement des neurones biologiques est très complexe, une modélisation fidèle requiert de monopoliser des ressources computationnelles importantes. Toutefois, il est possible de définir un modèle efficace inspiré du fonctionnement des neurones biologiques en admettant les (fausses) hypothèses suivantes:

- L'intégration des courants post-synaptiques est une transformation affine.
- L'information de phase des potentiels d'actions n'influe pas sur l'intégration des courants post-synaptiques.

- La fonction de réponse du neurone peut être assimilée à une fonction de notre choix  $g(x)$ .
- Tous les neurones peuvent être formalisés de la même manière.

La valeur de sortie d'un neurone en fonction de ses entrées ( $x = [x_1, \dots, x_n]$ ) devient donc ; avec  $W = [W_1, \dots, W_n]$  le vecteur des poids associés à chacune des entrées et  $b$  le biais, ou seuil, du neurone artificiel:

$$z(x) = g\left(\sum_{i=1}^n (w_i * x_i) + b\right) \quad (2.1)$$

Les fonctions d'activations  $g$  généralement utilisées sont: la sigmoïde et la tangente hyperbolique définies par les équations suivantes:

$$\mathbf{sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (2.2)$$

$$\mathbf{Tanh}(x) = \frac{e^{2x}-1}{e^{2x}+1} = 2x\mathbf{sigmoid}(2x) - 1 \quad (2.3)$$

Il est aussi possible de remplacer  $g$  par une fonction à base radiale, c'est à dire une fonction qui ne dépend que de la distance entre deux vecteurs, en l'occurrence ici  $\mathbf{x}$  et  $\mathbf{w}$ . La transformation affine est remplacée par cet opérateur de distance. Cela force une représentation locale: le neurone est activé seulement lorsque son entrée est au voisinage de  $\mathbf{w}$ . A l'inverse la transformation affine découpe l'espace en deux suivant l'hyperplan défini par l'équation:  $\mathbf{x} \cdot \mathbf{w}^t + \mathbf{b} = 0$ . Avec la fonction d'activation sigmoïde le neurone tend à être activé sur tout le demi-espace vérifiant  $\mathbf{x} \cdot \mathbf{w}^t + \mathbf{b} < 0$  Et à être désactivé sur tout le demi-espace vérifiant  $\mathbf{x} \cdot \mathbf{w}^t + \mathbf{b} > 0$ . On obtient une représentation distribuée.

Nous allons nous restreindre aux réseaux qui n'admettent pas de boucle dans leur graphe de calcul. Dans le cas contraire le réseau est dit récurrent et le modèle est dynamique. Ce type de modèle ne sera pas étudié dans cette thèse. Rajoutons les contraintes suivantes:

- Les réseaux de neurones sont arrangés en couches.
- Les neurones d'une couche admettent comme entrées uniquement les neurones de la couche précédente.

–Les couches, autres que celles de sortie et d'entrée du réseau, sont appelées couches cachées. Considérons  $z^0$  le vecteur d'entrée,  $z^i$  le vecteur des activations à la couche  $i$ ,  $s^i$  le vecteur des valeurs avant la fonction d'activation  $g$  ainsi que  $W^i$  et  $b^i$  re respectivement la matrice de poids et le vecteur de biais correspondants aux nueronsartificiels  $z^i$ . On obtient le modèle représenté à la figure II.5 et on a:

$$s^i = z^{i-1}w^i + b^i \quad (2.4)$$

$$z^i = g(s^i) \quad (2.5)$$

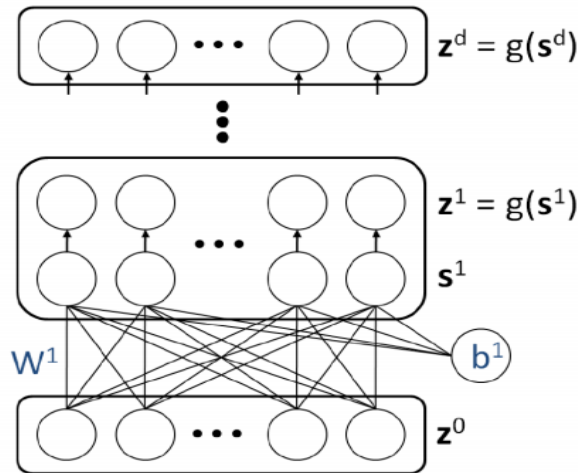


Figure .II.5: Structure générale d'un réseau de neurone artificiel(RNA)

## 7.2. Domaine d'application des réseaux de neurones

Les (RNA) aujourd'hui ont des applications dans des domaines variés parmi lesquelles on cite :

**7.2.1. L'application au traitement du signal :** Dans ce domaine l'application réalisées traitent essentiellement de la reconnaissance de signatures radar ou sonner. La société NESTOR a développe un réseau de neurones qui identifie une cible à coup sur (100% de réussite) et reconnaît du bruit avec un taux de réussite de 95%. Cette application à été construire 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 dans ces comme un filtre pour éliminer le bruit [21].

**7.2.2. 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 Hop Field et du modèle de Kohonen parait très utile pour la résolution des problèmes d'optimisation [21].

**7.2.3. L'application au contrôle :** Grâce à leur 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 classifiés en plusieurs méthode, Telles que : le contrôle supervisé. Le contrôle inverse et le contrôle neuronal adaptatif [21].

**7.2.4. 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. Exemples d'applications de chaque modèle [21].

**Tableau. II.2 :** Correspondance RNA – domaines d'application

Caractéristiques fonctionnelles	Type de RNA
Reconnaissance de forms	MLP, Hopefield, Kohonen, PNN
Mémoires associatives	Hopefield, MLP récurrents, Kohonen
Optimisation	Hopefield, ART, CNN
Approximation de fonctions	MLP, RBF
Modélisation et control	MLP, MLP récurrent, FLN
Traitement d'images	CNN, Hopefield
Classification et clustering	MLP, Kohonen, RBF, ART, PNN

### 7.3. Différentes configurations de réseaux

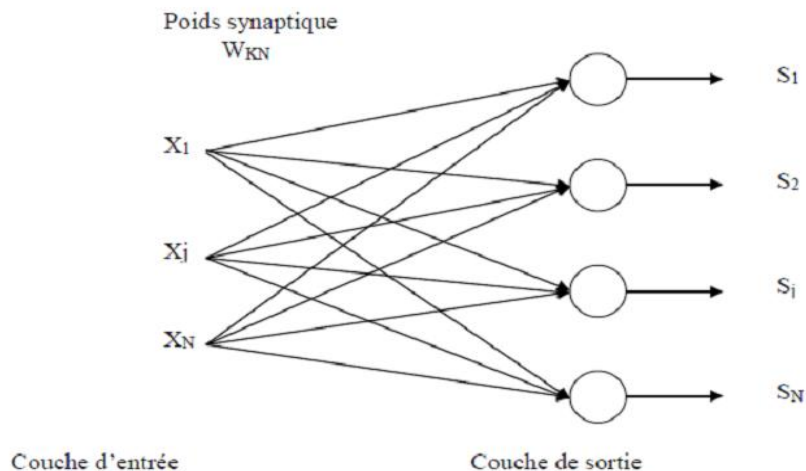
La Topologie du amélioré est décrite par les connexions comme les neurones qui les composent. Elle peut dépendre au hasard, simplement en chique on peut séparer une certain régularité.

#### 7.3.1. Perceptron à une seule couche

C'est le premier modèle qui fut présenté par F.Rosenblatt, dans ce type de réseaux, il y a une seule couche cachée qui lie les couches d'entrée aux couches de sortie. Seule la couche cachée est modifiable [23].

L'application avec un tel modèles sont très restreintes dans une même structure ayant en commun les mêmes entrées, chaque neurone agit indépendamment des autres et en particuliere reçoit aucune connexion en provenance des neurones de cette couche et traite des problèmes pour lesquels la sortie attendue n'est pas une seule valeur scalaire mais un vecteur scalaire.

L'exemple de réseau à une seule couche est donné par la Figure II.6.



**Figure .II.6 :** Structure d'une couche de neurones

##### 7.3.1.1. La convergence du perceptron

Au début de la convergence du Perceptron consiste à initialiser les petites. Valeurs aléatoires non nulle du poids synaptique  $W_i$  et le seuil  $\theta$  si l'erreur est

minimale et e réseau converge la sortie est acceptable, si non il fait varie le vecteur d'entrée.

La sortie est donne par [21] :

$$Y = f(\sum_{i=0}^{n-1} W_i X_i - \theta) \quad (2.6)$$

Avec :

$\theta$ : Le seuil.

$f$  : fonction d'activation.

### 7.3.2. Réseau multicouche

Cette topologie est affectée de neurones organisés en de çà taille de délivrance. Les connexions à cause les neurones se font avec les neurones des délivrances en aval (devise 18). Chaque neurone d'une flache est en conséquence connecté à intégraux les neurones de la flache suivante. Ce qui certains permettra d'intercaler la image de connotation de cheminement de l'inédite au torse d'un entrelacs et en conséquence déterminer les concepts de neurone d'recommandation, neurone de effloraison. On dit une flache d'recommandation aux neurones d'recommandation, flache d'effloraison aux neurones d'effloraison.

Les délivrance intermédiaires sont les délivrances cachées et n'ont annulé coudoirement avec l'extérieur.

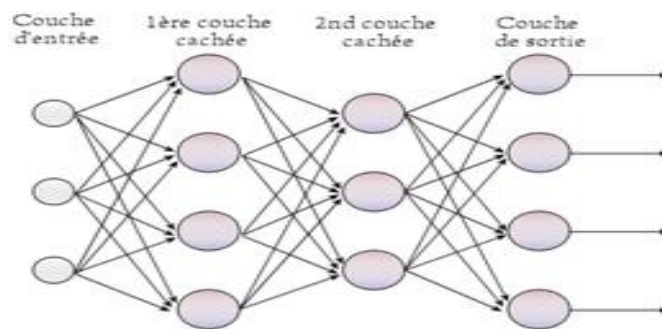
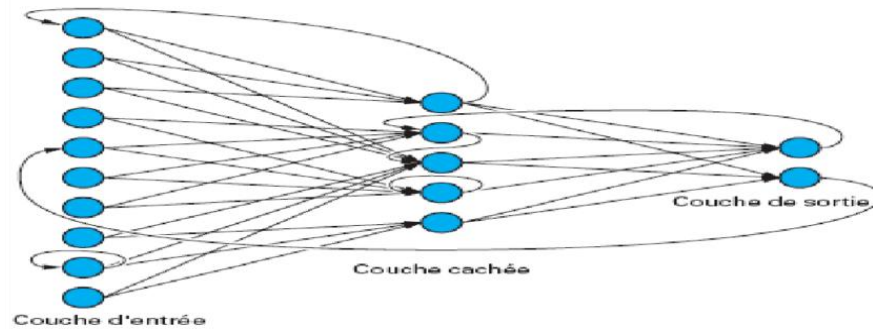


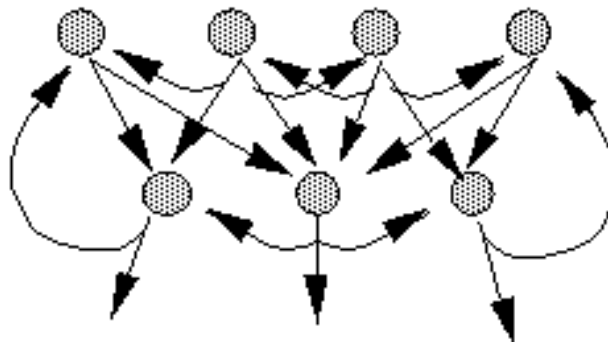
Figure .II.7: Un perceptron multicouche

a) **Réseau à connexions locales:** C'est une structure multicouche, comme larétine, permet conserver une certaine topologie. Ce n'est pas comme le cas Auparavant, nous avons ici chaque neurone connecté avec un nombre réduit et localisé neurones dans la couche de l'hirondelle (figure II.8), Les connexions sont alors moins beaucoup.



**Figure .II.8 :** Réseau à connexions locales

b) **Réseau à connexions récurrente:** Cette structure permet de prendre l'information derrière la direction de la propagation. Ils sont souvent des connexions locales (figure.II.9 ).



**Figure .II.9:** Réseau à connexions récurrentes

### 7.3.2.1. Les perceptrons multicouches(PMC)

Dans l'introduction aux réseaux neuronaux, nous avons souligné l'intérêt du Perceptron multiples dans l'évolution des réseaux neuronaux. Rappelez-vous, que l'origine de leur utilisation a été la découverte de règles d'apprentissage adaptées à son architecture. En effet, l'architecture Perceptrons multicouche (PMC et en anglais MLP pour *multi-layer perceptron*) est basé sur Perceptron, où des couches de neurones, appelées couches occultes, sont ajoutées

entre les entrée set la couche de sortie (figure.II.10). L'intérêt de cette évolution est de perceptron, pour optimiser la capacité d'apprendre des réseaux pour effectuer une tâche.

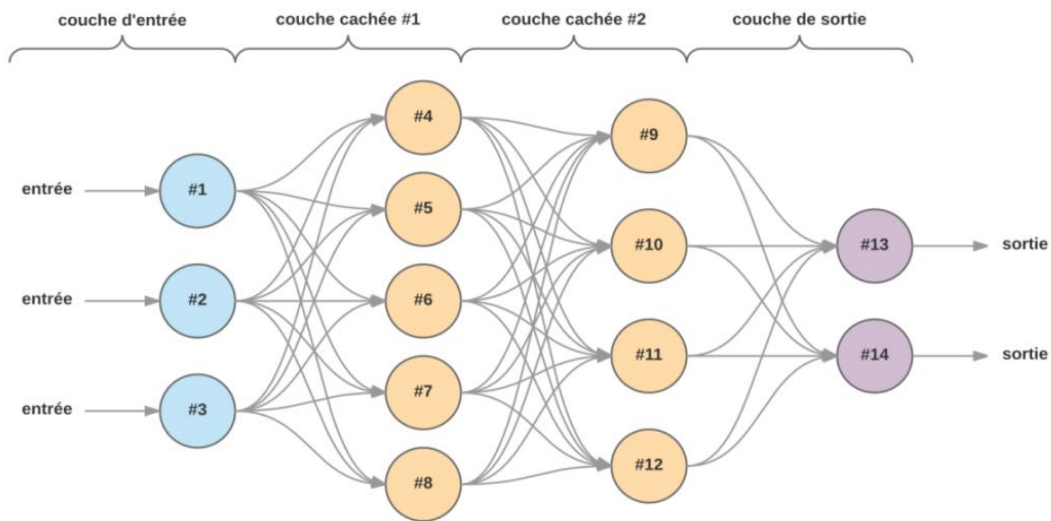


Figure .II.10: Les réseaux de neurones multicouches.

### 7.3.2.2. Etape d'apprentissage

Les réseaux de neurones sont l'outil puissant pour résoudre différents problèmes tels que la classification ou la prédiction. Les perceptrons multicouches ont été utilisés pour résoudre des problèmes difficiles par apprentissage supervisé à partir d'un algorithme connu sous le nom d'algorithme de rétro propagation du gradient qui se base sur la règle d'apprentissage par correction d'erreur. L'amélioration s'obtient avec le temps à partir de mesures réelles effectuées [24,25].

Un réseau de neurones apprend de son environnement à partir de processus itératifs en ajustant les poids synaptiques et les seuils. Théoriquement, le réseau s'améliore après chaque itération d'apprentissage [24,25].

Le principe d'apprentissage est l'estimation des paramètres  $\alpha_{j=0,p,k=1,q}$  et  $\beta_{k=0,q}$  par la minimisation d'une fonction coût (fonction perte) qui représente le but de l'apprentissage. Il faut noter que les deux paramètres  $\alpha$  et  $\beta$  sont les poids des entrées à estimer lors de la procédure d'apprentissage [26,27].

$$Q(\alpha, \beta) = \sum_{i=1}^n [y_i - \Phi(x; \alpha, \beta)]^2 \quad (2.7)$$

### 7.3.2.4. Rétro propagation de l'erreur

L'algorithme de rétro propagation de l'erreur du gradient est une généralisation de la règle de *Windows-Hoff* pour un réseau multicouche que l'on désigne par « back-propagation » [28].

Il figure aujourd'hui parmi les algorithmes d'apprentissage le plus utilisés. Il a été appliqué avec succès à une grande variété de problèmes tels que la prévision de consommation d'eau [29].

Il s'agit donc d'évaluer la dérivée de la fonction coût en une observation et par rapport aux différents paramètres.

$$\text{Soit : } y = \phi(x; \alpha, \beta) = \beta_0 + \beta' z \quad (2.8)$$

$$\text{Avec : } z_{ki} = f(\alpha_{k0} + \alpha_k x_i) \text{ et } z_i = \{z_{1i}, \dots, z_{qi}\} \quad (2.9)$$

Les dérivées partielles de la fonction perte quadratique s'écrivent [30] :

$$\frac{\partial Q_i}{\partial \beta_k} = -2(y_i - \phi(x_i))(\beta' z_i) z_{ki} = \delta_i z_{ki} \quad (2.10)$$

$$\frac{\partial Q_i}{\partial \alpha_{kj}} = -2(y_i - \phi(x_i))(\beta' z_i) \beta_k f'(\alpha_k x_i) x_{ip} = s_{ki} x_{ip} \quad (2.11)$$

Les termes  $X_i$  et  $s_{ki}$  sont respectivement les termes d'erreur du modèle courant à la sortie et sur chaque neurone caché. Ces termes d'erreur vérifient les équations dites de retro propagation [30]:

$$s_{ki} = f'(\alpha_k x_i) \beta_i \delta_i \quad (2.12)$$

Dont les termes sont évalués en deux passes. Une passe “*avant*”, avec les valeurs courantes des poids, l'application des différentes entrées  $x_i$  au réseau permet de déterminer les valeurs ajustées  $\hat{\phi}(x_i)$ .

La passe “*retour*” permet ensuite de déterminer les  $\delta_i$  qui sont “rétro-propagés” afin de calculer les  $s_{ki}$  et ainsi obtenir les évaluations des gradients.

### 7.3.2.5. Etape de généralisation

Une fois que le réseau neuronal a été formé (après l'apprentissage), il doit être testé avec une base de données différente de celle utilisée pour l'apprentissage. Ce test permet à la fois d'évaluer la performance du système neuronal et de détecter le type de données qui pose des problèmes. Si les résultats ne sont pas satisfaisants, il faudra modifier l'architecture du réseau ou la base d'apprentissage.

## 8. Algorithme de retro-propagation du gradient (generalized delta rule)

C'est un algorithme qui permet de stabiliser un réseau à plusieurs couches a été trouvé de façon indépendante par plusieurs équipes de chercheurs (Rumelhart, Parker et Le cun). La topologie d'un tel réseau est donc formée de plusieurs couches de neurones sans communication à l'intérieur d'une même couche.

L'apprentissage y supervisé, c'est a dire que l'on présente air réseau, en même temps, une forme et son modèle.

Le but est encore de minimiser l'erreur quadratique entre les sorties obtenues et celle souhaitées qui correspond de nouveau à une descente du gradient mais alors que dans un réseau à une seule couche on était assuré de la convergence de la méthode, ce n'est plus le cas dans un réseau multicouche, et on pourrait éventuellement rester bloqué sur un minimum local et non global, il s'avère qu'en fait ce la ne se produira pas dans la majorité des cas (mais cela reste expérimental).

Le principe est redistribué sur toutes les couches, y compris les couches cachées, une partie de l'erreur de manière récursive, en partant de la couche de sortie et en remontant vers la couche d'entrée [21].

## 9. Formalisation d'apprentissage

Pour un réseau multicouche à « m » entrées et « n » sorties, composé de L couches (couches cachées et couche de sortie), la somme de l'unité cachée [31,23] est :

$$S_j^L = \sum_{i=1}^n W_{ji}^L x_i + \theta_j^L \quad (2.13)$$

L : Indice de la couche cachée.

W : Poids de la connexion entrée.

$\theta_j^L$ : Bais ou la valeur de seuil interne du neurone.

$f$  : Est la fonction d'activation de ce neurone  $j$  tel que

$$I_j = f^L(S_j^L) = \sum_{j=1}^n W_{ji}^L x_1 + \theta_j^L \quad (2.14)$$

Les équations de la couche de sortie sont :

$$S_K^R = \sum_{j=1}^n w_{Kji}^R I_j + \theta_K^R \quad (2.15)$$

$$R_K = f_K^R(S_K^R) \quad (2.16)$$

$K$  : Numéro de neurone de la couche de sortie.

L'objectif de rétro-propagation est l'adaptation des paramètres  $W_{ij}$  de telle façon qu'on minimise la fonction de coût donnée par :

$$E_p = \frac{1}{2} \sum_{K=1}^T E = \frac{1}{2} \sum_{K=1}^T (\delta)^2 \quad (2.17)$$

Avec:

$$\delta_K = Y_K - R_K \quad (2.18)$$

Où :

$Y_K$ : La sortie désirée.

$R_K$ : La sortie actuelle du réseau.

$T$  : Longueur de l'ensemble d'apprentissage.

### 9.1. Adaptation des poids

Après avoir calculé la sortie  $R_K$  et l'erreur  $E$  correspondante a l'ensemble des entrées à partir des équations (2.16) et (2.17), les poids du réseau Sont alors ajuster par la méthode du gradient :

$$W_{ji}^L(n+1) = W_{ji}^L(n) + \Delta W_{ji}^L(n) \quad (2.19)$$

$$\Delta W_{ji}^L(n) = \mu \frac{\partial E}{\partial W_{ji}^L(n)} \quad (2.20)$$

Avec :

$N$ : Numéro de l'itération.

$M$  : Pas d'apprentissage représentant la vitesse de convergence, sa valeur est généralement choisie expérimentalement ( $0 < \mu < 1$ )

-Si  $\mu$  est trop petit, la convergence est lente mais la direction de la descente est optimale.

-Si  $\mu$  est trop grand, la convergence est rapide mais la précision est médiocre.

L'application des poids du réseau est faite tout d'abord pour la couche de sortie puis pour les couches cachées on a :

a) pour la couche de sortie

D'après (II-17)

$$E_p = \frac{1}{2} \sum_{K=1}^T (Y_K - R_K)^2 \quad (2.21)$$

La dérive de l'erreur  $E_p$  par rapport au poids synaptique  $W_K^R$  nous donne :

$$\frac{\partial E_p}{\partial w_{kj}^L} = (Y_K - R_K) \frac{\partial R_K^L}{\partial s_K^L} \frac{\partial s_K^L}{\partial w_{kj}^L} \quad (2.22)$$

Le dernier facteur de (II.22) est :

$$\frac{\partial s_K^R}{\partial w_{kj}^L} = \frac{\partial}{\partial w_{kj}^L} [W_{kj}^R I_j + \theta_K^L] = I_j \quad (2.23)$$

En combinant (II-21) et (II-22) ; on a pour l'opposé du gradient :

$$-\frac{\partial E_p}{\partial w_{kj}^L} (Y_K - R_K) f_K^R(S_K^R) I_j \quad (2.24)$$

L'amplitude de la variation du poids étant proportionnelle à l'opposé du gradient, les poids de couche de sortie son renouvelés selon :

$$W_{kj}^R(n+1) = W_{kj}^R(n) + \Delta W_{kj}^R(n) \quad (2.25)$$

$$\Delta W_{kj}^R = \mu (Y_K - R_K) f_K^R(S_K^R) I_j \quad (2.26)$$

Il y a deux formes de fonctions de sorties qui nous intéressent ici :

- $f_K^R(S_K^R) = S_K^R$
- $f_K^R(S_K^R) = [1 + \exp(-S_K^R)]^{-1}$

La première est une fonction linéaire. La deuxième appelée sigmoïde. Pour la fonction linéaire la dérivée est :  $f_K^R = 1$

L'équation (II.14) devient :

$$W_{kj}^R(n+1) = W_{kj}^R(n) + \mu (Y_K - R_K) f_K^R(S_K^R) I_j \quad (2.27)$$

Pour une fonction d'activation type sigmoïde, la dérivée est :

$$f_K^R = f_K^R(1 - f_K^R) = R_K(1 - R_K) \quad (2.28)$$

L'équation (II.25) devient :

$$W_{kj}^R(n+1) = W_{kj}^R(n) + \mu (Y_K - R_K) (1 - R_K) R_K I_j \quad (2.29)$$

En définissant la quantité :

$$\mathbf{S}_K^R = (\mathbf{Y}_K - \mathbf{R}_K) \mathbf{f}_K^R(\mathbf{S}_K^R) \quad (2.30)$$

Nous pouvons alors écrire l'équation d'adaptation des poids sous une forme indépendante de la fonction de sortie :

$$\mathbf{W}_{Kj}^R(\mathbf{1} + \mathbf{n}) = \mathbf{W}_{Kj}^R(\mathbf{n}) + \mu \delta_K^R \mathbf{I}_j \quad (2.31)$$

**b) pour les couches cachées**

Nous devons répéter le même calcul que pour la couche de sortie. L'erreur total doit être rapporté d'une façon ou d'une autre aux sorties de la couche cachée :

$$\mathbf{E}_p = \frac{1}{2} \sum_{k=1}^T (\mathbf{Y}_K - \mathbf{R}_K)^2 \quad (2.32)$$

$$= \frac{1}{2} \sum_{k=1}^T (\mathbf{Y}_K - \mathbf{f}_K^R(\mathbf{S}_K^R))^2 \quad (2.33)$$

$$= \frac{1}{2} \sum_{k=1}^T (\mathbf{Y}_K - \mathbf{f}_K^R(\sum \mathbf{W}_{Kj} \mathbf{I}_j + \boldsymbol{\theta}_K^R))^2 \quad (2.34)$$

Dans les équations (2-13) et (2-14), on constate que dépend des poids de la couche cachée, nous pouvons exploiter ce fait pour calculer le gradient de par rapport aux poids de la couche cachée :

$$\frac{\partial E_p}{\partial W_{ji}^L} = \frac{1}{2} \sum_k \frac{\partial}{\partial W_{ji}^k} (\mathbf{Y}_K - \mathbf{R}_K)^2 \quad (2.35)$$

$$= - \sum_k (\mathbf{Y}_K - \mathbf{R}_K) \frac{\partial \mathbf{R}_K}{\partial \mathbf{S}_K^R} \frac{\partial \mathbf{S}_K^R}{\partial I_j} \frac{\partial I_j}{\partial S_j^L} \frac{\partial S_j^L}{\partial W_{ji}^L} \quad (2.36)$$

Les facteurs de l'équation (2-31) peuvent être calculés à partir des équations précédentes pour donner :

$$\frac{\partial E_p}{\partial W_{ji}^L} = - \sum_K (\mathbf{Y}_K - \mathbf{R}_K) \mathbf{f}_K^R(\mathbf{S}_K^R) \mathbf{W}_{Kj}^R \mathbf{f}_j^L(\mathbf{S}_j^L) \mathbf{X}_j \quad (2.37)$$

Les poids des couches caches sont adaptés proportionnellement à l'opposé de l'équation (2.37).

$$\Delta \mathbf{W}_{Kj}^R = \mu \mathbf{f}_j^L(\mathbf{S}_j^L) \mathbf{X}_j \sum (\mathbf{Y}_K - \mathbf{R}_K) \mathbf{f}_K^R(\mathbf{S}_K^R) \mathbf{W}_{Kj}^R \quad (2.38)$$

Si on utilise la définition de, l'équation (2-38) devient :

$$\Delta \mathbf{W}_{Kj}^R = \mu \mathbf{f}_j^L(\mathbf{S}_j^L) \mathbf{X}_j \sum (\delta_K^R) \mathbf{W}_{Kj}^R = \mu \delta_j^L \mathbf{X}_j \quad (2.39)$$

Avec :

$$\delta_j^L = \mu \mathbf{f}_j^L(\mathbf{S}_j^L) \mathbf{X}_j \sum \delta_K^R \mathbf{W}_{Kj}^R \quad (2.40)$$

On constate que chaque adaptation des poids dans la couche cachée dépend de l'erreur totale de la couche de sortie ce qui conduit à la notion de rétro propagation.

L'équation d'adaptation des poids dans ce cas est :

$$W_{ji}^L(n + 1) = W_{ji}^L(n) + \mu \delta_j^L X_j \quad (2.41)$$

## 9.2 Algorithme de la rétro-propagation

**Etape 1 :** Initialiser les poids et les seuils internes des neurones à des petites valeurs aléatoires.

**Etape 2 :** Présenter le vecteur d'entrée et de sortie désirée.

**Etape 3 :** Calculer :

- La somme des entrées des neurones de la couche cachée en utilisant l'expression (2.6)
- Les sorties de neurones de la couche cachée en utilisant l'expression (2.13)
- La somme des entrées des neurones de la couche de sortie en utilisant l'expression (2.14)
- Les sorties des réseaux en utilisant l'expression (2.15).

**Etape4 :** Calculer l'erreur pour les neurones de la couche de sortie en utilisant l'expression (2.30).

**Etape5 :** Réinjecter l'erreur de sortie en utilisant l'expression (2.40)

**Etape6 :** Ajuster :

- Les poids de la couche de sortie en utilisant l'expression (2.31)
- Les poids de la couche cachée en utilisant l'expression (2.41)

**Etape 7 :** Calculer E l'erreur en utilisant l'expression (2.21)

**Etape8 :** Si la condition sur l'erreur E —  $E_p < \varepsilon$  est atteinte, aller à l'étape 9 sinon aller à l'étape 6 et refaire le calcul pour un autre époque.

**Etape 9 :** FIN.

## 10. Les applications des méthodes neuronales

Les grands domaines d'application des réseaux de neurones découlent naturellement de propriétés énoncées précédemment. Nous présentons dans les sections suivantes quelques exemples pour montrer le vaste étendu de leur applicabilité [19].

- **La régression non linéaire, ou modélisation de données statiques**

Une immense variété de phénomènes statiques peut être caractérisée par une relation déterministe entre des causes et des effets ; les réseaux de neurones sont de bons candidats pour modéliser de telles relations à partir d'observations expérimentales, sous réserve que celles-ci soient suffisamment nombreuses et représentatives [19].

- **La modélisation de processus dynamiques non linéaires**

Modéliser un processus consiste à déterminer un ensemble d'équations mathématiques qui décrivent le comportement dynamique du processus, c'est-à-dire l'évolution de ses sorties en fonction de l'évolution de ses entrées .Ce problème peut être avantageusement résolu par un réseau de neurones, si le phénomène que l'on désire modéliser est non linéaire [19].

- **La commande de processus**

La commande d'un processus consiste à concevoir un système comprenant un organe qui calcule la commande à appliquer au processus pour assurer un comportement dynamique spécifié par des cahiers de charges: régulation au voisinage d'un point de fonctionnement, poursuite d'une trajectoire de consigne, commande optimale....

L'ensemble commande / processus peut donc être considéré comme un système qui réalise une fonction (non linéaire) qu'un réseau de neurone peut approcher [19].

- **La classification**

Une autre grande catégorie de problème industriel consiste à attribuer de façon automatique un objet à une classe, parmi d'autres classes possibles. Et en raison de leur propriété d'approximateurs universels, les réseaux de neurones sont capables d'estimer de manière précise la probabilité d'appartenance d'un objet inconnu à une classe parmi plusieurs possibles. Aujourd'hui, les réseaux de neurones artificiels donnent de bons résultats dans différent

domaines d'applications, mais il existe des problèmes les techniques d'apprentissages classiques ne peuvent pas être résolues.

### 11. Les Machines d'apprentissage extrême (ELM)

L'apprentissage mécanique et intelligence artificielle n'ont apparemment jamais été aussi critiques et importantes pour les applications réelles telles qu'elles sont dans l'ère autonome et de grande taille des données d'aujourd'hui. Le succès de l'apprentissage machine et de l'intelligence artificielle repose sur la coexistence de trois conditions nécessaires: environnements informatiques puissants, riches et / ou de grandes données et des techniques efficaces d'apprentissage (algorithmes). La machine d'apprentissage extrême (ELM) comme une technique d'apprentissage émergente fournit une efficacité solutions unifiées aux réseaux avancés y compris mais sans s'y limiter (à la fois Multi-caché) réseaux de neurones, base radiale les réseaux de fonctions (RBF) [32].

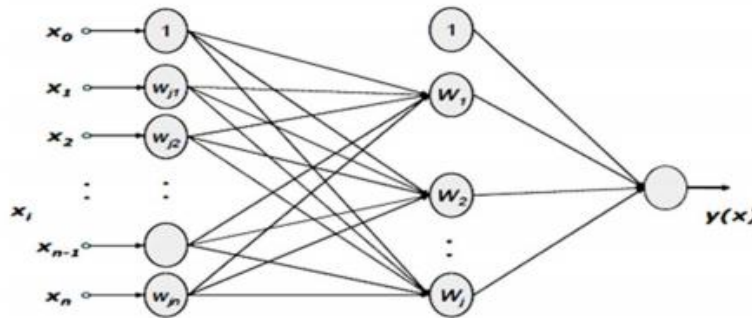


Figure .II.11: Réseau de neurones à apprentissage machine extrême (ELM).

#### 11.1. Formulation mathématique de l'ELM

La fonction de sortie d'ELM est représentée dans l'équation suivant [33] :

$$\sum_{i=1}^{\tilde{N}} \beta_i g(W_i \cdot X_j + b_i) = t_i, j = 1, \dots, N(2.42)$$

Où m : est le nombre de nœuds cachées et B est définie de la manière suivante:

$$B = [B_1 \dots \dots B_m]$$

B est le vecteur des pondérations de sortie entre la couche cachée de n nœuds et le nœud de sortie, et x est le vecteur de sortie de la couche cachée :

$$X = [X_1 \dots \dots X_n]$$

f: Fonction d'activation

$$H\beta = Y, \quad \beta=B^T \quad (2.43)$$

## 11.2. Caractéristiques d'ELM

Depuis l'apparition de réseau ELM en 2006, il a été exploité dans plusieurs applications, telle que la reconnaissance de forme, la reconnaissance d'action humaine, la classification des images, en raison de ces différentes caractéristiques que nous citons ci-dessous [34,35] :

- L'algorithme d'apprentissage ELM est beaucoup plus simple que la plupart des algorithmes d'apprentissage des réseaux de neurones feedforward.
- De nombreuses fonctions d'activation non linéaire peuvent être utilisées dans l'ELM, particulièrement, la fonction sigmoïde.
- Les poids d'entrée et les biais de la couche cachée, des réseaux de neurones perceptrons multicouches de type feedforward, peuvent être affectés aléatoirement sila fonction d'activation est infiniment différentiable.
- Un réseau de neurones perceptrons multicouches de type feedforward, peut être simplement considéré comme un système linéaire et les poids de sortie peuvent être déterminés analytiquement par une simple opération d'inverse généralisé de matrice de sortie de couche cachée.
- Le réseau ELM tend à avoir la plus petite norme de l'erreur quadratique, pour une meilleure performance de généralisation.
- La vitesse d'apprentissage de l'ELM peut être des milliers de fois plus rapide que les algorithmes classiques, comme la rétro propagation du gradient.
- ELM peut résoudre tous les problèmes de régression avec la précision souhaitée et adopté aux problèmes de classification.

- Les algorithmes classiques d'apprentissage basés sur le gradient peuvent faire face à plusieurs problèmes comme les minimums locaux, l'ELM tend à atteindre les solutions sans ces questions triviales.
- Il n'y a pas de paramètre à fixer sauf le nombre de neurones cachés.
- Le temps d'apprentissage de ELM est consacré au calcul du pseudo inverse de Moore-Penrose de la matrice de sortie de couche cachée  $H$ .
- Une des caractéristiques les plus marquées d'ELM est que les poids de la couche cachée n'ont pas besoin d'être ajustés.
- La performance de généralisation de ELM est très stable sur un large nombre de neurones cachés, bien que les performances de généralisation tendent à s'aggraver lorsque trop peu ou trop de neurones cachés.

### 11.3. L'algorithme de l'ELM

Étant donné un ensemble d'observations  $n$  entrées,  $m$  sorties,  $\tilde{N}$  neurones cachés ( $k = 1 \dots \tilde{N}$ ),

La fonction d'activation  $f(x)$  et  $N$  échantillons de données d'apprentissage [32].

**Étape 01 :** Assigner aléatoirement des valeurs aux poids d'entrée et aux biais.

**Étape 02 :** Calculer la matrice de la sortie de couche cachée.

**Étape 03 :** Calculer les poids de sortie, telle que.

### Conclusion

Ce chapitre a fait l'objet d'une présentation des techniques de développement des prédicteurs intelligents vus comme moyen intéressant de classer les MA. Les techniques utilisées pour développer ces régresseurs sont des méthodes paramétriques non linéaires qui relient des variables d'entrée à des variables de sortie, appliquées à la régression et la classification. Des généralités sur les réseaux RNAs à décrits à cet effet. Une attention particulière à la technique ELM, a été toutefois donnée en rappelant les concepts les plus importants pour sa mise en œuvre en tant qu'outil de régression et de classification.

# CHAPITRE III

## SIMULATION

### INTRODUCTION

Ce dernier chapitre est consacré à une étude en simulation et vise principalement l'évaluation des performances des techniques proposées comme étant une solution dans la conception de régresseurs intelligents utilisés dans le domaine de surveillance de la maladie d'Alzheimer. Ce travail présente donc une contribution à l'étude et au développement de classifier dans ce domaine. Une étude comparative entre les techniques MLP et ELM en termes de précision et d'autres paramètres de régression et de classification est alors présentée. L'objectif est de mettre en place un système intelligent à base des techniques d'apprentissage statistiques afin de prendre une décision adaptée au contrôle et au suivi des maladies d'Alzheimer.

### 1. BASE DE DONNÉES UTILISÉE

Afin de bien évaluer notre système proposé, deux différentes bases de données ont été exploitées pour sortir les performances globales. Ces données viennent de sources biomédicales réelles. Dans la suite, nous décrivons les caractéristiques de chaque base de données.

Données d'IRM cross-sectionnelles chez les jeunes adultes, les adultes d'âge moyen, les adultes âgés non déments et les adultes âgés déments : Cet ensemble se compose d'une collection transversale de 416 sujets âgés de 18 à 96 ans. Pour chaque sujet, trois ou quatre examens IRM pondérés en T1 obtenus en une seule séance sont inclus. Les sujets sont tous droitiers et comprennent à la fois des hommes et des femmes. 100 des sujets de plus de 60 ans ont reçu un diagnostic clinique de maladie d'Alzheimer (MA) très légère à modérée. De plus, un ensemble de données sur la fiabilité comprend 20 sujets non déments qui ont été photographiés lors d'une visite subséquente dans les 90 jours suivant leur séance initiale [36].

Données longitudinales d’IRM chez les adultes âgés non atteints et déments :Cet ensemble se compose d’une collection longitudinale de 150 sujets âgés de 60 à 96 ans. Chaque sujet a été scanné lors de deux visites ou plus, séparées par au moins un an pour un total de 373 séances d’imagerie. Pour chaque sujet, trois ou quatre examens IRM pondérés en T1 obtenus en une seule séance sont inclus. Les sujets sont tous droitiers et comprennent à la fois les hommes et les femmes. 72 des sujets ont été caractérisés comme non déments tout au long de l’étude. 64 des sujets inclus ont été caractérisés comme déments au moment de leurs premières visites et sont restés ainsi pour les examens subséquents, y compris 51 personnes avec la maladie d’Alzheimer légère à modérée. Quatorze autres sujets ont été qualifiés de non déments au moment de leur première visite et ont par la suite été qualifiés de déments lors d’une visite ultérieure [36].

**Tableau III.1 :** Informations d'attributs de la base Alzheimer.

N° de parameter	Abbreviation	Signification
01	Sex	1 = male; 0 = femelle
02	Age	âge en années
03	EDUC	Education Level
04	SES	Socio Economic status
05	MMSE	Mini Mental State Exomination
06	CDR	Clinical Dementia Rating(0 à 1)
07	Etiv	Estimated total
08	NWBv	Normalize Whole Brain Volume
09	ASF	Atlas Scaling Factor

## 2. Normalisation des données

La normalisation des données permet de s’affranchir des différences de « normes » des variables. En effet, des variables avec des grandes valeurs peuvent avoir une plus grande influence que des variables avec des petites valeurs, sans pour autant être plus significatives. La normalisation est un outil de prétraitement de données utilisé dans le système d'exploration de données [37]. Un

attribut d'un ensemble de données est normalisé en mettant ses valeurs à l'échelle afin qu'elles se situent dans une plage spécifiée de petite taille, telle que 0 à 1[38].

### 2.1. Normalisation avec Min-Max

La normalisation est particulièrement utile pour les algorithmes de classification et de régression. Dans le but d'optimiser les modèles de prédiction, nous introduisons la normalisation des données par la méthode min-max. Cette technique permet de fournir une transformation linéaire sur la plage de données originale. Le rééchelonnement est souvent accompli en utilisant une formule d'interprétation linéaire, tels que [38] :

- La fonction mathématique :

$$X_{i_N} = \frac{X_i - \min(X)}{\max(X) - \min(X)} \quad (3.1)$$

Avec :

$X_{i_N}$  : la valeur normalisée ;

$X = [X_1, X_2, \dots, X_i]$  : le vecteur à normaliser ;

$\min(X)$  : la valeur minimale de  $X$  ;

$\max(X)$  : la valeur maximale de  $X$ .

Lorsque la normalisation min-max est appliquée, chaque entité se trouvant dans la nouvelle plage de valeurs restera la même. La normalisation min-max a l'avantage de préserver exactement toutes les relations dans les données.

### 2.2. Normalisation avec Standard-scaler

Vous standardise une caractéristique en soustrayant la moyenne, puis en la redimensionnant à la variance unitaire. La variance unitaire signifie diviser toutes les valeurs par l'écart type. Standard Scaler ne répond pas à la définition stricte de l'échelle que j'ai présentée précédemment.

Standard Scaler donne une distribution avec un écart type égal à 1. La variance est égale à 1 également, car variance = écart type au carré. Et 1 au carré = 1.

Standard Scaler fait la moyenne de la distribution à 0. Environ 68% des valeurs seront comprises entre -1 et 1[39].

- La fonction mathématique [40] :

$$Z = \frac{x-\mu}{\sigma} \quad (3.2)$$

Avec :

$$\text{Mean: } \mu = \frac{1}{N} \sum_{i=1}^N (x_i) \quad (3.3)$$

$$\text{Standard déviation: } \sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (3.4)$$

### 3. Apprentissage et Test

Les méthodes de reconnaissance de formes telle que le MLP (PMC), ELM appliquées à la catégorisation (classification) des données. Elles sont utilisées entre les systèmes de limitation de hauts niveaux, et fondées sur l'analyse de données expérimentales. Il est ensuite dictatorial d'émaner au spicilège delà adresse la davantage adaptée, comme de science l'adjoindre complémentaiement pour un mascarade de diagnostic. Pour une hypocoristique des performances des couple modèles prématurément testés, définissons copieux mesures statistiques.

Soient :

- **Training:** la phase d'apprentissage (75%).
- **Testing:** la phase de test (25%).
- **Taux de classification:**  $T = \frac{NB}{N} * 100\%$

**NB:** nombre des échantillons les biens classées.

**N:** nombre totale des échantillons.

## 4. Prédiction

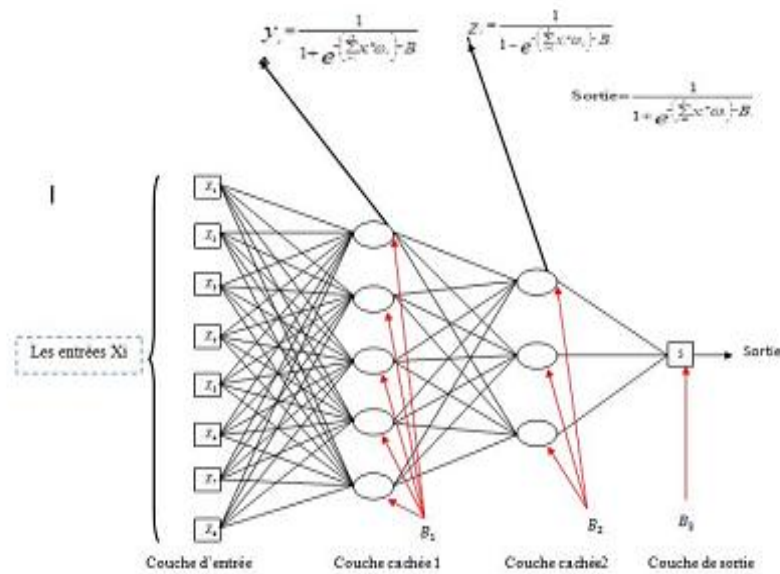
### 4.1. Prédiction avec MLP

L'apprentissage supervisé du modèles neuronaux MLP (figure III.1) consiste à déterminer les poids de celui-ci qui minimisent sur l'ensemble des données de la base d'apprentissage, les écarts

entre les valeurs de sortie désirée  $y_d$  et les valeurs de sortie calculée  $y_c$ . Ceci consiste à trouver le minimum du critère quadratique suivant [19] :

$$C_w = \frac{1}{N} \sum_{i=1}^N (y_{c_i} - y_{d_i})^2 \quad (3.5)$$

Où N est le nombre d'exemples de la base d'apprentissage.



**Figure.III.1 :** Architecture du réseau MLP.

Différentes architectures sont testées et évaluées pour déterminer le meilleur modèle pour les méthodes utilisées. On présente dans le tableau III.2, les résultats correspondant à la base de données utilisée. Les paramètres tels que le Nombre des Couches Caché (NCC) et le Nombre Neurone des Couches Caché (NNCC), le Taux de test(Taux\_test) sont indiqués pour les réseaux testés. En matière de reconnaissance.

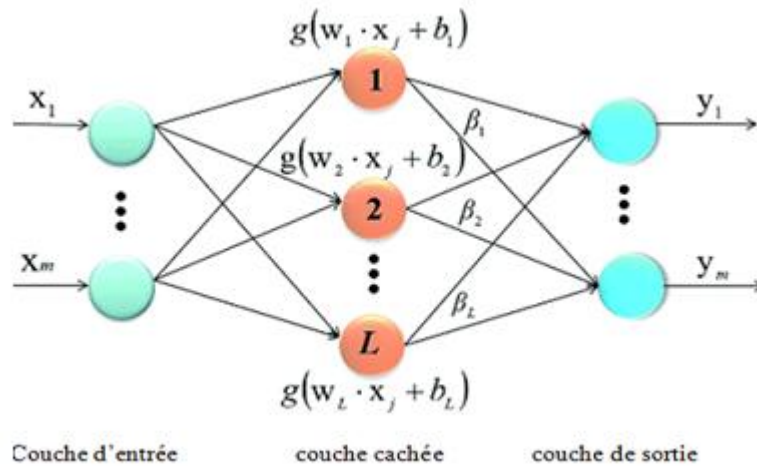
**Tableau.III.2 : Résultats de test de modèle MLP**

Méthode de Normalisation	NCC	NNCC	Taux_test (%)
Standard scaler	1	10	78
	2	10 – 8	82
	3	16 – 10 – 8	85,39
	3	32 – 16 – 8	85,39
Min Max	1	10	82
	2	10 – 8	78
	3	16 – 10 – 8	62
	3	64 – 16 – 8	86,51

D'après les résultats obtenus et pour l'ensemble des tests, un test supérieur de 86 % a été obtenu avec les réseaux testés des deux méthodes de normalisation. Nous constatons que les tests d'apprentissage offerts par la normalisation « Standard scaler » sont souvent les meilleurs. Cependant, nous devons souligner que le nombre de neurones cachés d'un réseau MLP avec une normalisation « Min-Max » sont beaucoup plus importants que ceux des réseaux de neurones classiques, et c'est la complexité du modèle. Une amélioration positive de taux de test se produit lorsqu'il y a une association accrue de neurones dans la couche de caché.

#### 4.2. Prédiction avec ELM

ELM est un algorithme d'apprentissage des réseaux de neurones à une seule couche cachée de type feedforward (voir la figure III.2). Dans le réseau ELM, le nombre de neurones de la couche cachée, est défini de façon adaptative, et les poids des neurones de la couche cachée sont assignés de façon aléatoire et restent constants. Cependant les poids des neurones de la couche de sortie, contrairement aux réseaux de neurones classiques, sont obtenus par la méthode des moindres carrés. Par conséquent, l'ELM génère une solution optimale unique. Le réseau ELM est un réseau de neurones d'architecture très simple et le seul paramètre à fixer est le nombre de neurones dans la couche cachée, en plus il est efficace pour la classification de données [35,41, 42].



**Figure.III.2 :**Architecture du réseau ELM.

Avec :

$\tilde{N}$ : le nombre de neurones de la couche cachée.

$X_N$ : les données d'entrée.

$W_{\tilde{N}}$ : les poids reliant le  $L^{i\grave{e}me}$  neurone caché et les neurones d'entrée.

$\beta_{\tilde{N}}$ : les poids reliant le  $L^{i\grave{e}me}$  neurone cachée et les neurones de sortie.

$b_{\tilde{N}}$ : le biais du  $L^{i\grave{e}me}$  neurone cachée.

$W_{\tilde{N}} \cdot X_N$ : désigne le produit scalaire entre  $W_{\tilde{N}}$  et  $X_N$ .

$g$ : est la fonction d'activation.

$t_{\tilde{m}}$ : Les sorties de réseau.

Dans notre cas, nous entraînerons ELM pour différentes valeurs de les paramètres tels que le Nombre Neurone de Couche Caché (NNCC) pour déterminer le meilleur modèle. On présente dans le tableau.III.3 les résultats d'apprentissage et de test correspondant aux paramètres d'évaluation tels que le taux de test (Taux\_test).

Tableau.III.3 : Résultats de test de modèle ELM

Méthode de Normalisation	NNCC	Taux_test(%)
<b>Standard scaler</b>	30	78
	<b>40</b>	<b>83,76</b>
	45	80,34
	50	82,9
	100	75,21
	150	75,21
	200	64
	250	75,21
<b>Min Max</b>	<b>40</b>	<b>79,48</b>
	45	76,92
	50	78,63
	100	76,06
	150	69,23
	200	63,24
	250	44

D'après les résultats obtenus et pour l'ensemble des tests, un test supérieur de 86 % a été obtenu avec les réseaux testés des deux modèles. Nous constatons que les tests offerts par la normalisation « Standard scaler » sont souvent les meilleurs. Nous constatons que le taux de test du réseau ELM avec une standardisation minimale à de mauvais résultats.

### 4.3. Comparaison

Tableau.III.4 : Temps d'entraînement des MLP et ELM avec deux méthodes de normalisation.

Les Réseaux Méthode de Normalisation	MLP(%)	ELM(%)
<b>Le Taux de test avec standard scaler</b>	85,39	83,76
<b>Le Taux de test avec Min-Max</b>	86,51	79,48

On résume dans le tableau III.4 un état comparatif des caractéristiques liées aux solutions envisageables dans l'utilisation de l'un des modèles (MLP et ELM) dans un système de classification.

**Tableau.III.5** : Tableau comparatif des modèles (MLP et ELM).

Propriétés	MLP	ELM
<b>Algorithme</b>	Apprentissage et généralisation	Apprentissage et généralisation
<b>Apprentissage</b>	Nombre de poids	Nombre de poids
<b>Ajustement des poids</b>	N'est pas stable	N'est pas stable
<b>Paramètres d'apprentissage</b>	Trop de paramètres	Trop de paramètres
<b>Test d'apprentissage</b>	86,51%	83,76%
<b>Inconvénients</b>	<ul style="list-style-type: none"> <li>- Classe les éléments qui n'appartiennent à aucune classe à la classe la plus proche</li> <li>- Le nombre de neurones et de couches caches est indéfini</li> </ul>	<ul style="list-style-type: none"> <li>- Nécessite une grande capacité de calcul</li> </ul>
<b>Avantages</b>	<ul style="list-style-type: none"> <li>- Accepte les données bruitées et la classification non-linéaire</li> <li>- Représentation globale de l'espace</li> <li>- Architecture simple</li> </ul>	<ul style="list-style-type: none"> <li>- La vitesse d'apprentissage de l'ELM est extrêmement rapide.</li> <li>- ELM est meilleures performances de généralisation que l'apprentissage classique</li> <li>- L'algorithme d'apprentissage ELM semble Beaucoup plus simple que la plupart des algorithmes d'apprentissage</li> </ul>

## 5. Discussion des résultats

### 5.1. Analyse et évaluation

Le problème de la généralisation est souvent vu sous trois perspectives différentes. Dans la première, la taille du réseau est fixée (en accord avec la complexité du problème). Dans notre cas, plusieurs neurones d'entrée, où chaque neurone représente un paramètre descripteur et un neurone réservé en sortie délivrant la décision. La question qui se pose est de : combien d'exemples d'apprentissage sont nécessaires pour atteindre une bonne généralisation ?

Cette perspective est intéressante dans les applications où l'on a la possibilité d'acquérir autant d'exemples que l'on veut. Dans le cas d'un système de diagnostic multi sensorielle (notre cas précis), on peut acquérir autant d'exemples que l'on veut, cependant une base de 1000 vecteurs par exemple est-elle suffisante ? La deuxième perspective c'est quand nous supposons que le nombre d'exemples d'apprentissage est fixé ; la question qui se pose dans ce cas est : quelle est la taille du réseau qui donne la meilleure généralisation de ces données ? Pour cette application, quel est le réseau pris parmi les différents réseaux testés qui donne la meilleure généralisation ? Est ce un réseau à une seule couche cachée ? À deux couches cachées ? Ou bien à trois couches cachées ? On est conduit à adopter finalement ce point de vue puisqu'on est devant l'impossibilité d'avoir une base de connaissance aussi complète qu'elle soit. Un enrichissement continu de cette base avec le temps est pratiquement indispensable. Il importe alors dans cette situation, de déterminer la taille du réseau qu'il faut pour décrire au mieux les données en notre possession. Cependant, tous les différents réseaux validés peuvent acquérir des données d'apprentissage, et l'erreur d'entraînement est acceptable presque dans tous ces réseaux. La variance de l'estimation due à la taille finie de l'échantillon induit un écart entre la capacité réelle de généralisation et la capacité estimée (risque empirique). [19].

Concernant l'utilisation de la technologie MLP ; Nous avons constaté que cette méthode donne de très bons résultats dans le cas de la classification binaire après normalisation des données. Les solutions trouvées se basent exclusivement sur les exemples (base de données) fournis en entrée. On peut dire à partir des résultats obtenus qu'un test d'entraînement élevé, après sélection des caractéristiques, confirme clairement l'adéquation de la technique pour ce type d'application.

## Conclusion

Ce dernier chapitre nous présente une étude de simulation liée à l'application des deux techniques d'apprentissage appliquées dans le domaine médical. Grâce à laquelle nous avons pu valider et évaluer les performances de chacune de ces méthodes. Nous avons mené une étude comparative dans le but de sélectionner avec soin la méthode d'application la plus appropriée. Les informations sur l'auto-reconnaissance et le taux de test étaient les facteurs pertinents qui nous ont permis d'évaluer les méthodes étudiées. La discussion des résultats obtenus a permis de choisir la technique MLP qui a conservé ses qualités et ses avantages adaptés au problème posé. Cette technique a donné de bons résultats et a démontré l'efficacité de cette approche.

# Conclusion General

Les travaux présentés dans ce mémoire ont été consacrés à la mise en œuvre de deux techniques d'apprentissage appliquées à la reconnaissance des formes sur le terrain médical. Cette étude découle des développements technologiques importants qui ont été enregistrés. Ces dernières années, avec l'objectif et l'intérêt d'une surveillance moderne et plus efficace patients. À cette fin, notre travail modeste peut être considéré comme une contribution à solutions proposées pour résoudre des problèmes d'intérêt stratégique nationales, en utilisant des outils modernes basés sur des techniques avancées de la santé.

L'objectif est de valider et d'évaluer les performances des techniques d'apprentissage de chacune des méthodes présentées à savoir MLP et ELM comme étant une solution pour le diagnostic médical en particulier, à la conception des modèles de classification à partir des paramètres descripteurs de différentes maladies étudiées telles que: La maladie d'Alzheimer. Cette étude a été précédée plusieurs travaux nous ont aidés à choisir les caractéristiques qui nous ont permis de déterminer les relations entre les variables de description afin de classifier la maladie.

Cette étude était structurée autour de trois chapitres essentiels. Le premier a été consacré à des généralités concernant la maladie d'Alzheimer, les différentes étapes de diagnostic ainsi que les causes, les différentes formes et les méthodes de traitements de la maladie d'Alzheimer. Dans le deuxième chapitre, il a été question des régresseurs en particulier, comme étant un moyen intéressant pouvant contourner l'usage de certains classifieurs intelligents. Deux modèles (Neuronal et ELM) à base de la conception de ces régresseurs ont été étudiés et évalués. Enfin, le troisième et dernier chapitre a été dédié à une application réelle justifiant la conception et la mise en œuvre d'un régresseur pour la classification de la maladie d'Alzheimer. La technique la mieux adaptée pour la conception de ce régresseur a été choisie sur la base d'une étude d'évaluation préliminaire des performances de chacune des méthodes présentées au chapitre deux.

D'après les résultats obtenus, il apparaît que sur le plan décisionnel, les deux modèles ont présenté des résultats acceptables pour une simulation sur des données biomédicales spécialement ELM et MLP. Le modèle MLP est plutôt mieux placé du point de vue taux de test, ce qui lui confère l'avantage d'une intégration dans un système de diagnostic dynamique.

## Conclusion general

Le principal problème dans l'application de ce modèle est d'obtenir la « meilleure » base de données. Cela implique évidemment le nombre et le type d'exemples à utiliser dans la bibliothèque d'apprentissage. Comme mentionné précédemment, dans ce cas, la présence d'experts (ou de systèmes experts) est indispensable.

Les perspectives d'application de cette méthode sont encore prometteuses. La prise de décision du système peut être améliorée en utilisant de nouveaux paramètres d'entrée. Il convient de noter que la sensibilité du domaine aux menaces imprévues nécessite des efforts accrus pour maximiser l'immunité du système et d'autres améliorations pour minimiser les risques pour la santé humaine. Enfin, l'application offre une alternative prometteuse pour notre pays à l'avenir, qui peut surveiller intelligemment, automatiquement et efficacement l'état du patient.

## **Résumé**

Les progrès technologiques ont facilité l'acquisition et la collecte de nombreuses données. Ces données peuvent être utilisées comme aide à la décision, conduisant au développement d'outils capables de les analyser et de les traiter. Les systèmes de soutien diagnostique sont considérés comme essentiels dans de nombreuses disciplines, basés sur des techniques d'intelligence artificielle, mais les problèmes les plus intéressants sont souvent basés sur de grandes données. Ces problèmes se rapportent à des situations où nous avons peu d'observations, alors que le nombre de variables explicatives est très élevé. La sélection des variables est devenue l'objet qui attire l'attention de nombreux chercheurs ces dernières années, cette sélection permet d'identifier et d'éliminer les variables qui pénalisent la performance d'un modèle complexe dans la mesure où elles peuvent être bruyantes, redondant ou non pertinent. En outre, l'identification des variables pertinentes facilite l'interprétation et la compréhension des aspects liés à l'application; ainsi, il permet d'améliorer la performance prédictive des méthodes de classification et de ne pas tenir compte de l'organisme nuisible de la dimensionnalité élevée de ces données.

Actuellement, l'approche par filtre est couramment utilisée pour analyser les données biologiques, cette approche consiste à examiner la sélection des variables avant le processus d'apprentissage et ne conserve que des caractéristiques informationnelles.

L'objectif de ce travail est de contribuer à la recherche et au développement de systèmes innovants d'aide au diagnostic médical. Ce travail de simulation se concentre sur l'application de techniques d'apprentissage statistique comme solution à la conception de ces systèmes par la reconnaissance des formes. Dans le cadre de l'apprentissage contrôlé, la sélection des fonctions permet d'obtenir des classificateurs précis.

Afin de mener une étude comparative permettant un choix décisif de la méthode la mieux adaptée à l'application proposée, on évaluera pour les méthodes exposées les paramètres liés au taux de reconnaissance, au temps d'apprentissage et à l'erreur d'entraînement.

***Mots clés*** : Diagnostic médical, Sélection des variables, Classification, RNA, ELM, MLP, PMC Simulation.

## **Abstract**

Advances in technology have made it easier to acquire and collect a lot of data. These data can be used as decision support, leading to the development of tools capable of analyzing and processing them. Diagnostic support systems are considered essential in many disciplines, based on artificial intelligence techniques, but the most interesting problems are often based on big data. These problems relate to situations where we have few observations, while the number of explanatory variables is very high. The selection of variables has become the object that attracts the attention of many researchers in recent years, this selection allows to identify and eliminate the variables that penalize the performance of a complex model insofar as they can be noisy, redundant or irrelevant. In addition, the identification of the relevant variables facilitates the interpretation and understanding of aspects related to the application; thus, it improves the predictive performance of classification methods and ignores the pest of the high dimensionality of this data.

Currently, the filter approach is commonly used to analyze biological data, this approach consists of examining the selection of variables before the learning process and retains only informational characteristics.

The objective of this work is to contribute to the research and development of innovative medical diagnostic aid systems. This simulation work focuses on the application of statistical learning techniques as a solution to the design of these systems through pattern recognition. In controlled learning, the selection of functions results in precise classifiers.

In order to conduct a comparative study allowing a decisive choice of the method best suited to the proposed application, the parameters related to the recognition rate, the learning time and the training error will be evaluated for the methods presented.

## ملخص

جعلت التطورات التكنولوجية من السهل الحصول على الكثير من البيانات وجمعها. يمكن استخدام هذه البيانات كدعم للقرار ، مما يؤدي إلى تطور أدوات قادرة على تحليلها ومعالجتها. تعتبر أنظمة الدعم التشخيصي ضرورية في العديد من التخصصات ، بناءً على تقنيات الذكاء الاصطناعي ، لكن المشكلات الأكثر إثارة للاهتمام غالبًا ما تعتمد على البيانات الضخمة. تتعلق هذه المشاكل بالحالات التي لدينا فيها القليل من الملاحظات ، في حين أن عدد المتغيرات التفسيرية مرتفع للغاية. أصبح اختيار المتغيرات هو الشيء الذي يجذب انتباه العديد من الباحثين في السنوات الأخيرة، ويسمح هذا الاختيار بتحديد المتغيرات التي تعاقب أداء نموذج معقد والقضاء عليها بقدر ما يمكن أن تكون صاخبة أو زائدة عن الحاجة أو غير ذات صلة. بالإضافة إلى ذلك، فإن تحديد المتغيرات ذات الصلة يسهل تفسير وفهم الجوانب المتعلقة بالتطبيق وبالتالي فإنه يحسن الأداء التنبؤي لطرق التصنيف ويتجاهل الأبعاد العالية لهذه البيانات.

حاليًا ، يتم استخدام نهج التصفية بشكل شائع لتحليل البيانات البيولوجية ، ويتكون هذا النهج من فحص اختيار المتغيرات قبل عملية التعلم ويحتفظ فقط بالخصائص الإعلامية.

الهدف من هذا العمل هو المساهمة في البحث والتطوير لأنظمة مساعدة التشخيص الطبي المبتكرة. يركز عمل المحاكاة هذا على تطبيق تقنيات التعلم الإحصائي كحل لتصميم هذه الأنظمة من خلال التعرف على الأنماط. في التعلم المتحكم فيه ، ينتج عن اختيار الوظائف مصنفات دقيقة.

من أجل إجراء دراسة مقارنة تسمح باختيار حاسم للطريقة الأنسب للتطبيق المقترح ، سيتم تقييم المعلمات المتعلقة بمعدل التعرف ووقت التعلم وخطأ التدريب للطرق المقدمه.

# REFERENCES

- [01] A.Brahime, A.Oussama , "approche de sélection des données biomédicales dans un problème de reconnaissance de pathologies en utilisant les support vector machines"
- [02] N. Settouti, A. Hafa, "Approche Filtre pour la sélection des gènes pertinents des données biopuces du Cancer du Côlon" , Université Abou Bekr Belkaid – Tlemcen, 2013.
- [03]- N. BOUDIA, " Le recuit simulé pour la sélection de variables des puces à ADN", Mémoire De Master, université Abdelhamid ibn badis – Mostaganem, 2016.
- [04] Rouaud, Olivier, Jean-Francois Demonet. "La maladie d'Alzheimer et les maladies apparentées." Forum Médical Suisse. Vol. 18. No. 11. EMH Media, 2018
- [05] <https://alzheimer-recherche.org/la-maladie-alzheimer/quest-maladie-dalzheimer/definition-et-chiffres/>
- [06] <https://alzheimer-recherche.org/la-maladie-alzheimer/symptomes-et-diagnostic/diagnostic/>
- [07] <https://www.capretraite.fr/prevenir-dependance/maladie-d-alzheimer/les-tests-de-depistage-d-alzheimer/>
- [08] <https://www.capretraite.fr/prevenir-dependance/maladie-d-alzheimer/les-causes-de-la-maladie-d-alzheimer/>
- [09] <https://www.doctissimo.fr/html/dossiers/alzheimer/11708-alzheimer-differentes-formes.htm>
- [10] Alzheimer's Disease International. 2019. Rapport Mondial Alzheimer 2019 : Attitudes à l'égard de la démence. Londres : Alzheimer's Disease International
- [11] <https://www.vidal.fr/maladies/systeme-nerveux/maladie-alzheimer/traitements-psycho-sociaux.html>
- [12] Breton, Marie-Claude, Mélanie Turgeon, and Cathy Gosselin. "Traitements pharmacologiques de la maladie d'Alzheimer et des maladies apparentées." Revue systématique (2015).
- [13] Mr. Ali DJAIDJA".Etude de la classification supervisée des données environnementales à l'aide de réseaux de neurones de fonctions à base radiales".
- [14] Benkirane, Chaimaa. Modélisation d'un canal à 60 GHz en utilisant les réseaux de neurones perceptron multicouches. Diss. Université du Québec en Abitibi-Témiscamingue, 2013.
- [15] François B. et Michel V. « Les réseaux de neurones artificiels » Presses Universitaire, Paris, 1996..

- [16] Bernard G. « Application de réseaux de neurones artificiels a la reconnaissance au de caractères manuscrits» Thèse de Doctorat, Faculté polytechnique de Mons ,1996.
- [17] Gérard D. « Reducing the complexity of neural nets for industrial applications and biological models » Neuronal Information Processing - From Biological Data to Modelling and Applications (World Scientific, 1998).
- [18] MC Culloch, W., Pitis, W. - « A Logical Calculus for the Ideas Immanent in NervousActivity » - Bulletin Mathematics and Biophysics - 1943, N°. 5 - p. 115-133.:
- [19] CHARIK Khalissa& CHARIK Loubna." Approche Filtre par la sélection dedonnées multi-sensorielles pour l'aide au diagnostic médical".
- [20] Katrina B. "Fuzzy neural predictive control of highly nonlinear ph process Computers chem.Engng, vol21, , 1997".
- [21] M.MEDJILI Fayçal." Modélisation par Réseaux de Neurones Artificiels (RNA) et commande Prédictive non linéaire d'une station de production d'eau froide. mémoire de magister en automatique, université de annaba 2010.
- [22] Jerbi, Jacin. Evaluation des options et gestion des risques financiers par les réseaux de neurones et par les modèles à volatilité stochastique. Diss. Université Panthéon-Sorbonne-Paris I, 2006.
- [23] François B. et Michel V. « Les réseaux de neurones artificiels » PressesUniversitaire, Paris, 1996.
- [24] S.H. WOO, et al., &quot;On-line estimation of key process variables based on kernel partial least squares in an industrial cokes wastewater treatment plant &quot;; J. Hazard. Mater, Vol. 161, pp. 538-544,2009.
- [25] L. Bao, X. Yuan, Z. Ge, &quot;Co-training partial least squares model for semi-supervised soft sensor development &quot;; Chemometrics and Intelligent Laboratory Systems, Vol. 147, pp. 75-85, 2015.
- [26] M. Ladjal, &quot;Contribution au développement de systèmes de surveillance innovants dédiés au contrôle de la qualité des eaux potables&quot;; Thèse de doctorat, Université de M'sila, Algérie,2013.
- [27] M.W. Lee, et al., &quot;Real-time remote monitoring of small-scaled biological wastewater treatment plants by a multivariate statistical process control and neural network-based software sensors&quot;; Process. Biochem, Vol. 43, pp. 1107-1113, 2008.
- [28] E. Davalo, P. Naïm, &quot;Des réseaux de neurones&quot;; Edition Eyrolles, 2ème édition, Paris, France,1993.

- [29] X. Ding, S. Canu, T. Denoeux, &quot;Neural network models for forecasting&quot;; In J. G. Taylor, editor, Neural Networks and their Applications, John Wiley and Sons, pp. 243-252, Chichester, UK, 1996.
- [30] P. Besse, B. Laurent, &quot;Apprentissage Statistique : Pr evision et data mining&quot;; Institut de Math ematiques de Toulouse, France, 2012.
- [31] A. ASSOUM, «  tude de la tol erance aux al eas logiques des r eseaux de neurones Artificiels », th ese de doctorat, institut national polytechnique de Grenoble, France. 1997.]
- [32] A. Hadj Kaddour & A. Djedid  valuation des performances des techniques RNA et ELM utilis ees dans le d veloppement des capteurs logiciels pour la surveillance de la qualit  de l'eau. M'sila-2017.
- [33] S. Sojasi, "Caract risation de min raux indicateurs par imagerie hyperspectrale et traitement de l'image dans l'infrarouge proche et l'infrarouge lointain", Th ese de doctorat, Univ de Laval, Qu ebe , Canada, 2016.
- [34] Guang-Bin Huang. Introduction to Extreme learning machine. School of Electrical and Electronic Engineering Machines, University Singapore, cours Nov 2006.
- [35] HUANG, Guang-Bin, BAI, Zuo, KASUN, Liyanaarachchi Lekamalage Chamara, et al. Local receptive fields based extreme learning machine. IEEE Computational Intelligence Magazine, vol. 10, no 2, p. 18-29, 2015.
- [36] <https://www.kaggle.com/jboysen/mri-and-alzheimers>
- [37] J. Sola, J. Sevilla, "Importance of Input Data Normalization for the Application of Neural Networks to Complex Industrial problems", IEEE Transactions on Nuclear Science, Vol.44, No.03, pp.1464-1468, 1997.
- [38] M.DJRIOUI Contribution au D veloppement de Syst mes Multicapteurs Intelligents D di s   la Surveillance et au Contr le de la Qualit  des Eaux Propres. M'sila-2019.
- [39] Towardsdatascience <https://towardsdatascience.com/scale-standardize-or-normalize-with-scikit-learn-6ccc7d176a02>.
- [40] Qastack <https://qastack.fr/programming/40758562/can-anyone-explain-me-standardscaler>.
- [41] DING, Shifei, ZHAO, Han, ZHANG, Yanan, et al. Extreme learning machine: algorithm, theory and applications. Artificial Intelligence Review, vol. 44, no 1, p. 103-115, 2013.
- [42] HUANG, Guang-Bin, ZHOU, Hongming, DING, Xiaojian, et al. Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no 2, p. 513-529, 2012.