

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ IBN-KHALDOUN DE TIARET

FACULTÉ DES SCIENCES APPLIQUEES
DÉPARTEMENT DE GENIE ELECTRIQUE



MEMOIRE DE FIN D'ETUDES

Pour l'obtention du diplôme de Master

Domaine : Sciences et Technologie

Filière : Génie Electrique

Spécialité : Informatique industrielle

THÈME

**APPROCHE NEURONALE POUR LA CLASSIFICATION TEMPORELLE
DES SIGNAUX ECG**

**Préparé par : *MEBKHOUT Radouane*
*ZITOUNI Abdelkader***

Devant le Jury :

Nom et prénoms	Grade	Qualité
BENABID Houari	MAA	Président
BENSATTALAH Aissa	MAA	Examineur 1
MAASSKRI Mostapha	MAA	Examineur 2
GOISMI Mohamed	MAA	Encadreur

PROMOTION 2015 /2016

Dédicaces

A mon très cher père pour son soutien inconditionnel, son encouragement pour m'avoir permis de réaliser mes études dans les meilleures conditions et que dieu lui procure bonne santé et longue vie.

A ma défunte mère, à mon frère Mourad, à ma deuxième mère, à la famille Boutarene, à la famille Megnouche, à la famille Baza, à la famille Nedjar, à Yahia Rabah et à tous mes amis.

A tous ceux que j'ai connus durant toutes mes études.

A tous ceux pour qui je compte et qui comptent pour moi.

Je leur dédie ce travail.

Radouane

DÉDICACES

je dédie

À mes très chers parents à qui je dois tout le bonheur de ma vie,

À tous les membres de ma famille, petits et grands

Veillez trouver dans ce modeste travail l'expression de mon affection la plus sincère.

À mon cher père

À ma chère mère

À mes frères et sœur

À tous ceux qui m'ont soutenu et qui se reconnaîtront.

À tout mes collègues sans exception

Zitouni

Remerciement

Avant tous, Nous remercions en premier lieu ALLAH AZZA WA DJAL de nous avoir donné la santé et la patience pour avoir terminé ce travail.

Que le salut d'**ALLAH** soit sur notre prophète **MOHAMED SALLA ALLAH AALAYHI W SALLEM**, notre éducateur et guide de bon chemin.

Nous remercions très vivement Monsieur **GOISMI Mohamed**, de nous avoir confié le sujet, et de l'aide précieuse qu'il a toujours bien voulu nous apporter. Qu'il trouve ici l'expression de notre sincère reconnaissance.

Nous adressons nos chaleureux remerciements vont également aux membres du jury qui ont accepté d'évaluer ce travail et ce mémoire :

Monsieur **BENABID Houari**, Maître de conférences à l'Université Ibn Khaldoun de Tiaret, qui nous a fait l'honneur de présider le jury de mémoire.

Monsieur **BENSATTALAH Aissa**, Maître assistant à l'université de Tiaret qui nous fait l'honneur de participer à ce jury en qualité d'examineur de mémoire.

Monsieur **MAASSKRI Mostapha**, Maître assistant à l'université de Tiaret qui nous fait l'honneur de participer à ce jury en qualité d'examineur de mémoire.

Nos vifs remerciements vont également à l'ensemble des enseignants de l'université de Tiaret pour la formation qu'ils nous ont assuré au long de notre cursus universitaire.

Nos vifs remerciements vont également à tous ceux qui ont participé de près ou de loin à l'élaboration de ce travail.

SOMMAIRE:

Liste des figures	i
Liste des tableaux	iii
Introduction générale	1
Chapitre I: L'électrocardiographie	
I.1. Introduction	5
I.2. Le cœur	5
I.2.1. Anatomie du coeur	5
I.2.2. Fonctionnement électrique	6
I.3. L'électrocardiographie	7
I.3.1. Principe de l'électrocardiogramme (ECG)	7
I.3.2. Trace électrique du cœur	8
I.4. Les pathologies cardiaques	9
I.4.1. Éléments fondamentaux de la conduction électrique cardiaque	9
I.4.1.1. La conduction électrique	10
I.4.2. Le battement cardiaque standard et ses caractéristiques	10
I.5. Extrasystole ventriculaire (ESV)	12
I.6. Diagnostic à partir du rythme	13
I.6.1. Bradycardie	13
I.6.1.1. Bradycardie sinusale	13
I.6.1.2. Bradycardie d'origine fonctionnelle	13
I.6.2. La tachycardie	14
I.6.2.1. Tachycardie sinusale	14
I.6.2.2. Tachycardie auriculaire et nodale AV	14
I.6.2.3. Tachycardie ventriculaire (TV)	14
I.7. Arythmies, ou troubles de la régularité	15
I.8. Conclusion	15

Chapitre II : Les modèles de classification

II.1.	Introduction	17
II.2.	La classification en IA	17
II.2.1.	Chaine de classification d'objets	18
II.2.1.1.	Les principaux modules de classification d'objets	18
II.2.1.2.	L'extraction des descripteurs	18
II.2.1.3.	Procédure de résolution par apprentissage	19
II.2.1.3.1.	Phase d'apprentissage	19
II.2.1.3.2.	Phase de test	20
II.2.2.	Formalisation mathématique d'un problème de classification	21
II.3.	Méthodes de classification	21
II.3.1.	La logique floue en classification	21
II.3.1.1.	Base des règles floues	22
II.3.1.2.	Avantages et Inconvénients	22
II.3.1.3.	Les arbres de décision	23
II.3.1.4.	Exemple	23
II.3.2.	Avantages et Inconvénients	24
II.3.2.1.	Réseaux bayésiens	25
II.3.2.2.	Exemple	25
II.3.2.3.	Avantages et Inconvénients	25
II.3.3.	Machine a Vecteurs de Support (SVM)	26
II.3.3.1.	Avantages et Inconvénients	27
II.3.4.	Les carte de kohonen	28
II.3.4.1.	Avantages et Inconvénients	29
II.4.	Conclusion	29

Chapitre III : Les réseaux de neurones

III.1.	Introduction	31
--------	--------------	----

III.2.	Historique sur les réseaux de neurones	31
III.3.	Les réseaux de neurones	33
III.3.1.	Le neurone biologique	33
III.3.2.	Le neurone artificiel et principe de fonctionnement	33
III.3.3.	Réseau de neurones	35
III.3.4.	Architecture des réseaux de neurones	35
III.3.4.1.	Les réseaux de neurones non bouclés -feed forward	35
III.3.4.1.1.	Présentation	35
III.3.4.1.2.	Un peu d'histoire	36
III.3.4.1.3.	Composition du perceptron	36
III.3.4.2.	Les réseaux de neurones bouclés	37
III.3.5.	Apprentissage des réseaux de neurones	37
III.3.5.1.	Type d'apprentissage	37
III.3.5.1.1.	Les réseaux à apprentissage supervisé	37
III.3.5.1.2.	Les réseaux à apprentissage non supervisé	38
III.3.5.2.	Algorithmes d'apprentissage	38
III.3.5.2.1.	Algorithme de rétro-propagation	38
III.4.	Conception et mise en œuvre des réseaux de neurones	39
III.4.1.	Détermination des entrées/sorties du réseau de neurones	39
III.4.2.	Choix et préparation des échantillons	40
III.4.3.	Elaboration de la structure du réseau	40
III.4.4.	Apprentissage	41
III.4.5.	Validation et Tests	41
III.5.	Apprentissage par les réseaux multicouches	42
III.5.1.	Algorithme d'apprentissage	42
III.5.1.1.	Algorithme de la rétro-propagation du gradient	43
III.6.	Conclusion.	44

Chapitre IV : Contribution Et Réalisation	44
IV.1. Introduction	46
IV.2. Classification neuronale des arythmies cardiaques	46
IV.2.1. Sélection de la base d'exemples	47
IV.2.1.1. Sélection des descripteurs temporaires d'un cycle cardiaque	48
IV.2.2. Architecture du réseau de neurones	50
IV.2.2.1. Algorithme d'apprentissage	50
IV.2.2.2. Dimensionnement du réseau pour chaque classifieur	50
IV.2.3. Résultats expérimentaux des classifications	51
IV.3. Architecture des descripteurs temporels	52
IV.4. Conclusion	55
Conclusion générale	57
Références bibliographiques	59
Annexe	61
Annexe I : Base de données MIT-BIH	

Table des figures

Chapitre I: L'électrocardiographie

Figure I. 1: Schéma fonctionnel du cœur.	5
Figure I. 2: L'impulsion électrique se propage dans le muscle cardiaque et induit sa contraction.	6
Figure I. 3: Déflexion du vecteur électrique	7
Figure I. 4: La contraction des oreillettes se traduit sur l'ECG	8
Figure I. 5: Eléments fondamentaux de la conduction électrique cardiaque	9
Figure I. 6: Paramètres d'intérêt pour la description d'un battement.	11
Figure I. 7 : Extrasystole ventriculaire (ESV).	13

Chapitre II : Les modèles de classification

Figure II. 1: Arbre de décision	24
Figure II. 2: Structure de causalité	25
Figure II. 3: Schéma d'une carte de Kohonen connectée à un stimulus. Les connections sont complètes	28

Chapitre III : Les réseaux de neurones

Figure III. 1: Le neurone biologique	33
Figure III. 2: Le neurone artificiel.	34
Figure III. 3: Fonctions d'activations.	34
Figure III. 4: Exemple de réseau non-bouclé	36
Figure III. 5: Exemple de réseau bouclé	37
Figure III. 6: Organigramme de conception d'un réseau de neurones	42

Chapitre IV: Contribution Et Réalisation

Figure IV. 1: Schéma global de la classification	47
Figure IV. 2: Battements normaux et intervalles d'intérêt	49
Figure IV. 3: Architecture des descripteurs temporels.	52
Figure IV. 4: Interface principale du classifieur temporel	52
Figure IV. 5: Paramétrage du réseau de neurone	53
Figure IV. 6: Interface d'apprentissage du réseau de neurone	53
Figure IV. 7: Graphe de performance	54
Figure IV. 8: Les graphes des gradient et mu et validation de teste.	54
Figure IV. 9: Validation de test.	55

Annexes

Chapitre II:

Tableau II. 1: Les définitions des grandeurs VP, VN, FP et FN.	21
Tableau II. 2: Base des règles floues	22
Tableau II. 3: Exemple des données	23

Chapitre IV:

Tableau IV. 1: Les enregistrements choisis de la base de données MIT-BIH.	48
Tableau IV. 2: Performances du classifieur neuronal à données temporelles.	51

INTRODUCTION

GENERALE

Introduction générale

Les maladies cardio-vasculaires sont responsables de 17 millions de décès à travers le monde. Elles sont suivies par le cancer qui tue 7,6 millions de personnes chaque année à travers le monde.

Le premier rapport de l'Organisation mondiale de la santé (OMS) sur les maladies non transmissibles (MNT) rendu public le 27 avril 2011, révèle que les pays en voie de développement sont les plus touchés par les MNT.

« En 2008, 36,1 millions de personnes sont décédées d'affections telles que les cardiopathies, les accidents vasculaires cérébraux (AVC), les pneumo-pathies chroniques, les cancers et le diabète. Plus de 80% de ces décès sont survenus dans les pays à revenu faible ou moyen», indique le communiqué de l'OMS. Les maladies cardio-vasculaires et le cancer sont à l'origine de plusieurs millions de décès chaque année à travers le monde.

Les maladies cardiovasculaires sont responsables de la plupart des décès par MNT, à savoir 17 millions de personnes chaque année ; elles sont suivies par le cancer (7,6 millions), les maladies respiratoires (4,2 millions) et le diabète (1,3 million). Ces quatre groupes de maladies représentent près de 80% de l'ensemble des décès par MNT et ont en commun quatre facteurs à risque : le tabagisme, la sédentarité, l'usage nocif de l'alcool et une alimentation déséquilibrée, renseigne le rapport de l'OMS.

La prise en charge de ces maladies, reconnaît le directeur général de l'OMS, constitue un énorme défi pour beaucoup de pays, où la prévalence est alarmante. «Pour certains pays, il n'est pas exagéré de décrire la situation comme une catastrophe imminente, pour la santé, pour la société et surtout pour les économies nationales», souligne le directeur général de l'OMS, le docteur Chan.

La proportion des personnes en âge d'activité affectées est importante. Selon le sous-directeur général de l'OMS, le docteur Ala Alwan, près de 30% des personnes qui meurent de maladies non transmissibles dans les pays à faibles revenus sont âgées de moins de 60 ans. Ces décès, dit-il, sont d'autant plus tragiques qu'ils sont en grande partie évitables.

Pour diagnostiquer des troubles du rythme cardiaque, les médecins ont pour habitude d'utiliser principalement l'électrocardiogramme (ECG). Lorsque l'ECG est bruité, le cerveau humain est capable, dans une certaine mesure, de retrouver les informations exactes parmi les informations bruitées ou de chercher dans d'autres sources de données (par exemple sur la morphologie de l'ECG) la confirmation de ces informations.

L'automatisation de cette recherche d'informations n'est pas triviale, parce qu'il est difficile d'extraire du signal ECG toutes les informations dont se sert le médecin pour diagnostiquer des

arythmies, mais aussi parce que le système de monitoring doit être capable, si l'ECG est trop bruité pour pouvoir être utilisé tel quel, de trouver l'information nécessaire dans les autres sources de données, même si elles ne sont à priori pas suffisantes pour permettre un diagnostic fiable. En outre, fournir à un système suffisamment de connaissances expertes pour automatiser le diagnostic est une tâche extrêmement coûteuse et fastidieuse.

L'analyse du rythme figure parmi les applications les plus prometteuses de l'ingénierie biomédicale : elle consiste, pour l'essentiel, à détecter les anomalies éventuelles, à identifier leur nature et à préciser un traitement approprié, d'où l'intérêt de développer un outil automatisé d'aide au diagnostic du signal ECG. Cet objectif comporte plusieurs volets, qui sont d'ailleurs communs à tous les outils d'aide à la décision.

1. L'acquisition du signal ECG. L'instrumentation doit permettre un enregistrement fidèle de ce signal, tout en minimisant l'incidence des signaux parasites (bruits venants du secteur, de l'interface peau-électrode et des parties du corps autres que le cœur).
2. Le prétraitement du signal ECG. L'objectif de cette étape est la transformation du signal brut en un signal interprétable par filtrage de l'ECG et la détection de ses différentes ondes.
3. La caractérisation du signal ECG. Cette étape consiste en le choix de caractéristique et paramètres permettant une description informative du signal ECG.
4. L'analyse des données et l'interprétation des résultats. Les informations obtenues lors des étapes précédentes doivent être intégrées dans des algorithmes permettant de faire la distinction entre les différents types de signaux. La classification obtenue doit avoir une signification (séparer les signaux normaux des ECGs pathologiques par exemple) et doit pouvoir être interprétée par le médecin cardiologue.

Le but de ce mémoire est d'une part d'appliquer des techniques permettant d'interpréter de manière automatique un signal ECG en basant sur une voie essentielle :

Il est nécessaire d'extraire des caractéristiques temporelles telles que la fréquence et la période du temps de certaines ondes (P, Q, R, S, T, le complexe QRS, la période RR, etc..) de l'ECG, cette voie est dite : voie interprétation par caractéristiques temporelles.

Le manuscrit de ce mémoire est divisé en quatre chapitres :

Le premier chapitre consiste sur l'anatomie et la biologie du cœur, les dérivations cardiaques pour enregistrer cette activité électrique sont ensuite présentées. Les ondes élémentaires d'un signal ECG sont décrites ainsi que les intervalles, durées, et segments que l'on mesure. Et en dernière section, le chapitre décrit brièvement les différentes arythmies et pathologies cardiaques susceptibles d'être détectées dans un enregistrement ECG, le but, ici, est de mettre en relation

certaines observations, anormales du tracé ECG avec les anomalies cardiaques les plus courantes. Il s'agit de sensibiliser le lecteur à l'importance diagnostique de l'étude des caractéristiques des ondes P, Q, R, S et T (Formes, distances, ...), au-delà d'une simple de rythme.

Le deuxième chapitre définit ce qu'est un problème de classification. Parmi la variété de problèmes de classification, nous analysons les avantages et les inconvénients de quelques méthodes de classification telles que (la logique floue, les arbres de décision, les réseaux bayésiens, les machines à vecteur de support (SVM) et les cartes de kohonen).

Le troisième chapitre permet de détailler le modèle de classification utilisé, à savoir les réseaux de neurones. En donnant l'historique de ce modèle, ces architectures, ces composants essentiels, les algorithmes d'apprentissage.

Le quatrième chapitre décrit notre application, les modèles neuronaux pour la classification des battements cardiaques, le modèle neuronal à base caractéristiques temporelles extraites depuis le signal ECG.

Finalement, on terminera ce mémoire par une conclusion générale qui résume les résultats obtenus

CHAPITRE I

L'ELECTROCARDIOG RAPHIE

I.1. Introduction

Le cœur est l'élément central du système cardiovasculaire qui permet l'alimentation en oxygène et en nutriments des organes. Il est principalement composé de quatre cavités, deux oreillettes et deux ventricules. À chaque battement, leurs contractions successives peuvent être suivies depuis l'extérieur du corps par des électrodes, collées à la surface de la peau, qui mesurent l'activité électrique des fibres musculaires cardiaques.

Cet examen non invasif s'appelle l'électrocardiogramme (ECG) est effectué pendant quelques minutes à l'hôpital, quand à l'examen Holter est effectué sur 24 heures.

L'analyse de ces enregistrements permet, à elle seule, de diagnostiquer un grand nombre de pathologies et de les localiser. En effet, chaque couple de cavités (les oreillettes et les ventricules) possède sa propre signature électrique :

- onde **P** : systole auriculaire,
- ondes **Q, R** et **S** : systole ventriculaire,
- onde **T** : repolarisation ventriculaire.

I.2. Le cœur

Le cœur est l'élément central du système cardiovasculaire. Nous décrivons dans la suite du chapitre l'anatomie et le fonctionnement électrique d'un cœur sain.

I.2.1. Anatomie du cœur

Le cœur propulse le sang grâce aux contractions de son tissu musculaire appelé myocarde. Une épaisse cloison le divise en deux moitiés (cœur gauche/cœur droit), et chacune d'elles comporte deux cavités : l'oreillette et le ventricule. À chaque battement, le myocarde suit la même séquence de mouvement, le sang pauvre en oxygène arrive au cœur par la veine cave.

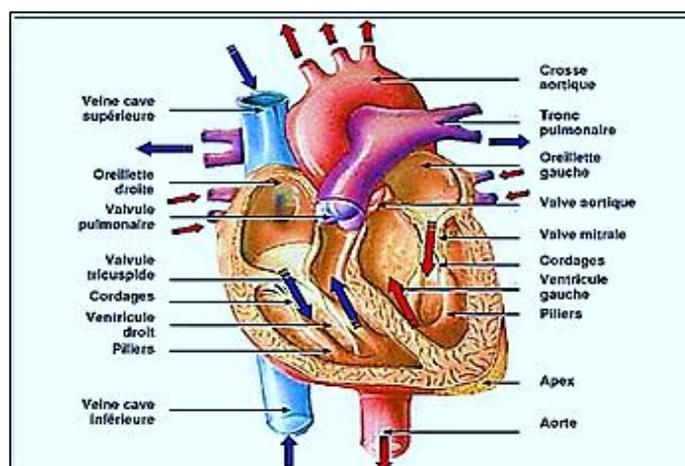


Figure I. 1: Schéma fonctionnel du cœur.

Il y entre par l'oreillette droite, et en est chassé par sa contraction appelée systole auriculaire qui le déplace dans le ventricule droit. La systole ventriculaire (contraction des ventricules) propulse à son tour le sang du ventricule droit vers les poumons où il va se charger en oxygène. De retour au cœur par les veines pulmonaires, le sang s'accumule dans l'oreillette gauche puis, lors de la systole auriculaire, passe dans le ventricule gauche qui lors de la systole ventriculaire l'envoie vers les organes par l'artère aorte (figure I.1).

I.2.2. Fonctionnement électrique

Comme pour tous les muscles du corps, la contraction du myocarde est provoquée par la propagation d'une impulsion électrique le long des fibres musculaires cardiaques induite par la dépolarisation des cellules musculaires. Dans le cœur, la dépolarisation prend normalement naissance dans le haut de l'oreillette droite (le sinus), et se propage ensuite dans les oreillettes, induisant la systole auriculaire (figure I.2) qui est suivie d'une diastole (décontraction du muscle).

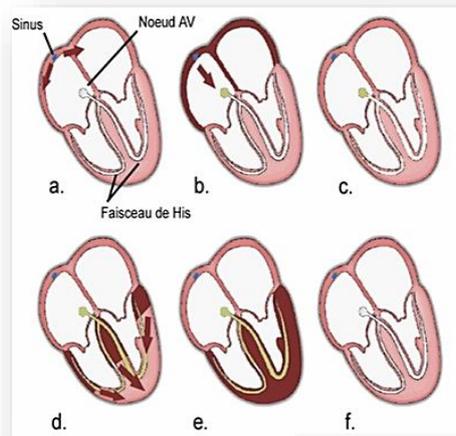


Figure I. 2: L'impulsion électrique se propage dans le muscle cardiaque et induit sa contraction.

L'impulsion électrique arrive alors au nœud auriculo-ventriculaire (AV), seul point de passage possible pour le courant électrique entre les oreillettes et les ventricules. L'impulsion électrique subit une courte pause permettant au sang de pénétrer dans les ventricules. Elle emprunte alors le faisceau de His, qui est composé de deux branches principales allant chacune dans un ventricule. Les fibres constituant ce faisceau, complétées par les fibres de Purkinje, grâce à leur conduction rapide, propagent l'impulsion électrique en plusieurs points des ventricules, et permettent ainsi une dépolarisation quasi instantanée de l'ensemble du muscle ventriculaire, malgré sa taille importante, ce qui assure une efficacité optimale dans la propulsion du sang, cette contraction constitue la phase de systole ventriculaire. Puis suit la diastole ventriculaire (décontraction du muscle), les fibres musculaires se repolarisent et reviennent ainsi dans leur état initial.

I.3. L'électrocardiographie

Ou l'art d'enregistrer l'activité électrique du cœur.

I.3.1. Principe de l'électrocardiogramme (ECG)

Vers 1880, E Marey et Augustus Waller montrèrent que l'activité électrique du cœur, découverte quelques années plus tôt, pouvait être suivie à partir de la peau, et vers 1890, Willem Einthoven réalisa le premier enregistrement cardiographique[1] .

Le courant mesuré par des électrodes sur le torse du patient mettait en mouvement un mince fil d'argent tendu entre les pôles d'un gros aimant, ces déflexions étaient enregistrées sur du papier photographique qui se déroulait face à un rayon lumineux.

« Maintenant, nous pouvons enregistrer l'activité électrique du cœur anormale et la comparer à l'activité normale». Ainsi naît l'électrocardiogramme à l'aube du XXIème siècle.

Le principe de l'enregistrement moderne est, à peu de chose près, celui qui fut proposé par Einthoven, grâce à deux électrodes collées à la surface de la peau, on enregistre la différence de potentiel entre deux points diamétralement opposés par rapport au cœur, ce signal étant directement corrélé au déplacement de l'impulsion électrique dans les fibres du muscle cardiaque.

L'activité électrique instantanée peut être définie par un vecteur orienté suivant la différence de potentiel présente dans le cœur, et de module proportionnel à celle-ci. Le couple d'électrodes enregistre à chaque instant l'amplitude de la projection de ce vecteur suivant leur axe ainsi, lorsque le vecteur électrique est orienté de l'électrode - à l'électrode +, on observe sur l'enregistreur une déflexion positive, et lorsque le vecteur est orienté en sens inverse, la déflexion est négative. À l'échelle d'une cellule, le détail des tracés est indiqué sur la figure I.3.

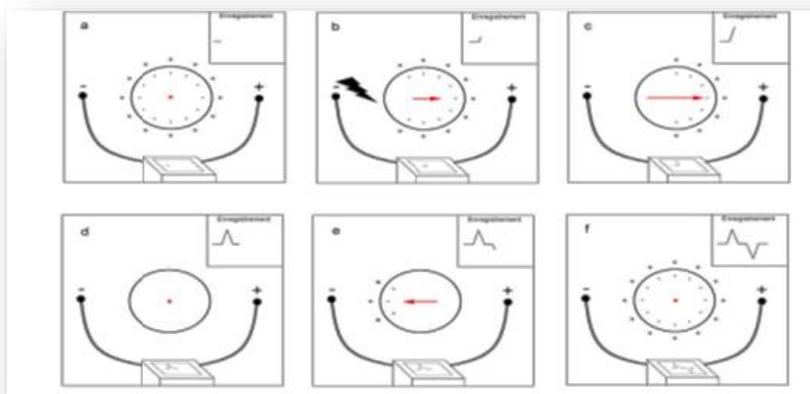


Figure I. 3: Déflexion de vecteur électrique

I.3.2. Trace électrique du cœur

Le battement cardiaque peut donc être suivi grâce à l'enregistrement en surface du signal électrique qui l'accompagne. En effet, chaque phase du battement possède une trace électrique particulière. Un œil exercé peut donc, dans la plupart des cas, différencier de manière rapide la trace d'une contraction auriculaire de la trace d'une contraction ventriculaire.

L'impulsion initiale vient du sinus, elle n'est pas visible sur l'ECG. L'onde électrique qui se propage ensuite dans les oreillettes, entraînant leurs contractions, laisse la trace d'une petite déflexion positive sur l'ECG l'onde P.

L'impulsion arrive alors au nœud auriculo-ventriculaire (AV), où se produit la courte pause qui se traduit sur l'ECG par un petit segment plat, puis elle emprunte les voies de conductions rapides (le faisceau de His) pour entraîner la contraction des ventricules, suivie de leur repolarisation. Cette propagation de l'impulsion, la contraction brève et puissante de l'ensemble du muscle ventriculaire, dessinent sur l'ECG une succession de 3 ondes (Q, R et S) appelé complexe QRS. L'onde Q est la première : c'est une onde dirigée vers le bas, qui n'est pas toujours visible sur le tracé, la seconde est l'onde R : elle est de grande amplitude et dirigée vers le haut, la dernière est dirigée vers le bas, c'est l'onde S.

C'est l'ensemble de ces trois ondes qui constitue le complexe QRS. Après chaque complexe QRS, on observe sur l'ECG une onde appelée onde T. Entre cette onde et l'onde précédente, on note une courte pause appelée le segment ST, dont l'étude est très importante pour l'identification de certaines pathologies. L'onde T traduit la phase de repolarisation des cellules constituant les ventricules, c'est un phénomène purement électrique et pendant cette phase le cœur est mécaniquement inactif (figure I.4)[2].

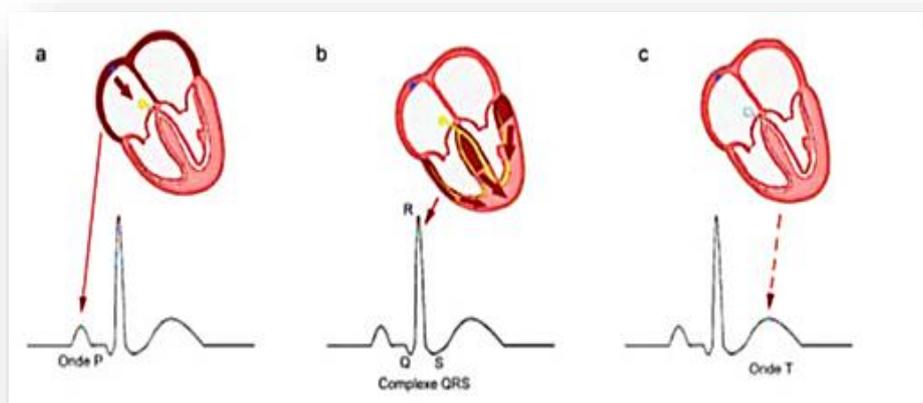


Figure I. 4: La contraction des oreillettes se traduit sur l'ECG

I.4. Les pathologies cardiaques

I.4.1. Eléments fondamentaux de la conduction électrique cardiaque

Cette section est volontairement limitée aux notions utiles à la compréhension des chapitres suivants. Le lecteur intéressé pourra cependant trouver des informations complémentaires détaillées et de manière plus intuitive dans [3][4].

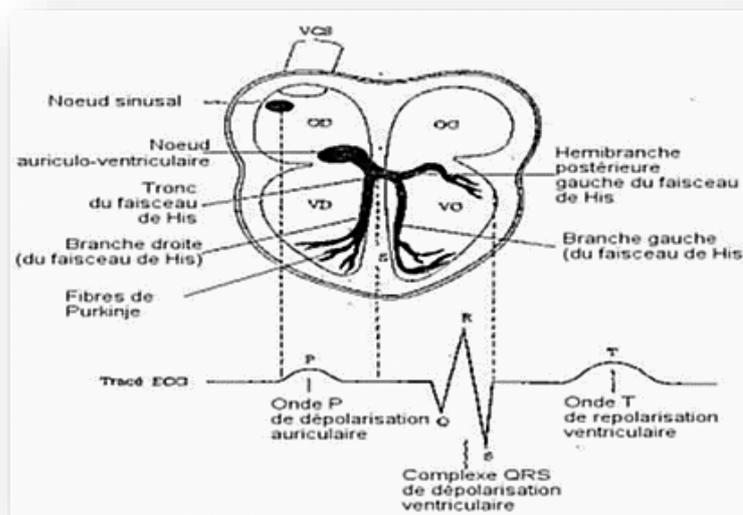


Figure I. 5: Eléments fondamentaux de la conduction électrique cardiaque

La contraction du muscle cardiaque (ou myocarde) a pour origine la propagation d'une onde électrique qui excite les cellules musculaires dans un ordre bien établi afin que la contraction soit la plus efficace possible. Le système de conduction électrique comprend : le nœud sinusal, les voies spécialisées inter nodales, le nœud auriculo-ventriculaire, le faisceau de His, les branches gauche et droite et les fibres de Purkinje qui terminent les deux branches. Dans le cas normal, le stimulus physiologique à l'origine de l'onde électrique (et donc du battement cardiaque) provient du nœud sinusal. Ce stimulus est fourni rythmiquement à une fréquence comprise entre 60 et 100/minute. L'impulsion cardiaque initiée dans le nœud sinusal est ensuite transmise aux deux oreillettes au moyen des voies inter-nodales qui relient le nœud sinusal au nœud auriculo-ventriculaire. L'onde se propage ensuite à travers le faisceau de His puis vers les branches gauche et droite jusqu'aux fibres de Purkinje. La propagation de l'onde se fait par une succession de dépolarisation des tissus du myocarde.

Après le passage de l'onde, la membrane des tissus se repolarise. Cette succession de dépolarisations et de repolarisations des tissus cardiaques est visible sous plusieurs angles sur un électrocardiogramme (ECG).

I.4.1.1. La conduction électrique

L'électrocardiogramme L'ECG reflète l'activité électrique cardiaque. L'électrocardiographie moderne s'appuie sur l'étude de douze dérivations (ou dérivations standards). Ces dernières sont obtenues en plaçant des électrodes sur le thorax près du cœur, les précordiales (6 voies de V1 à V6), ou sur les bras et les jambes (dérivations bipolaires D1 à D3 et unipolaires aVR, aVL et aVF) .

Ces différentes dérivations permettent d'avoir plusieurs vues de la propagation électrique. Lorsqu'il n'y a aucun dysfonctionnement, ces vues sont redondantes, on se sert alors principalement des voies bipolaires I et II. L'ensemble des vues est cependant nécessaire en cas de dysfonctionnement pour localiser précisément la source du problème. Lors de la propagation de l'onde électrique dans le cœur, les électrodes permettent d'enregistrer un motif électrique composé d'une succession d'ondes caractéristiques désignées par les lettres de l'alphabet à partir de P[5].

L'onde P initiale représente la dépolarisation des oreillettes. L'ensemble QRS (appelé aussi complexe QRS) traduit la dépolarisation des ventricules tandis que l'onde T correspond à leur repolarisation (la repolarisation des oreillettes est masquée par celle des ventricules). L'onde U est inconstante et de faible amplitude, elle a une signification encore discutée

I.4.2. Le battement cardiaque standard et ses caractéristiques

L'étude d'un enregistrement ECG est fondée sur l'analyse de quelques battements cardiaques successifs, l'étude d'un seul battement ne fournit que peu d'indications pour la pose d'un diagnostic, mais les variations des paramètres caractéristiques de chaque battement au cours de l'enregistrement constituent une source d'information essentielle. Ces paramètres caractéristiques sont :

Les durées des ondes P, Q, R, S et T, et les amplitudes mesurées par rapport à la ligne de base (figure I.6). La ligne de base, est la ligne isoélectrique du cœur au repos, qui est prise comme référence pour mesurer l'amplitude des ondes pendant l'inactivité cardiaque, le potentiel mesuré est donc normalement nul par rapport à cette référence. C'est le cas au niveau de l'intervalle entre les ondes T et P de deux battements successifs, de l'intervalle entre les ondes P et Q d'un même battement, de l'intervalle entre les ondes S et T (en l'absence de pathologie) Les distances entre ces ondes.

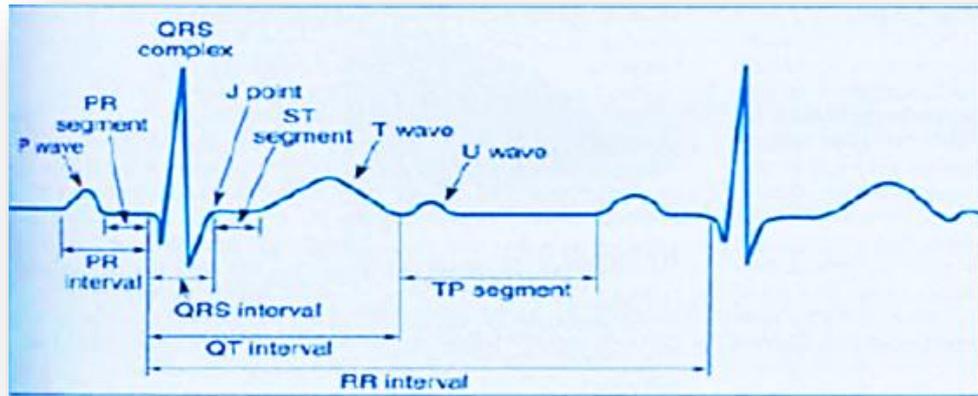


Figure I. 6: Paramètres d'intérêt pour la description d'un battement.

Pour caractériser le rythme cardiaque, les cardiologues se fient à la forme générale de l'ECG. Ils portent une attention particulière à la forme des ondes citées précédemment et aux intervalles de temps qui les séparent. La FIG 1.6 répertorie les intervalles qui seront utilisés dans la suite de cette étude. Les valeurs numériques utilisées proviennent de la littérature médicale et notamment de [6]

- ❖ l'intervalle RR (ou RR1) sépare les sommets de deux ondes R consécutives et définit le rythme ventriculaire. Un intervalle RR normal est généralement mesuré entre 600 et 1000ms
- ❖ l'intervalle PP (ou PP1) sépare les sommets de deux ondes P consécutives et définit le rythme auriculaire.
- ❖ l'intervalle PR (ou PR1) s'étend du début de l'onde P au début du complexe QRS. Un intervalle PR normal mesure entre 120 et 200 ms.
- ❖ l'intervalle RP (ou RP1) s'étend du début du complexe QRS jusqu'à l'onde P suivante
- ❖ les intervalles RR2, PP2 séparent respectivement les sommets de deux QRS (resp. onde P) séparés par un QRS (resp. une onde P).

Les instants des ondes Q, S et T sont encore très difficiles à détecter automatiquement et ne seront donc pas utilisés dans cette étude. La régularité du rythme est mesurée en termes de variation de durée d'un intervalle RR à l'intervalle suivant.

Les intervalles de temps définis entre deux ondes caractéristiques de l'ECG fournissent d'importants indicateurs pour le diagnostic de maladies cardiaques car ils sont le reflet du processus physiologique. Par exemple, l'intervalle **R-R**, représentatif de la période cardiaque, est généralement obtenu en détectant le pic de l'onde **R**. De par sa grande amplitude, la détection de cette onde est la plus accessible et nous permet souvent de segmenter l'ECG avant d'estimer toute autre onde.

L'intervalle **Q-T**, reflétant la durée de la dépolarisation et de la repolarisation ventriculaire, est l'un des intervalles les plus importants de l'ECG. En effet, sa prolongation peut être associée à des risques d'arythmie ventriculaire et de mort subite. Il est classiquement calculé entre le début du complexe **QRS** et la fin de l'onde **T**. Cependant, la fin de l'onde **T** est difficile à détecter à cause de la morphologie très variée de l'onde **T**, et surtout de la perturbation du bruit notamment sur la fin de l'onde.

Tout comme pour l'onde **T**, la détection de l'onde **P** n'est pas évidente, voire davantage complexe. Effectivement, l'onde **P** est souvent de faible amplitude et noyée dans le bruit.

L'estimation des intervalles **P-R** est donc très difficile. Alors que beaucoup d'études portent sur la détection de l'intervalle **Q-T**.

I.5. Extrasystole ventriculaire (ESV)

L'extrasystole ventriculaire (ESV) est un battement anormal suffisamment répandu pour être présenté en introduction. Les ESV s'observent sur quasiment tous les enregistrements, principalement en période de récupération après un effort. Bien que leur présence n'indique aucune pathologie particulière, si, de façon récurrente, leur nombre par minute est supérieur à 6, elles peuvent être un signe précurseur d'une tachycardie ventriculaire qui, elle, constitue une pathologie majeure.

Contrairement aux battements normaux qui ont pour origine la dépolarisation des cellules sinusales, l'ESV naît de la dépolarisation spontanée d'un petit groupe de cellules ventriculaires, appelé alors foyer ectopique ventriculaire. L'impulsion électrique créée n'emprunte pas la voie normale de conduction (faisceau de His), et se propage donc plus lentement dans les ventricules. La contraction ventriculaire ainsi étalée dans le temps perd de son efficacité.

Le tracé d'un battement ESV est caractérisé par deux propriétés : l'onde **R** n'est pas précédée d'une onde **P**, puisqu'il n'y a pas eu d'activité auriculaire préalable, et la durée du complexe est supérieure à la durée d'un complexe **QRS** normal (figure I.7).

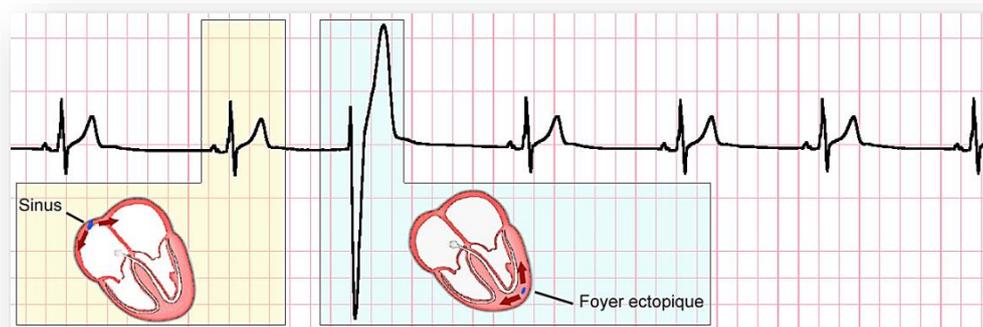


Figure I. 7 : Extrasystole ventriculaire (ESV).

I.6. Diagnostic à partir du rythme

Pour dresser un diagnostic à partir d'un enregistrement Holter, deux points essentiels sont à étudier : le rythme cardiaque (fréquence et régularité) et la forme des ondes (distances relatives, amplitudes, durées,...).

I.6.1. Bradycardie

La bradycardie est caractérisée par le fait que la fréquence cardiaque est inférieure à 60 bpm, elle est dite d'origine sinusale, fonctionnelle, ou ventriculaire, selon le site d'initiation de l'impulsion électrique à l'origine des battements considérés.

I.6.1.1. Bradycardie sinusale

Outre sa fréquence basse, la bradycardie sinusale est caractérisée par la présence systématique d'une onde P avant les complexes QRS (Figure 4) car l'origine de la contraction ventriculaire reste la dépolarisation du sinus et des oreillettes, comme lors de battements normaux. Les causes d'une telle arythmie sont multiples et souvent extérieures au système cardiovasculaire : traitement médicamenteux (bêtabloquant, dioxine,...), hypothermie, urémie,...

I.6.1.2. Bradycardie d'origine fonctionnelle

En cas de dysfonctionnement sinusal avec ou sans dépolarisation auriculaire, le nœud AV peut assurer le rôle de pacemaker de secours à une fréquence de l'ordre de 30 à 60 bpm.

L'impulsion électrique prend alors naissance dans le nœud auriculo-ventriculaire (AV) et suit le chemin de conduction habituel : la morphologie des ondes QRS est identique à celle du battement normal ; en revanche, l'onde P peut être absente ou, si elle est présente, elle peut être désynchronisée de la systole ventriculaire.

Lorsque le problème de conduction intervient en aval du nœud AV, c'est un groupe de cellules du muscle ventriculaire qui est susceptible de devenir pacemaker (foyer ectopique ventriculaire) ; le rythme devient alors une succession d'ESV à une fréquence très lente, entre 15 et 40 bpm, appelé rythme d'échappement ventriculaire. Ce type de pathologie constitue une indication typique de la pose d'un stimulateur cardiaque pour éviter une possible syncope.

I.6.2. La tachycardie

À l'inverse de la bradycardie, la tachycardie est caractérisée par le fait que la fréquence est supérieure à 100 bpm, elle peut être d'origine sinusale, auriculaire, ou ventriculaire.

I.6.2.1. Tachycardie sinusale

La tachycardie sinusale correspond à un rythme sinusal dont la fréquence est comprise entre 100 et 180 bpm, on observe, avant chaque complexe QRS, une onde P de forme identique à celle observée lors du rythme normal.

Une tachycardie sinusale est considérée comme pathologique si elle est de longue durée et indépendante du contexte.

Les causes de cette pathologie sont habituellement extracardiaques, et incluent tous les facteurs de stimulations du système nerveux : surmenage, anxiété, ... et l'influence de différentes substances comme l'adrénaline ou la caféine.

I.6.2.2. Tachycardie auriculaire et nodale AV

La tachycardie auriculaire peut avoir pour origine un foyer ectopique, une boucle de stimulation (flutter) ou une voie qui court-circuite la voie AV, dite voie accessoire, avec réentrée par le nœud AV.

I.6.2.3. Tachycardie ventriculaire (TV)

La tachycardie ventriculaire a pour origine un ou plusieurs foyer(s) ectopique(s) ventriculaire(s) (qui se dépolarisent à tour de rôle). Les battements ont donc la forme d'extrasystoles ventriculaires très rapprochées. Ce type de rythme est dangereux à cause de sa possible évolution en fibrillation ventriculaire qui, elle, conduit au décès du patient si elle n'est pas traitée à l'aide d'un défibrillateur dans les quelques minutes qui suivent son apparition.

I.7. Arythmies, ou troubles de la régularité

L'absence de régularité des battements cardiaques est une caractéristique du rythme importante pour le diagnostic, elle est souvent associée à un trouble de la production ou de la conduction de l'impulsion électrique (foyers ectopiques, blocs, boucles,...). Les arythmies permanentes ou sporadiques nécessitent un suivi médical et sont des indications typiques à la pose régulière d'un enregistreur Holter

I.8. Conclusion

L'analyse automatique du tracé Holter comme elle est effectuée depuis déjà quelques années permet de mettre en évidence un grand nombre de pathologies cardiaques, notamment celles où la fréquence cardiaque et/ou la régularité des battements sont anormales, (bradycardies, tachycardies, bigéminismes,...). De plus une étude simple de la forme des complexes QRS permet de distinguer les complexes d'origine supraventriculaire des extrasystoles ventriculaires.

Cependant la caractérisation précise des pathologies sous-jacentes aux anomalies détectées n'est pas encore vraiment accessible, car elle nécessite, en général, l'analyse de la présence des ondes autres que le QRS. Citons, par exemple, la distinction entre une bradycardie sinusale et une bradycardie jonctionnelle qui impose l'analyse de la présence ou non de l'onde P. En outre, toutes les pathologies qui se traduisent uniquement par des distances anormales entre les ondes et/ou des absences d'ondes ne peuvent aujourd'hui être décelées automatiquement.

Ainsi, la méthode proposée dans ce qui suit, permettant un repérage précis et continu de toutes les ondes caractéristiques P, Q, R, S et T du battement, devra permettre de localiser plus précisément toutes les zones du signal susceptibles de porter la trace d'un comportement anormal du cœur sur les 24 heures de l'enregistrement Holter.

Chapitre II

LES MODELES DE CLASSIFICATION

II.1. Introduction

Un diagnostic médical est le résultat du raisonnement d'un médecin, décision très souvent prise à partir d'informations incertaines et/ou incomplètes. De nombreuses techniques d'intelligence artificielle (IA) ont été appliquées pour essayer de modéliser ce raisonnement[7]. Ainsi présente l'utilisation détaillée de plusieurs systèmes experts en médecine. Citons par exemple, des systèmes à base de règles comme MYCIN[8].

En amont de ce raisonnement, il faut aussi être capable de modéliser ces informations incertaines et/ou incomplètes. Les performances des ordinateurs sont de plus en plus remarquables, et avec elles, la nécessité et la possibilité de développer des méthodes de classification automatiques capables et robustes augmentent également. Dans ce contexte, de nombreuses méthodes de classification automatiques sont apparues. Les plus connues sont : les méthodes de la logique floue, les réseaux de neurones, les réseaux bayésiens, les arbres de décision, les SVM (Support Vector Machine), les cartes de kohonen, etc...

II.2. La classification en IA

La classification rassemble une famille de méthodes permettant d'automatiser le processus de reconnaissance, elle peut être considérée comme un domaine permettant de définir des algorithmes et des techniques susceptibles de classer des objets dont l'aspect est variable par rapport à un objet type. Le but de la classification est donc de classer des observations sur la base d'une série de caractéristiques prédéfinies par apprentissage. Aujourd'hui on ne peut pas parler d'une théorie de la classification d'objets, mais d'approches différentes dont le but commun est de mettre en évidence, de façon automatique, une structure commune à l'ensemble des objets en présence, par recherche de similarités. Le champ d'application de la classification est particulièrement vaste. Il couvre principalement le domaine du traitement des signaux et d'images. On distingue habituellement deux types de classification :

- ❖ La classification avec ensemble d'apprentissage connu et parfaitement expertisé, ou classification en mode supervisé.
- ❖ La classification en mode non supervisé, dans ce dernier cas, on ne dispose pas d'un ensemble d'apprentissage préalablement expertisé.

II.2.1. Chaîne de classification d'objets

II.2.1.1. Les principaux modules de classification d'objets

Un dispositif de classification ou de reconnaissance automatique de formes est généralement conçu comme une chaîne de modules de traitement. Ainsi, un système de reconnaissance de formes comporte habituellement les modules suivants :

- 1- **Un module d'acquisition:** des capteurs mesurent des grandeurs caractéristiques de l'objet à classer (signal, image). Cet ensemble de grandeurs constitue la première représentation de l'objet.
- 2- **Un module de prétraitement :** il peut être judicieux de modifier des grandeurs brutes issues des capteurs par un algorithme afin de tenir compte des connaissances qui peuvent être disponibles a priori sur le problème. Par exemple à partir de la réponse d'un capteur on peut appliquer un ensemble de filtres destinés à éliminer des bruits indésirables. Ainsi, on obtient une nouvelle représentation de l'objet, plus adéquate pour la classification envisagée. D'autres modules de traitement peuvent élaborer des représentations successives de l'objet, ces différentes représentations ont généralement pour objectif de réduire la dimension de la représentation, c'est-à-dire de diminuer le nombre de descripteurs de l'objet et d'élaborer des descripteurs de plus en plus pertinents pour la tâche de discrimination à accomplir.
- 3- **Un module de classification :** l'algorithme de classification considère la dernière représentation de l'objet et décide d'affecter celui-ci à une classe. Cet algorithme peut fournir soit une réponse binaire à valeurs discrètes (appartenance ou non à une classe) soit une réponse à valeurs continues[9].

II.2.1.2. L'extraction des descripteurs

Habituellement une classification de formes ne se fait pas directement sur des formes brutes (morphologique de l'ECG par exemple), mais plutôt à partir de descripteurs ou paramètres caractérisant les formes. Chacune des N formes F_i est représentée par un point X_i qui est l'espace des paramètres. Il existe deux approches pour caractériser une forme :

- a. **L'approche purement mathématique** (analyse en composantes principales, prédiction linéaire, transformée de Fourier, etc.), qui consiste à retenir comme paramètres certains coefficients pertinents.
- b. **L'approche intuitive** qui laisse au spécialiste le soin de définir les descripteurs qui lui semblent importants. Cette approche donne souvent de meilleurs résultats, car les paramètres choisis résultent d'une grande expérience et peuvent être plus discriminants [10].

II.2.1.3. Procédure de résolution par apprentissage

La classe reçoit une définition qui peut être une définition purement descriptive ou une interprétation, par exemple du type diagnostic médical. Le classifieur qui réalisera le classement des formes doit passer par deux phases, une phase d'apprentissage et une phase de test[11].

II.2.1.3.1. Phase d'apprentissage

Le but de l'apprentissage est de découvrir les règles (généralement non déterministes) qui gouvernent et régissent des formes. L'apprentissage est un processus calculatoire qui doit être capable d'amener à une certaine prédiction et à une certaine généralisation. Il existe principalement deux types d'apprentissage, supervisé et non supervisé. Dans le premier cas, on doit apprendre des associations (individus, classes) ; dans le deuxième cas on ne fournit pas d'indications sur les classes[12].

1) Apprentissage supervisé :

L'apprentissage est dit supervisé si les différentes familles de formes, ou classes, sont connues a priori ainsi que l'affectation de chaque forme à telle ou telle famille. Le processus d'apprentissage en mode supervisé peut se résumer comme suit : on dispose d'un ensemble de formes ou d'objets, qui est un ensemble d'apprentissage A d'objets parfaitement classés, ensemble que l'on se propose de reconnaître ; il peut être présenté sous la forme de paires (X_i, C_j) , $1 \leq i \leq N$, $1 \leq j \leq m$, $m < N$, et interprété, « l'objet X_i appartient à la classe C_j ».

On construit dans un premier temps une « machine » sachant classer correctement les objets de l'ensemble d'apprentissage A . La description de chaque forme est prise en compte par la tâche d'apprentissage. Le travail d'apprentissage consiste à analyser les ressemblances entre les formes d'une même classe et les dissemblances entre les formes de classes différentes pour en déduire les meilleures séparations possibles entre classes. Ensuite, on fait en sorte que cette « machine » ait une bonne capacité à prédire la classe d'une forme qui n'est pas dans l'ensemble d'apprentissages A . d'où l'exigence de tester cette machine sur un ensemble de validation et d'objets parfaitement classés. Cette phase de validation est nécessaire car elle aussi de savoir si la machine élaborée n'a pas été entraînée à apprendre « par cœur » sans possibilités de généralisation (sur-apprentissage ou over fitting) [11][9].

2) Apprentissage non supervisé

On l'appelle aussi, suivant l'approche utilisée, classification automatique, ou apprentissage sans professeur ou encore apprentissage par corrélation. Ce type d'apprentissage est utilisé dans le cas où on dispose d'une base d'apprentissage dont les classes ne sont pas définies à l'avance:

soit constitué d'une part d'une liste d'objets X_i et d'autre part d'une donnée du nombre de catégories, soit enfin uniquement composé d'un ensemble d'objets X_i , $1 \leq i \leq N$ [11].

Il s'agit alors de regrouper les différentes formes en classes en fonction d'un critère de similarité choisi a priori. Ce type d'apprentissage permet la construction automatique des classes sans intervention d'un opérateur. On distingue essentiellement deux types de méthodes : les processus de coalescence et la classification hiérarchique. Cette approche nécessite cependant d'avoir une bonne estimation du nombre de classes [9].

II.2.1.3.2. Phase de test

Cette phase doit permettre l'affectation d'un nouvel objet à l'une des classes, au moyen d'une règle de décision intégrant les résultats de la phase d'apprentissage. L'objectif est d'obtenir une estimation la plus fidèle possible du comportement du classifieur dans des conditions réelles d'utilisation. Pour cela, des critères classiques comme les taux de classification et les taux d'erreur sont presque systématiquement utilisés. Mais d'autres critères, comme la spécificité et la sensibilité, apportent aussi des informations utiles [11].

a. Taux de classification et taux d'erreurs

Les taux de classification et d'erreurs permettent d'évaluer la qualité du classifieur C par rapport au problème pour lequel il a été conçu. Ces taux sont évalués grâce à une base de test qui contient des formes décrites dans le même espace de représentation E que celles utilisées pour l'apprentissage.

Elles sont aussi étiquetées par leur classe réelle d'appartenance afin de pouvoir vérifier les réponses du classifieur. Pour que l'estimation du taux de reconnaissance soit la plus fiable possible, il est important que le classifieur n'ait jamais utilisé les échantillons de cette base pour faire son apprentissage (la base de test ne doit avoir aucun objet en commun avec la base d'apprentissage et les éventuelles bases de validation). De plus, cette base de test doit être suffisamment représentative du problème de classification [13].

b. Sensibilité et spécificité

L'évaluation des performances d'un classifieur peut être réalisée par l'appréciation de deux lois statistiques, qui sont la sensibilité $S_e(i)$ et la spécificité :

$$S_e(i) = \frac{VP(i)}{VP(i) + FN(i)} \quad (\text{II. 1})$$

$$S_p(i) = \frac{VN(i)}{VN(i) + FP(i)} \quad (\text{II. 2})$$

Où les grandeurs $VP(i)$, $FN(i)$, $VN(i)$ et $FP(i)$ sont définies dans le tableau suivant :

	Présence d'événement de classe i	Absence d'événement de classe i
Classification Positive	Vrai Positif $VP(i)$	Faux Positif $FP(i)$
Classification Négative	Faux Négatif $FN(i)$	Vrai Négatif $VN(i)$

Tableau II. 1: Les définitions des grandeurs VP, VN, FP et FN.

La sensibilité $S_e(i)$ représente la probabilité de bonne classification de la classe i et la spécificité $S_p(i)$ est une mesure indirecte de la probabilité de fausse alarme puisque cette dernière n'est rien que $1 - S_p(i)$.

II.2.2. Formalisation mathématique d'un problème de classification

La classification apparaît comme une tâche qui consiste à ranger des formes ou objets décrits par un ensemble de variables descriptives en un certain nombre de catégories ou classes définies a priori. Traduit en termes mathématiques, un problème de classification comporte les éléments suivants :

- ✓ **Une population de N objets** $O_i, 1 \leq i \leq N$,
- ✓ **P variables descriptives** X_{id} , qui permettent de décrire les objets ; elles sont aussi appelées plus simplement descripteurs, $1 \leq d \leq P$,
- ✓ **C classes** C_k dans lesquelles on cherche à ranger les objets (k variant de 1 à C). Résoudre un problème de classification, c'est trouver une application de l'ensemble des objets à classer, décrits par les variables descriptives choisies, dans l'ensemble des classes. L'algorithme ou la procédure qui réalise cette application est appelé classifieur. Les variables descriptives considérées ici sont celles qui sont fournies à l'algorithme de classification. Comme indiqué plus haut, elles peuvent être le résultat d'un prétraitement des variables initiales[10].

II.3. Méthodes de classification

Il existe de nombreuses méthodes de classification en IA appliquées dans le domaine médical, parmi celles-ci on trouve la logique floue, les arbres de décision, les réseaux de neurones (notre méthode de classification choisie pour la cardiologie), les réseaux bayésiens, les systèmes expert, les SVM, etc.

II.3.1. La logique floue en classification

Les systèmes d'aide au diagnostic constituent aujourd'hui des moyens techniques indispensables dans le domaine médical notamment la cardiologie. Dans un tel domaine, les méthodes classiques de classification étant inadéquates, la logique floue est considérée comme étant une technique

efficace pour la résolution d'un problème de classification. Le but de cet article est de présenter une approche floue pour la classification des arythmies cardiaques de la famille fibrillation auriculaire. Les résultats obtenus ont montré les bonnes performances du système flou conçu.

Le travail s'intéresse à l'étude d'un classificateur supervisé basé sur les concepts de base de la logique floue. C'est un système de classification floue dont la base de connaissances est modélisée sous forme de règles floues du type « si-alors ». L'objectif principal est la conception d'un système d'aide au diagnostic permettant, à partir d'ECG réels, de détecter trois familles de fibrillation auriculaire à savoir la fibrillation lente (FL), l'arythmie complète par fibrillation auriculaire (ACFA) et la tachycardie complète par fibrillation auriculaire (TACFA)[14].

II.3.1.1. Base des règles floues

La base des règles floues définie comporte 4 règles (tableau II. 2). Elle n'est pas complète car étant donné que le vecteur d'entrée est de dimension 4 et que chaque paramètre peut prendre au moins deux valeurs possibles, il est évident que le nombre de règles qu'on peut former est nettement supérieur à 4.

Ces règles ont été élaborées manuellement. Une fois encore, nous avons adopté l'approche intuitive pour la construction de la base de connaissances bien que plusieurs méthodes de génération et d'apprentissage des règles à partir de données numériques aient été proposées.

Règle	F	RR	P	QRS	Diagnostic
r1	Faible	Irrégulier	Absente	Fin	FL
r2	Moyenne	Irrégulier	Absente	Fin	ACFA
r3	Elevée	Irrégulier	Absente	Fin	TACFA
r4	Moyenne	Régulier	Présente	Moyen	N

Tableau II. 2 : Base des règles floues

II.3.1.2. Avantages et Inconvénients

Les systèmes à base de logique floue ont des formalismes robustes qui ne nécessitent aucun modèle mathématique, ils permettent l'intégration de connaissances à priori sous forme de règles, ils garantissent le suivi et l'interprétation des processus et des résultats, ils sont facilement implémentable et permettent l'extraction des règles.

Par contre, Ils n'ont aucun apprentissage possible, ils ont une adaptation relativement difficile aux modifications des données, ils nécessitent un processus de défusification et l'inconvénient majeur pour nous est qu'ils sont sensibles aux données bruitées, le cas de notre base d'apprentissage.

II.3.1.3. Les arbres de décision

La construction des arbres de décision à partir de données est une discipline déjà ancienne. Les statisticiens en attribuent la paternité à Morgan et Sonquist (1963) qui, les premiers, ont utilisé les arbres de régression dans un processus de prédiction et d'explication (AID Automatic Interaction Detection). Il s'en est suivi toute une famille de méthodes, étendues jusqu'aux problèmes de discrimination et classement, qui s'appuyaient sur le même paradigme de la représentation par arbres (THAID Morgan et Messenger, 1973, CHAID Kass, 1980). On considère généralement que cette approche a connu son apogée avec la méthode CART (Classification and Regression Tree) de Breiman et al (1984) décrite en détail dans une monographie qui fait encore référence aujourd'hui.

II.3.1.4. Exemple

La popularité de la méthode repose en grande partie sur sa simplicité. Il s'agit de trouver un partitionnement des individus que l'on représente sous la forme d'un arbre de décision. L'objectif est de produire des groupes d'individus les plus homogènes possibles du point de vue de la variable à prédire. Il est d'usage de représenter la distribution empirique de l'attribut à prédire sur chaque sommet (nœud) de l'arbre.

Pour mieux appréhender la démarche, nous allons reprendre et dérouler un exemple qui est présenté dans l'ouvrage de Quinlan (1993)[15]. Le fichier est composé de 14 observations, il s'agit d'expliquer le comportement des individus par rapport à un jeu {jouer, ne pas jouer} à partir des prévisions météorologiques.

numéro	Ensoleillement	Température(F)	Humidité(%)	Vent	Jouer
1	Soleil	75	70	OUI	OUI
2	Soleil	80	90	OUI	NON
3	Soleil	85	85	NON	NON
4	Soleil	72	95	NON	NON
5	Soleil	69	70	NON	OUI
6	couvert	72	90	OUI	OUI
7	Couvert	83	78	NON	OUI
8	Couvert	64	65	OUI	OUI
9	couvert	81	75	NON	OUI
10	Pluie	71	80	OUI	NON
11	Pluie	65	70	OUI	NON
12	Pluie	75	80	NON	OUI
13	Pluie	68	80	NON	OUI
14	Pluie	70	96	NON	OUI

Tableau II.3 : Exemple de données

L'arbre de décision correspondant est décrit ci-dessous (figure II.1).

- Le premier sommet est appelé la « racine » de l'arbre. Il est situé sur le premier niveau. Nous y observons la distribution de fréquence de la variable à prédire « Jouer ». Nous constatons qu'il y a bien 14 observations, dont 9 « oui » (ils vont jouer) et 5 « non ».
- La variable « ensoleillement » est la première variable utilisée, on parle de variable de segmentation. Comme elle est composée de 3 modalités {soleil, couvert, pluie}, elle produit donc 3 sommets enfants.
- • La première arête (la première branche), à gauche, sur le deuxième niveau, est produite à partir de la modalité « soleil » de la variable « ensoleillement ». Le sommet qui en résulte couvre 5 observations correspondant aux individus {1, 2, 3, 4, 5}, la distribution de fréquence nous indique qu'il y a 2 « jouer = oui » et 3 « jouer = non ».
- La seconde arête, au centre, correspond à la modalité « couvert » de la variable de segmentation « ensoleillement » ; le sommet correspondant couvre 4 observations, tous ont décidé de jouer (dans le tableau ce sont les individus n°6 à 9). Ce sommet n'ayant plus de sommets enfants, ce qui est normal puisqu'il est « pur » du point de vue de la variable à prédire, il n'y a pas de contre-exemples. On dit qu'il s'agit d'une feuille de l'arbre[15].

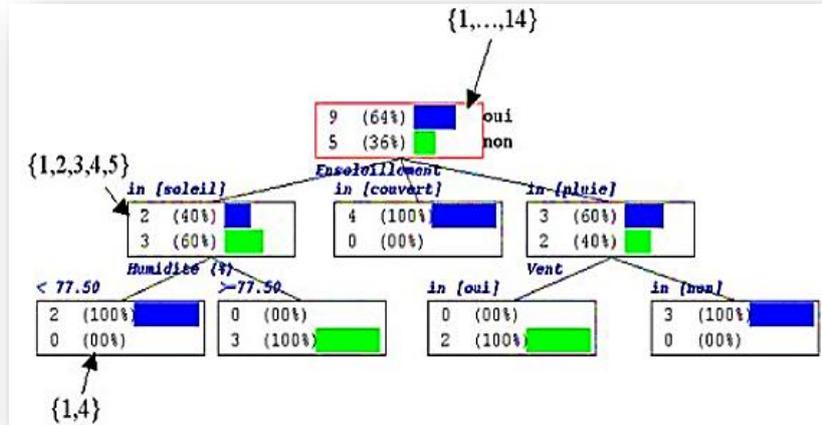


Figure II. 1: Arbre de décision

II.3.2. Avantages et Inconvénients

Les arbres de décisions ont beaucoup d'avantages tels que la manipulation facile des données symboliques, la sollicitation des variables d'amplitudes très différentes et la bonne interprétation des résultats. Par contre ils sont très sensibles au bruit et points aberrants en plus, ils sont valables que pour les échantillons de petite taille et le changement d'une variable changera l'arbre complètement (Problème de mise à jour).

II.3.2.1. Réseaux bayésiens

Les réseaux bayésiens constituent la technologie la plus puissante de ces 10 dernières années en IA et en apprentissage automatique. Les RB constituent un langage graphique et une méthodologie, simples et corrects, pour exprimer pratiquement ce de quoi on est certain ou incertain. Ils reposent sur la formule de Bayes reliant des probabilités conditionnelles avec des probabilités jointes [15].

II.3.2.2. Exemple

Un opérateur travaillant sur une machine risque de se blesser, s'il l'utilise mal. Ce risque dépend de l'expérience de l'opérateur et de la complexité de la machine. «Expérience» et «Complexité» sont deux facteurs déterminants de ce risque (figure II.2) : Structure de causalité. Bien sûr, ces facteurs ne permettent pas de créer un modèle déterministe. Si l'opérateur est expérimenté, et la machine simple, cela ne garantit pas qu'il n'y aura pas d'accident. D'autres facteurs peuvent jouer : l'opérateur peut être fatigué, dérangé, etc. La survenance du risque est toujours aléatoire, mais la probabilité de survenance dépend des facteurs identifiés.

Le schéma ci-dessous représente la structure de causalité de ce modèle

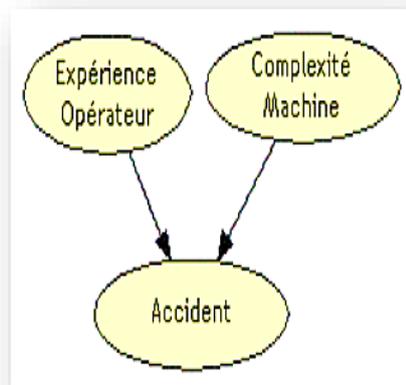


Figure II. 2: Structure de causalité

II.3.2.3. Avantages et Inconvénients

Les aspects suivants des réseaux bayésiens les rendent, dans de nombreux cas, préférables à d'autres modèles :

- ❖ **Acquisition des connaissances** : Les réseaux bayésiens donnent la possibilité de rassembler et de fusionner des connaissances de diverses natures dans un même modèle.
- ❖ **Représentation des connaissances** : La représentation graphique d'un réseau bayésien est explicite, intuitive et compréhensible.

- ❖ **Utilisation des connaissances** : Un réseau bayésien est polyvalent : on peut se servir du même modèle pour évaluer, prévoir, diagnostiquer, ou optimiser des décisions. Mais les réseaux bayésiens ont pas mal d'inconvénients surtout la complexité des algorithmes :
- ❖ **Les variables continues**: La plupart des algorithmes développés pour l'inférence et l'apprentissage dans les réseaux bayésiens, aussi bien que les outils disponibles sur le marché pour mettre en œuvre ces algorithmes utilisent des variables discrètes.
- ❖ **La complexité des algorithmes** : La généralité du formalisme des réseaux bayésiens aussi bien en termes de représentation que d'utilisation les rend difficiles à manipuler à partir d'une certaine taille.

II.3.3. MACHINE A VECTEURS DE SUPPORT (SVM)

Les Support Vector Machines souvent traduit par l'appellation de Séparateur à Vaste Marge (SVM) sont une classe d'algorithmes d'apprentissage initialement définis pour la discrimination c'est-à-dire la prévision d'une variable qualitative binaire. Ils ont été ensuite généralisés à la prévision d'une variable quantitative. Dans le cas de la discrimination d'une variable dichotomique, ils sont basés sur la recherche de l'hyperplan de marge optimale qui, lorsque c'est possible, classe ou sépare correctement les données tout en étant le plus éloigné possible de toutes les observations. Le principe est donc de trouver un classifieur, ou une fonction de discrimination, dont la capacité de généralisation (qualité de prévision) est la plus grande possible.

Cette approche découle directement des travaux de Vapnik en théorie de l'apprentissage à partir de 1995. Elle s'est focalisée sur les propriétés de généralisation (ou prévision) d'un modèle en contrôlant sa complexité. Voir à ce sujet le chapitre concernant la dimension de Vapnik-Chernovenkis qui est un indicateur du pouvoir séparateur d'une famille de fonctions associée à un modèle et qui en contrôle la qualité de prévision. Le principe fondateur des SVM est justement d'intégrer à l'estimation le contrôle de la complexité c'est-à-dire le nombre de paramètres qui est associé dans ce cas au nombre de vecteurs supports. L'autre idée directrice de Vapnik dans ce développement, est d'éviter de substituer à l'objectif initial : la discrimination, un ou des problèmes qui s'avèrent finalement plus complexes à résoudre comme par exemple l'estimation non-paramétrique de la densité d'une loi multidimensionnelle en analyse discriminante[16].

Le principe de base des SVM consiste de ramener le problème de la discrimination à celui, linéaire, de la recherche d'un hyperplan optimal. Deux idées ou astuces permettent d'atteindre cet objectif :

- ❖ La première consiste à définir l'hyperplan comme solution d'un problème d'optimisation sous contraintes dont la fonction objectif ne s'exprime qu'à l'aide de produits scalaires entre vecteurs et dans lequel le nombre de contraintes "actives" ou vecteurs supports contrôle la complexité du modèle.
- ❖ Le passage à la recherche de surfaces séparatrices non linéaires est obtenu par l'introduction d'une fonction noyau (kernel) dans le produit scalaire induisant implicitement une transformation non linéaire des données vers un espace intermédiaire (feature space) de plus grande dimension. D'où l'appellation couramment rencontrée de machine à noyau ou kernel machine. Sur le plan théorique, la fonction noyau définit un espace hilbertien, dit auto-reproduisant et isométrique par la transformation non linéaire de l'espace initial et dans lequel est résolu le problème linéaire

II.3.3.1. Avantages et Inconvénients

L'avantage des SVM est qu'ils s'adaptent facilement aux problèmes non linéairement séparables. Avant de procéder à l'apprentissage de la meilleure séparation linéaire, on transforme les vecteurs d'entrée en vecteurs de caractéristiques de dimension plus élevée. De cette façon, un séparateur linéaire trouvé par un SVM dans ce nouvel espace vectoriel devient un séparateur non linéaire dans l'espace original. Cette transformation des vecteurs se fait à l'aide de noyaux («kernels»).

Dans le cas de la classification de textes, les entrées sont des documents et les sorties sont des catégories. En considérant un classificateur binaire, on voudra lui faire apprendre l'hyperplan qui sépare les documents appartenant à la catégorie et ceux qui n'en font pas partie. Selon les SVM conviennent bien pour la classification de textes parce que, premièrement, une dimension élevée ne les affecte pas puisqu'ils se protègent contre le sur-apprentissage. Dans le même sens, il affirme que peu d'attributs sont totalement inutiles à la tâche de classification et que les SVM permettent d'éviter une sélection agressive qui aurait comme résultat une perte d'information. On peut se permettre de conserver plus d'attributs. Également, une caractéristique des documents textuels est que lorsqu'ils sont représentés par des vecteurs, une majorité des entrées sont nulles. Or, les SVM conviennent bien à des vecteurs dits clairsemés. Un autre aspect positif des SVM est qu'aucun ajustement de paramètres manuel n'est requis, car ils ont l'habileté de trouver automatiquement des paramètres adéquats.

II.3.4. Les carte de kohonen

Un réseau de neurones est un ensemble de cellules interconnectées par des liens ajustables. Cette propriété lui permet de trouver une représentation interne d'un problème donné, le réseau grâce à cette représentation, est capable de reproduire la sortie appropriée pour une entrée qui lui est présentée[17].

L'une des caractéristiques les plus importantes des réseaux de neurones est leur capacité à apprendre, en effet, nous n'avons pas besoin d'une formulation rigoureuse d'un problème donné pour le résoudre, tout ce dont nous avons besoin c'est d'une collection d'exemples représentative de la fonction désirée, ensuite le réseau s'adapte pour reproduire les sorties souhaitées quand un exemple d'entrée est présenté[18].

Les réseaux de neurones à auto-organisation, et en particulier les cartes auto-organisatrices de Kohonen sont une classe importante des réseaux de neurones. Ils ont reçu une importance particulière depuis les travaux de Von der Malsburg et Kohonen[19].

Ces réseaux sont composés d'une grille de neurones (ou nœuds, ou unités) auxquels seront présentés des stimuli. Un stimulus est un vecteur, de dimension d , qui décrit un objet à classer. Ce vecteur peut aussi bien être une description des caractéristiques physiques des objets stimuli qui fait référence à des caractéristiques telle que la présence ou l'absence d'un mot-clé dans un document. Chaque unité de la grille est reliée au vecteur d'entrée (stimulus) par l'intermédiaire de d synapses de poids w . En fait, à chaque unité est associée à un vecteur de dimension d qui contient les poids w (figure II.3)[19].

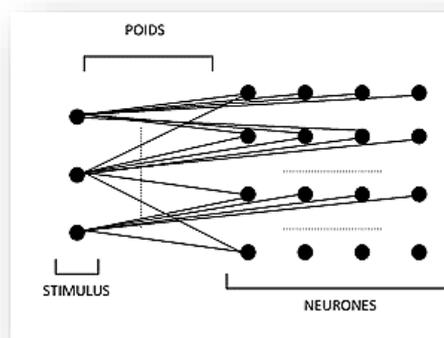


Figure II. 3: Schéma d'une carte de Kohonen.

Les cartes de Kohonen permettent le classement d'objets sur lesquels aucune information relative à leur classification n'est disponible à priori. Une fois le réseau organisé (une fois l'apprentissage effectué), chaque objet est classé dans une zone qui correspond à son « gagnant ».

L'utilisation des cartes de Kohonen dans la recherche documentaire consiste à présenter l'ensemble des documents à la carte pour les classer. Ensuite représenter le classement sous forme graphique compréhensible à l'être humain et qui permettra à l'utilisateur de naviguer dans le fond documentaire pour rechercher l'information désirée[20].

II.3.4.1. Avantages et Inconvénients

Les ancêtres des cartes de Kohonen, les algorithmes comme "k-moyennes", réalisent la discrétisation de l'espace d'entrée en ne modifiant à chaque cycle d'adaptation qu'un seul vecteur référent. Leur processus d'apprentissage est donc très long. L'algorithme de Kohonen profite des relations de voisinage dans la grille pour réaliser une discrétisation dans un temps très court. On suppose que l'espace n'est pas constitué de zones isolées, mais de sous-ensembles compacts. Donc en déplaçant un vecteur référent vers une zone, on peut se dire qu'il y a probablement d'autres zones dans la même direction qui doivent être représentées par des vecteurs référents. Cela justifie le fait de déplacer les neurones proches du vainqueur dans la grille dans cette même direction, avec une amplitude de déplacement moins importante. L'algorithme présente des opérations simples ; il est donc très léger en termes de coût de calculs. Par contre, le voisinage dans les cartes auto adaptatives est malheureusement fixe, et une liaison entre neurones ne peut être cassée même pour mieux représenter des données discontinues.

II.4. Conclusion

Dans ce chapitre, nous avons introduit la notion de classification, formalisme d'un problème de classification, les différentes méthodes de classification existantes (La logique floue, les arbres de décision, les réseaux bayésiens, les SVM- Support Vector Machine et les cartes de Kohonen) et qui permettent de traiter de manière proche ou similaire de point de vue résultats obtenus avec celles qu'on a choisi pour faire la classification des arythmies cardiaques à savoir le réseau de neurones multicouches

Ainsi, Nous avons donné le strict minimum pour chaque méthode en illustrant chacune d'elles par un exemple

Chapitre III

Les réseaux de neurones

III.1. Introduction

Un réseau de neurones est un outil d'analyse statistique permettant de construire un modèle de comportement à partir de données qui sont des exemples de ce comportement. Un réseau de neurones est constitué d'un graphe pondéré orienté dont les nœuds symbolisent les neurones. Ces neurones possèdent une fonction d'activation qui permet d'influencer les autres neurones du réseau. Les connexions entre les neurones, que l'on nomme liens synaptiques, propagent l'activité des neurones avec une pondération caractéristique de la connexion. On appelle poids synaptique la pondération des liens synaptiques.

Notre projet a pour but ici de mettre en place des programmes de réseaux de neurones sous Matlab, qui utilise la rétropropagation du gradient pour identifier différentes courbes.

Les neurones peuvent être organisés de différentes manières, c'est ce qui définit l'architecture et le modèle du réseau. L'architecture la plus courante est celle dite du perceptron multicouche. Avant de pouvoir utiliser les capacités de classification d'un réseau de neurones, il faut le construire. Ceci se déroule en quatre temps:

1. La construction de la structure du réseau (généralement empirique).
2. 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).
3. 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.
4. 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.

III.2. Historique sur les réseaux de neurones

Nous nous intéressons ici à une branche de l'Informatique fondamentale qui, sous l'appellation d'Intelligence Artificielle, a pour objectif de simuler des comportements du cerveau humain. Les premières tentatives de modélisation du cerveau sont anciennes et précèdent même l'ère informatique. C'est en 1943 que Mc Culloch (neurophysiologiste) et Pitts (logicien) ont proposé les

premières notions de neurone formel. Ce concept fut ensuite mis en réseau avec une couche d'entrée et une sortie par Rosenblatt en 1959 pour simuler le fonctionnement rétinien et tacher de reconnaître des formes. C'est l'origine du perceptron. Cette approche dite connexionniste a atteint ses limites technologiques, compte tenu de la puissance de calcul de l'époque, mais aussi théoriques au début des années 70.

L'approche connexionniste à connaissance répartie a alors été supplantée par l'approche symbolique ou séquentielle qui promouvait les systèmes experts à connaissance localisée. L'objectif était alors d'automatiser le principe de l'expertise humaine en associant trois concepts:

- Une base de connaissance dans laquelle étaient regroupées "toutes" les connaissances d'experts humains sous forme de propositions logiques élémentaires ou plus élaborées en utilisant des quantificateurs (logique du premier ordre).
- Une base de faits contenant les observations du cas à traiter comme, par exemple, des résultats d'examens, d'analyses de sang, de salive pour des applications biomédicales de choix d'un antibiotique,
- Un moteur d'inférence chargé d'appliquer les règles expertes sur la base de faits afin d'en déduire de nouveaux faits jusqu'à la réalisation d'un objectif comme l'élaboration du traitement d'une infection bactérienne.

Face aux difficultés rencontrées lors de la modélisation des connaissances d'un expert humain, au volume considérable des bases de connaissance qui en découlaient et au caractère exponentiel de la complexité des algorithmes d'inférence mis en jeu, cette approche s'est éteinte avec les années 80. En effet, il a été montré que les systèmes basés sur le calcul des prédicats du premier ordre conduisaient à des problèmes NP complets.

L'essor technologique et quelques avancées théoriques :

- ❖ Estimation du gradient par rétro-propagation de l'erreur (Hopkins, 1982),
- ❖ Analogie de la phase d'apprentissage avec les modèles markoviens de systèmes de particules de la mécanique statistique (verres de spin) par (Hopfield, 1982).

Au début des années 80 ont permis de relancer l'approche connexionniste. Celle-ci a connu au début des années 90 un développement considérable si l'on considère le nombre de publications et de congrès qui lui ont été consacrés mais aussi les domaines d'applications très divers où elle apparaît. Sur de nombreux objectifs, justement ceux propres au data mining, les réseaux neuronaux ne rentrent pas nécessairement en concurrence avec des méthodes statistiques bientôt centenaires mais apportent un point de vue complémentaire qu'il est important de considérer.

Finalement, la motivation initiale de simulation du cortex cérébral a été rapidement abandonnée alors que les méthodes qui en découlaient ont trouvé leur propre intérêt de développement méthodologique et leurs champs d'applications[21].

III.3. Les réseaux de neurones

III.3.1. Le neurone biologique

Le neurone est une cellule composée d'un corps cellulaire et d'un noyau. Le corps cellulaire se ramifie pour former ce que l'on nomme les dendrites. Celles-ci sont parfois si nombreuses que l'on parle alors de chevelure dendritique ou d'arborisation dendritique. C'est par les dendrites que l'information est acheminée de l'extérieur vers le soma, corps du neurone.

L'information traitée par le neurone chemine ensuite le long de l'axone (unique) pour être transmise aux autres neurones. La transmission entre deux neurones n'est pas directe. En fait, il existe un espace intercellulaire de quelques dizaines d'Angstrom (10^{-9} m) entre l'axone du neurone afférent et les dendrites (on dit une dendrite) du neurone efférent. La jonction entre deux neurones est appelée la synapse (figure III.1).

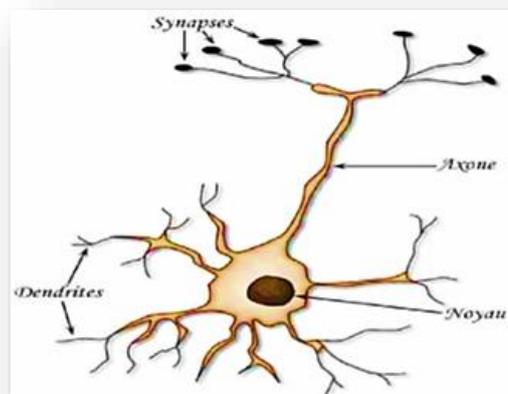


Figure III. 1: Le neurone biologique

III.3.2. Le neurone artificiel et principe de fonctionnement

Le neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones amont. A chacune de ces entrées est associé un poids w abréviation de weight (poids en anglais) représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals. A chaque connexion est associé un poids.

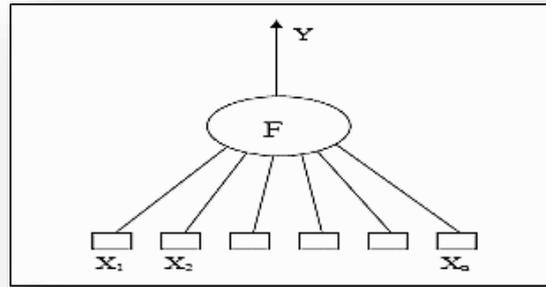


Figure III. 2: Le neurone artificiel.

Le neurone réalise alors trois opérations sur ses entrées :

1. **Pondération** : multiplication de chaque entrée par un paramètre appelé poids de connexion.
2. **Sommation** : une sommation des entrées pondérées est effectuée.
3. **Activation** : passage de cette somme dans une fonction, appelée fonction d'activation.

La valeur calculée est la sortie du neurone qui est transmise aux neurones suivants.

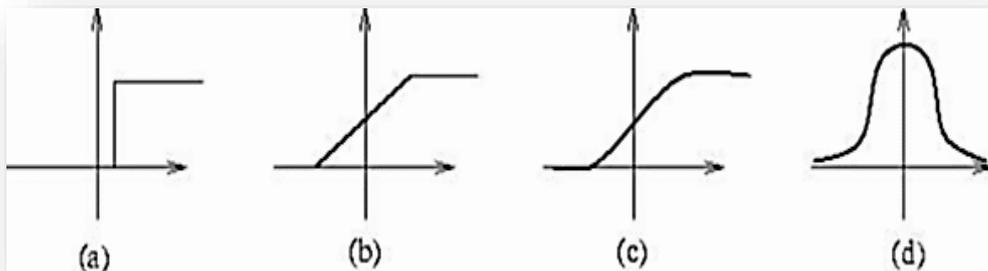


Figure III. 3: Fonctions d'activations.

La fonction d'activation opère une transformation d'une combinaison affine des signaux d'entrée, α_0 étant appelé le biais du neurone. Cette combinaison affine est déterminée par un vecteur de poids $[\alpha_0 ; \alpha_p]$ associé à chaque neurone et dont les valeurs sont estimées dans la phase d'apprentissage. Ils constituent "la mémoire" ou "connaissance répartie" du réseau.

Les différents types de neurones se distinguent par la nature f de leur fonction d'activation. Les principaux types sont :

- a. Seuil
- b. Linéaire par morceau
- c. Sigmoide
- d. Gaussienne

Les modèles linéaires et sigmoïdaux sont bien adaptés aux algorithmes d'apprentissage impliquant (cf. ci-dessous) une rétro-propagation du gradient car leur fonction d'activation est différentiable, ce sont les plus utilisés. Le modèle à seuil est sans doute plus conforme à la "réalité" biologique mais pose des problèmes d'apprentissage. Enfin le modèle stochastique est utilisé pour des problèmes d'optimisation globale de fonctions perturbées ou encore pour les analogies avec les systèmes de particules. On ne le rencontre pas en *data mining*.

III.3.3. Réseau de neurones

Un réseau de neurones artificiel est une structure composée d'entités capables de calcul et interagissant entre eux, les neurones. Il permet de traiter, par le biais de l'informatique, des problèmes de différentes natures que les outils classiques ont du mal à résoudre. En effet, son fonctionnement s'inspire de celui des cellules neuronales animales, et est donc différent des méthodes de calcul analytiques que l'on utilise ordinairement. Il s'avère très puissant dans des problèmes de reconnaissance, classification, approximation ou prévision. Étudions de plus près le fonctionnement d'un réseau neuronal classique.

III.3.4. Architecture des réseaux de neurones

Il existe principalement deux sortes de réseaux :

III.3.4.1. Les réseaux de neurones non bouclés -feed forward

III.3.4.1.1 Présentation

Un réseau de neurones non bouclés est un assemblage de neurones organisés en couches chaque neurone d'une couche peut avoir des connexions à des neurones de la couche inférieure et de la couche supérieure, mais pas aux autres neurones (pas même ceux de la même couche).

La première couche constitue les neurones d'entrée (c'est la couche d'entrée, ou couche sensorielle), et la dernière couche est la couche de sortie.

S'il n'y a que la couche d'entrée et celle de sortie, on parle de réseau monocouche, sinon de réseau multicouche, et les couches intermédiaires sont dites couches cachées.

III.3.4.1.2 Un peu d'histoire

En 1958, F. Rosenblatt créa le premier réseau de neurones cherchant à comprendre et copier le modèle biologique le Perceptron.

La théorie présentée ici traite d'une classe de (modèles cérébraux) appelé perceptron. Par modèle c'cébral, nous entendons tout système théorique qui tente d'expliquer le fonctionnement psychologique du cerveau en termes de lois physiques et mathématiques connues et de faits connus en neuro anatomie et physiologie. Les perceptrons sont :

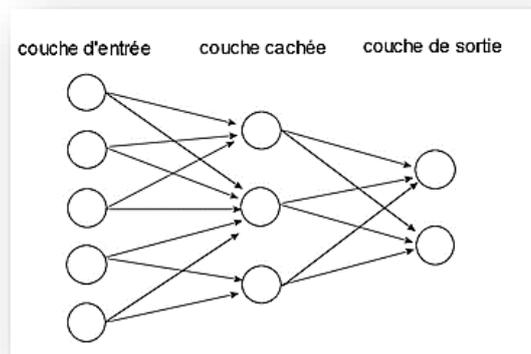


Figure III. 4: Exemple de réseau non-bouclé

III.3.4.1.3 Composition du perceptron

Dans les modèles de base, le Perceptron est formé des éléments suivants :

- Des unités sensibles sensibles à des stimulations de l'environnement. En général, si les stimulus dépassent un certain seuil θ , le signal +1 est renvoyé, sinon on retourne le signal 0.
- Des unités d'association disposant de connexions en entrée et en sortie. Les entrées sont pondérées, et si la somme des entrées pondérées dépasse le seuil θ , alors le signal +1 est retourné.
- Des unités de réponse qui agissent comme les unités d'association mais qui sont en bout de chaîne du perceptron.
- Et enfin une matrice d'interaction qui définit les coefficients de couplage (Aussi appelés poids de pondération).

III.3.4.2. Les réseaux de neurones bouclés

Les réseaux de neurones bouclés sont quasiment identiques aux réseaux non-bouclés à ceci près que les différents neurones peuvent s'interconnecter comme ils le souhaitent, sans aucune restriction de couche. Le fait qu'il puisse ainsi se former des cycles entre les différents neurones permet de stocker de l'information au cœur du réseau et ainsi obtenir une mémoire vive. Le fonctionnement et l'étude de tels réseaux en devient forcément plus complexe c'est pourquoi nous ne traiterons pas cette partie des réseaux de neurones.

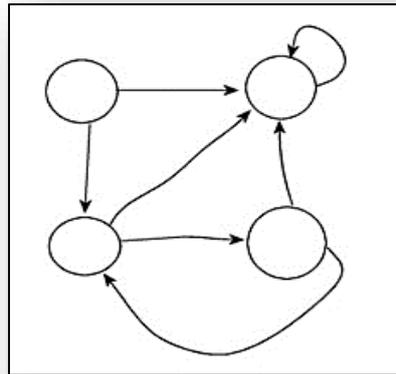


Figure III. 5: Exemple de réseau bouclé

III.3.5. Apprentissage des réseaux de neurones

L'apprentissage d'un réseau de neurones consiste à déterminer les poids optimaux (de la force optimale des connexions) suivant le problème à résoudre. Rappelons que l'architecture du réseau est déjà déterminée à ce stade.

III.3.5.1. Type d'apprentissage

On distingue deux types de réseaux de neurones en fonction du type d'apprentissage auquel ils sont soumis :

III.3.5.1.1 Les réseaux à apprentissage supervisé

Les réseaux à apprentissage supervisé qui sont généralement destinés à reproduire un processus quelconque (chimique, mécanique, financier...) dont on connaît seulement quelques variables et les résultats correspondant.

L'apprentissage supervisé d'un réseau de neurones consiste à en modifier les poids tant que la réponse correspondant à chaque entrée n'est pas assez proche de la réponse souhaitée. Nous allons ici examiner un cas simple d'apprentissage, celui du perceptron et de l'adaline. Puis, nous nous intéresserons à la méthode d'apprentissage non supervisée la plus utilisée : la rétro-propagation du gradient des réseaux multicouches.

III.3.5.1.2 Les réseaux à apprentissage non supervisé

Les réseaux à apprentissage non supervisé utilisés par exemple en classification lorsque les classes auxquelles doivent appartenir les données ne sont pas connus a priori. Les données à l'apprentissage sont donc différentes dans les deux cas, puisqu'elles consistent en une série de couples (entrée, sortie désirée correspondante) pour les apprentissages supervisés, tandis que pour les apprentissages non supervisés, les données sont uniquement des entrées.

L'apprentissage non supervisé consiste à ajuster les poids à partir d'un seul ensemble d'apprentissage formé uniquement de données. Aucun résultat désiré n'est fourni au réseau.

Qu'est-ce que le réseau apprend exactement dans ce cas ? L'apprentissage consiste à détecter les similarités et les différences dans l'ensemble d'apprentissage. Les poids et les sorties du réseau convergent, en théorie, vers les représentations qui capturent les régularités statistiques des données[22]. Ce type d'apprentissage est également dit compétitif et (ou) coopératif. L'avantage de ce type d'apprentissage réside dans sa grande capacité d'adaptation reconnue comme une auto-organisation, « *self-organizing* ». L'apprentissage non supervisé est surtout utilisé pour le traitement du signal et l'analyse factorielle[23].

III.3.5.2. Algorithmes d'apprentissage

L'algorithme d'apprentissage est la méthode mathématique qui va modifier les poids de connexions afin de converger vers une solution qui permettra au réseau d'accomplir la tâche désirée. L'apprentissage est une méthode d'identification paramétrique qui permet d'optimiser les valeurs des poids du réseau.

Plusieurs algorithmes itératifs peuvent être mis en œuvre, parmi lesquels on note : l'algorithme de rétro-propagation (ce que nous allons détailler ci-après), la Méthode Quasi-Newton, La méthode de Newton, etc...

III.3.5.2.1 Algorithme de rétro-propagation

L'algorithme de rétro-propagation ou de propagation arrière « *back propagation* » est l'exemple d'apprentissage supervisé le plus utilisé à cause de l'écho médiatique de certaines applications spectaculaires telles que la démonstration de *Sejnowski* et *Rosenberg* (1987) dans laquelle l'ARP est utilisé dans un système qui apprend à lire un texte. Un autre succès fut la prédiction des cours du marché boursier et plus récemment la détection de la fraude dans les opérations par cartes de crédit.

La technique de rétro-propagation du gradient (*Back propagation* en anglais) est une méthode qui permet de calculer le gradient de l'erreur pour chaque neurone du réseau, de la dernière couche

vers la première. L'historique des publications montre que l'ARP a été découvert indépendamment par différents auteurs mais sous différentes appellations. Le principe de la rétro-propagation peut être décrit en trois étapes fondamentales : acheminement de l'information à travers le réseau, rétro-propagation des sensibilités et calcul du gradient, ajustement des paramètres par la règle du gradient approximé[24].

III.4. Conception et mise en œuvre des réseaux de neurones

Les réseaux de neurones réalisent des fonctions non linéaires paramétrées. Leurs mises en œuvre nécessitent :

- ✚ La détermination des entrées et des sorties pertinentes, c'est à dire les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser (classification,...).
- ✚ La collecte des données nécessaires à l'apprentissage et à l'évaluation des performances du réseau de neurones.
- ✚ La détermination du nombre de neurones cachés nécessaires pour obtenir une approximation satisfaisante.
- ✚ La réalisation de l'apprentissage.
- ✚ L'évaluation des performances du réseau de neurones à l'issue de l'apprentissage.

III.4.1. Détermination des entrées/sorties du réseau de neurones

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

- Premièrement, la dimension intrinsèque du vecteur des entrées doit être aussi petite que possible, en d'autre terme, la représentation des entrées doit être la plus compacte possible, tout en conservant pour l'essentiel la même quantité d'information, et en gardant à l'esprit que les différentes entrées doivent être indépendantes.
- En second lieu, toutes les informations présentées dans les entrées doivent être pertinentes pour la grandeur que l'on cherche à modéliser : elles doivent donc avoir une influence réelle sur la valeur de la sortie.

III.4.2. Choix et préparation des échantillons

Le processus d'élaboration d'un réseau de neurones commence toujours par le choix et la préparation des échantillons de données. La façon dont se présente l'échantillon conditionne le type de réseau, le nombre de cellules d'entrée, le nombre de cellules de sortie et la façon dont il faudra mener l'apprentissage, les tests et la validation. Il faut donc déterminer les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser[25].

Lorsque la grandeur que l'on veut modéliser dépend de nombreux facteurs, c'est-à-dire lorsque le modèle possède de nombreuses entrées, il n'est pas possible de réaliser un « pavage » régulier dans tout le domaine de variation des entrées : il faut donc trouver une méthode permettant de réaliser uniquement des expériences qui apportent une information significative pour l'apprentissage du modèle. Cet objectif peut être obtenu en mettant en œuvre un plan d'expériences. Pour les modèles linéaires, l'élaboration de plans d'expériences est bien maîtrisée, par ailleurs, ce n'est pas le cas pour les modèles non linéaires.

Afin de développer une application à base de réseaux de neurones, il est nécessaire de disposer de deux bases de données, une pour effectuer l'apprentissage et l'autre pour tester le réseau obtenu et déterminer ses performances.

Notons qu'il n'y a pas de règle pour déterminer ce partage d'une manière quantitative, néanmoins chaque base doit satisfaire aux contraintes de représentativité de chaque classe de données et doit généralement refléter la distribution réelle, c'est à dire la probabilité d'occurrence des diverses classes[26].

III.4.3. Elaboration de la structure du réseau

La structure du réseau dépend étroitement du type des échantillons. Il faut d'abord choisir le type de réseau : un perceptron standard, un réseau de *Hopfield*, un réseau à décalage temporel, un réseau de *Kohonen*, un Perceptron Multicouches, etc.

Par exemple, dans le cas de ce dernier, il faudra aussi bien choisir le nombre de couches cachées que le nombre de neurones dans cette couche.

a. Nombre de couches cachées

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

b. Nombre de neurones cachés

Chaque neurone peut prendre en compte des profils spécifiques de neurones d'entrée. Un nombre plus important permet donc de mieux "coller" aux données présentées mais diminue la capacité de généralisation du réseau. Il faut alors trouver le nombre adéquat de neurones cachés nécessaire pour obtenir une approximation satisfaisante.

Il n'existe pas, à ce jour, de résultat théorique permettant de prévoir le nombre de neurones cachés nécessaires pour obtenir une performance spécifique du modèle, compte tenu des modèles disponibles. Il faut donc nécessairement mettre en œuvre une procédure numérique de conception de modèle.

III.4.4. Apprentissage

L'apprentissage est un problème numérique d'optimisation. Il consiste à calculer les pondérations optimales des différentes liaisons, en utilisant un échantillon.

La méthode la plus utilisée est la rétro-propagation, qui est généralement plus économe que les autres en termes de nombres d'opérations arithmétiques à effectuer pour évaluer le gradient.

Pour rendre l'optimisation plus performante, on peut utiliser des méthodes de second ordre. Le calcul est très efficace, mais lourd. Elles ont de nombreuses limitations, quant aux conditions de convergence, sur les dérivées secondes. Des corrections sont proposées pour éviter ce problème, et sont prises en compte par les méthodes dites de Quasi-Newton ou de Newton modifiée.

III.4.5. Validation et Tests

Alors que les tests concernent la vérification des performances d'un réseau de neurones hors échantillon et sa capacité de généralisation, la validation est parfois utilisée lors de l'apprentissage. Une fois le réseau de neurones développé, des tests s'imposent afin de vérifier la qualité des prévisions du modèle neuronal.

Cette dernière étape doit permettre d'estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage. Une validation rigoureuse du modèle développé se traduit par une proportion importante de prédictions exactes sur l'ensemble de la validation. Si les performances du réseau ne sont pas satisfaisantes, il faudra, soit modifier l'architecture du réseau, soit modifier la base d'apprentissage.

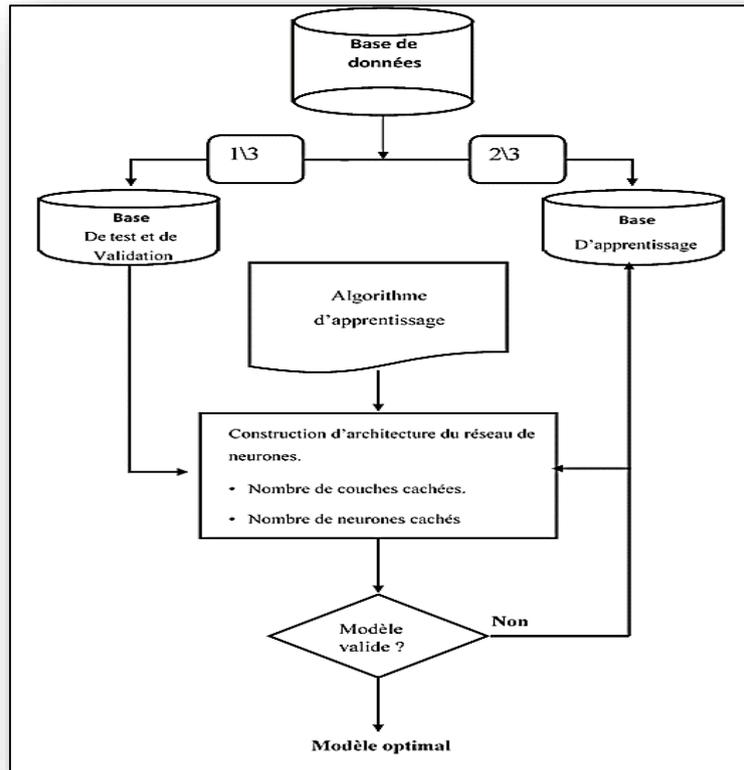


Figure III. 6: Organigramme de conception d'un réseau de neurones

III.5. Apprentissage par les réseaux multicouches

Un réseau de neurones multicouches (ou Perceptron Multicouches « PMC») est composé d'une couche d'entrée, d'une ou plusieurs couches cachées et d'une couche de sortie.

III.5.1. Algorithme d'apprentissage

L'algorithme d'apprentissage par rétro-propagation impose l'utilisation d'une fonction d'activation continue. Cet algorithme consiste à comparer la sortie obtenue et la sortie désirée, puis à propager la différence en sens inverse en modifiant les poids entre chaque couche. C'est un algorithme itératif, qui nécessite la présentation de l'ensemble d'apprentissage un grand nombre de fois. Nous présentons ci-dessous les différentes étapes de cet algorithme :

1. Initialisation des poids du réseau.
2. Présentation du vecteur d'apprentissage à l'entrée du réseau (couche d'entrées).
3. Calcul du vecteur de sortie S: le système propage les activités neuronales (état excité ou inhibé des neurones) à travers le réseau. Chaque neurone calcule la somme pondérée de ses entrées et transmet le résultat par une fonction de type sigmoïde pour produire sa valeur de sortie. Le vecteur de sortie S est le résultat du calcul de la dernière couche (couche de sortie).

4. Calcul de l'erreur: On compare alors les valeurs de sortie actuelles S avec les valeurs désirées de R appelé vecteur de référence (ou encore vecteur cible). On en déduit l'erreur J , définie comme la somme des carrés des différences observée sur chaque unité de la couche de sortie.
5. Rétro-propagation de l'erreur: l'algorithme de rétro-propagation du gradient permet récursivement de rétro-propager l'erreur de la couche de sortie vers les couches cachées jusqu'à la première couche du réseau.
6. Modification des poids: les poids de chaque neurone sont modifiés, soit à chaque présentation d'un vecteur d'apprentissage (méthode du gradient stochastique), soit après cumul de l'erreur pour un certain nombre de vecteurs d'apprentissage (méthode du gradient standard).

Il y a principalement deux facteurs qui influent sur l'apprentissage. Ce sont la qualité de l'échantillonnage d'apprentissage (les exemples qui constituent la base d'apprentissage) et la diversité des valeurs. En effet, le réseau de neurones généralisera mieux si la qualité de l'échantillonnage est meilleure et si les données des exemples d'apprentissage sont variées. Intuitivement, on est conscient que s'il sait répondre correctement pour un nombre fini de situations les plus diverses, il sera alors plus proche de ce que l'on veut dans une situation nouvelle.

III.5.1.1. Algorithme de la rétro-propagation du gradient

L'algorithme va être donné dans sa version la plus générale, c'est à dire avec plusieurs couches cachées. On notera f la fonction d'activation. Une démonstration de l'efficacité de cet algorithme exige de la fonction d'activation qu'elle soit indéfiniment dérivable. On notera S_i la sortie du neurone i de la couche de sortie et y_i la sortie attendue pour ce même neurone. Enfin, pour des neurones d'une couche cachée, on notera O_i la sortie calculée du neurone.

Entrée : un exemple, sous la forme (vecteur_x, vecteur_y);
 Epsilon le taux d'apprentissage
 Un Perceptron Multicouches avec $q-1$ couches cachées C_1, \dots, C_{q-1} , une couche de sortie C_q .

Répéter

Prendre un exemple (vecteur_x, vecteur_y) et calculer $f(\text{vecteur}_x)$

Pour toute cellule de sortie i $d_i \leftarrow S_i(1 - S_i)(y_i - S_i)$ **fin Pour**

Pour chaque couche de $q-1$ à 1

Pour chaque cellule i de la couche courante

$d_i = O_i(1 - O_i) * \text{Somme [pour } k \text{ appartenant aux}$

indices des neurones prenant en entrée la

sortie du neurone i] de $d_k * w_{ki}$

Fin Pour

Fin Pour

Pour tout poids $w_{ij} \leftarrow w_{ij} + \text{Epsilon} * d_i * x_{ij}$ **fin Pour**

Fin Répéter

La variable "di" apparaît deux fois dans le code. Il s'agit de deux variables différentes, car en fait on suppose que les neurones sont numérotés de sorte que l'on puisse associer à un identifiant à un neurone et réciproquement. Par conséquent, le 'i' de "di" identifie un neurone et ainsi on peut effectuer la dernière boucle de manière uniforme sans différencier pour la couche de sortie et les couches cachées.

La méthode donne de bons résultats pratiques. Dans la plupart des cas, on rencontre peu de problèmes dus aux minima locaux comme va l'expliquer ci-après, mais il y en a. Toutefois, il est moins performant que d'autres algorithmes de propagation d'erreur : il tend moins rapidement vers des poids plus ou moins optimaux.

III.6. Conclusion.

Tout ce passe comme si on utilisait le réseau dans deux directions, un calcul 'en avant' pour déterminer la réponse du réseau à une 'stimulation' puis un calcul en 'arrière' pour évaluer les modifications à apporter aux connexions afin d'améliorer le comportement global du système. Il est important de remarquer que le réseau ne peut influencer que sur les poids des connexions et (éventuellement) sur les seuils des cellules. Il ne peut, en aucun cas modifier son architecture ni les fonctions de transferts utilisées.

De plus pour que le calcul soit possible il est impératif que les fonctions de transfert soient continués (et donc dérivables).

Par ailleurs, du fait même des propriétés sur le calcul du gradient on n'a plus la garantie de la convergence, c'est pourquoi on a besoin d'introduire une borne sur le nombre d'itérations (époques). De plus, comme on effectue des calculs sur machine, on sait que ces calculs sont entachés d'erreur, il devient donc illusoire de vouloir atteindre les valeurs optimales, d'où la nécessité d'introduire un seuil minimal d'erreur (ϵ). Enfin, comme le processus d'apprentissage est initialisé aléatoirement, il est nécessaire d'effectuer plusieurs simulations avec les mêmes paramètres initiaux - ceci nous assure une stabilité statistique de l'apprentissage.

On peut remarquer que le réseau est sujet à de très nombreux paramètres, la plupart sont à la charge du concepteur et peu de latitude est laissée au réseau. Une fois choisies l'architecture (nature du réseau, nombre de couches, nombre de cellules par couches) et les bases d'apprentissage, de test et de validation, il reste encore de nombreux paramètres à déterminer, parmi ceux-ci on peut citer :

- 1.-La fonction de transfert
- 2.-La fonction d'erreur
- 3.-Le pas d'apprentissage

Chapitre IV
Contribution Et
Réalisation

IV.1. Introduction

Ce chapitre a pour objectif de présenter le travail réalisé pour donner une solution de la problématique posée - Une approche neuronale pour la classification distribuée par descripteurs temporels des arythmies cardiaques.

Nous allons donc, voir comment développer un réseau de neurones artificiels (ex, PMC perceptron Multicouches) de l'étape de sélection de la base d'apprentissage et de test, sélection des descripteurs temporels pertinents d'un cycle cardiaque étudié et l'architecture des réseaux de neurones (Algorithme d'apprentissage, dimension de ce réseau) et bien sûr discuter les résultats expérimentaux des classifieurs.

IV.2. Classification neuronale des arythmies cardiaques

Comme tout modèle de traitement, les classifieurs utilisés en reconnaissance des formes ou d'objets sont amenés à effectuer des tâches complexes (notamment lors de l'apprentissage), à partir d'une base de données réelles.

Pour que le classifieur soit opérationnel en phase d'exploitation, il est primordial que la nature des données et notamment leurs variations soient correctement prises en compte. Le problème se profile à plusieurs niveaux, comportant chacun son lot de difficultés théoriques et pratiques. Il faut disposer premièrement d'un système d'acquisition efficace, qui nous fournira des données de qualité (avec un minimum de bruit). Il faut ensuite pouvoir représenter cette information (possiblement la réduire et la transformer) de manière efficace et compatible avec les étapes de traitement ultérieur. Cette représentation et ses modalités constituent le modèle avec lequel le mécanisme de reconnaissance devra travailler, afin de produire le résultat désiré.

Dans certains domaines tels que le médical, le principal obstacle auquel est confronté le concepteur lors de la mise en œuvre d'un système de classification est la caractérisation des données par des paramètres fiables et pertinents.

L'analyse de la littérature présentée dans le chapitre des modèles de classification montre que, lorsqu'une approche de reconnaissance des arythmies cardiaques est adoptée, les réseaux de neurones artificiels sont une bonne technique de classification.

Pour notre étude, le réseau de neurones utilisé est optimisé par des algorithmes de type supervisé. La stratégie que nous envisageons en œuvre est présentée dans la figure IV.1.

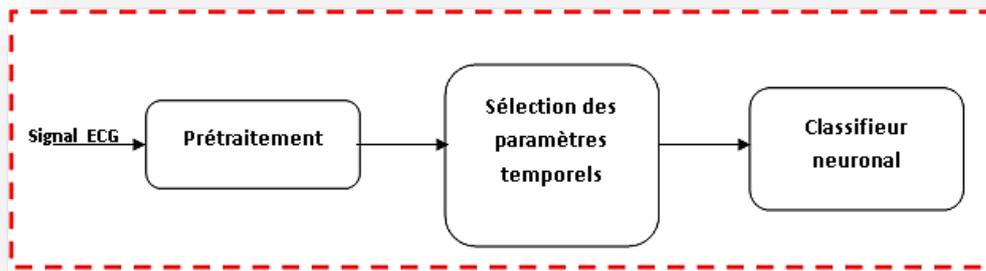


Figure IV. 1: Schéma global de la classification

Les étapes : « Prétraitement », « Sélection des paramètres temporels » et « Classifieur neuronal » font partie de la classification neuronale en basant sur des caractéristiques temporelles extraites depuis le signal ECG de l'entrée, puisqu'après le prétraitement (dé bruitage,...), on passe directement vers l'extraction des paramètres temporelles telles que : les périodes RR0, RRs, PP, le complexe QRS et le rapport RRs/RR0

Et comme nous l'avons précédemment mentionné dans le chapitre III, dans le paragraphe III.4, la mise en œuvre de chaque réseau de neurones nécessite :

1. La détermination des entrées et des sorties pertinentes, c'est à dire les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.
2. La collecte des données nécessaires à l'apprentissage et à l'évaluation des performances du réseau de neurones c'est-à-dire la sélection de base d'exemple et base de test ou de validation.
3. La détermination du nombre de neurones cachés nécessaires pour obtenir une approximation satisfaisante.
4. La réalisation de l'apprentissage.
5. L'évaluation des performances du réseau de neurones à l'issue de l'apprentissage.

IV.2.1 Sélection de la base d'exemples

Nous avons utilisé, dans le cadre de ce mémoire, la base de données MIT-BIH. Cette dernière se compose de quarante-huit enregistrements (deux voies ECG ambulatoire d'environ trente minutes), dont les instants d'apparition et la classe de tous les cycles sont annotés par deux cardiologues différents (pour plus d'information, voir annexe B). Les enregistrements comprenant suffisamment de battements normaux ont été choisis pour construire la base d'exemple, voir tableau IV.1.

Nous avons rassemblé les battements de chaque enregistrement en deux groupes (classes) :

- ✚ Le groupe N est composé des battements différents des « Normal ».
- ✚ Le groupe A est composé des battements autres que normal.

Il est important de remarquer que tous les enregistrements choisis pour faire l'expérimentation sont des battements normaux, mais ils contiennent des cycles cardiaques anormaux.

Toutes ces informations sont récapitulées dans le tableau IV.1 suivant :

Enregistrement	Nombre de battements de type « N »	Nombre de battements de type « A »
100	2239	34
101	1860	05
103	2082	02
105	2526	46
112	2537	02
116	2302	110
203	2529	511
205	2571	85
209	2621	384
210	2423	205
213	2641	610
215	3195	168
230	2255	01

Tableau IV. 1: Les enregistrements choisis de la base de données MIT-BIH.

La base d'apprentissage affecte la performance des classifieurs neuronaux. Le critère général à respecter est de sélectionner des exemples représentatifs de toutes les classes. Dans notre cas, nous avons préparé deux bases, une d'apprentissage et l'autre de validation ou de test.

Notre expérimentation est composée de 60 cycles de Type N (battement normal) et 36 cycles de Type « A » (mélange entre les BVP, BBDR, BBG), qui vaut 96 cycles pour chaque base choisie parmi les 13 enregistrements sélectionnés.

IV.2.1.1 Sélection des descripteurs temporaires d'un cycle cardiaque

Pour obtenir une meilleure performance d'un classifieur, il faut choisir des données pertinentes permettant de décrire de manière assez complète l'échantillon à classer. Ces données sont appelées caractéristiques ou descripteurs.

Dans notre cas, en plus le complexe QRS, des descripteurs permettant de caractériser un battement cardiaque sont retenus. Ils sont comparables aux paramètres sur lesquels sont construites les connaissances médicales.

❖ Intervalle RRp

Nous appelons RRp la distance entre le pic R du présent battement et le pic R du battement précédent. Nous avons divisé la valeur de la distance RRp par la moyenne des distances des huit intervalles RR précédents. Ce rapport est au voisinage de la valeur 1 pour un battement Normal et un peu plus petit que 1 pour un battement de type « A ».

❖ Rapport des intervalles RR

Nous appelons RRs la distance entre le pic R du présent battement et le pic R du battement suivant. Dans le cas d'un rythme régulier, le rapport RRs / RRp est un paramètre qui caractérise une classe donnée. Dans le cas d'un rythme régulier « N », ce rapport est voisin de 1, mais il peut largement dépasser cette valeur dans le cas « A ».

❖ Largeur du complexe QRS

Ce paramètre est important pour l'identification des battements cardiaques, la mesure de ce paramètre a été réalisée en utilisant un algorithme implémenté sous matlab, au laboratoire L.T.S.I à l'université de Rennes 1.

❖ Distance PP entre deux points minimum et maximum du QRS

Ce paramètre a été aussi retenu comme valeur informationnelle potentiellement pertinente pour chaque cycle cardiaque, sachant que cette distance varie énormément avec la déformation du complexe QRS.

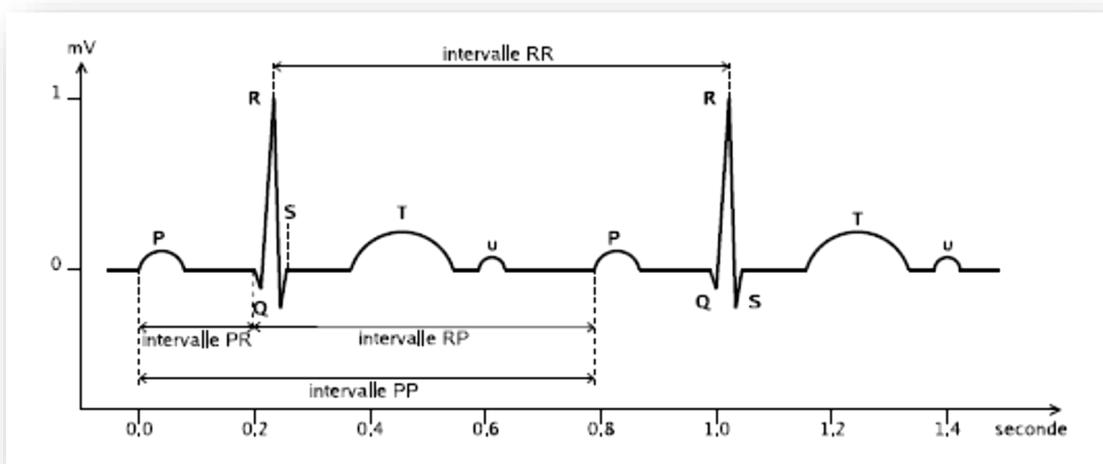


Figure IV. 2: Battements normaux et intervalles d'intérêt

IV.2.2 Architecture du réseau de neurones

IV.2.2.1 Algorithme d'apprentissage

On a développé nos réseaux de neurones (Par des descripteurs morphologiques et d'autres temporels) sous Matlab version R2012a. Au cours de ce travail nous avons utilisé principalement trois algorithmes très classiques en optimisation non linéaire.

1. Quasi-Newton.
2. Levenberg-Marquardt.
3. Méthode du gradient à pas.

En les confrontant sur plusieurs problèmes, il apparaît qu'aucune de ces trois méthodes ne prend un avantage considérable sur l'autre. Chaque méthode d'optimisation possède des avantages et des inconvénients. La méthode de Levenberg-Marquardt présente un intérêt pratique car elle peut être utilisée sans avoir à choisir le pas.

Bien entendu, aucune méthode ne conduit à coup sûr au minimum global. Il convient donc de se placer dans des conditions où les minima locaux sont aussi peu nombreux que possible ou correspondent à une erreur suffisamment faible.

IV.2.2.2 Dimensionnement du réseau pour chaque classifieur

La structure neuronale est déterminée par le nombre de neurones de la couche d'entrée, N_e et celui de la couche cachée, N_c , tous les deux sont variables.

- Couche d'entrée: le nombre de neurones sur cette couche est lié directement au vecteur de caractéristiques
- Le classifieur temporaire, le nombre de neurones dans la couche d'entrée est 96 neurones, puisque, nous avons choisi des descripteurs donnant un vecteur de taille est égale à ce nombre (96).
- Couche cachée : le choix de la taille de la couche cachée est réalisée souvent de façon heuristique et expérimentale, car aucune formule théorique ne peut nous aider à l'optimiser. Donc, nous avons mené une série d'essais pour arrêter le nombre N_c à mettre sur la couche cachée afin d'avoir un meilleur résultat possible. On a constaté que plus le nombre de neurones cachés est important, plus le nombre d'exemples bien classés est grand, car le réseau, grâce au nombre de paramètres croissant, se spécialise sur la base d'apprentissage. Les fonctions d'activation des neurones cachés sont des sigmoïdes.

- Le nombre d'itérations, fixé dans ces deux essais à 30000, a permis de réaliser un apprentissage satisfaisant. Par ailleurs, il est clair que ce dernier paramètre est important s'il est trop grand, le réseau risque de réaliser un sur-apprentissage, au contraire, s'il est trop petit l'apprentissage ne sera pas conduit à son terme.

IV.2.3 Résultats expérimentaux des classifications

Dans ce qui suit, on présente de façon résumée les résultats des d'un classifieur obtenu. Les performances sont représentées respectivement par le taux de classification t_{class} , la sensibilité S_e et la spécificité S_p déjà définis dans le chapitre des modèles de classification.

Le tableau suivant récapitule les résultats de classification faite sur une base d'apprentissage, celui-ci est une base contenant les descripteurs temporels

Enregistrement	t_{class} (%)	S_e (%)	S_p (%)
100	92.44	97.12	98.34
101	89.14	96.22	97.33
103	91.87	90.29	99.45
105	93.34	88.45	93.34
112	75.34	87.17	93.40
116	89.77	90.55	96.66
203	95.02	97.33	99.35
205	97,67	79,88	87,77
209	96,77	96,99	90,89
210	92,59	98,77	89,07
213	96,34	90,02	99,9
215	98,56	97,09	100
230	97,55	91,30	95,45
Moyenne	92.40	92.49	96.83

Tableau IV. 2: Performances du classifieur neuronal à données temporelles.

IV.3. Architecture des descripteurs temporels

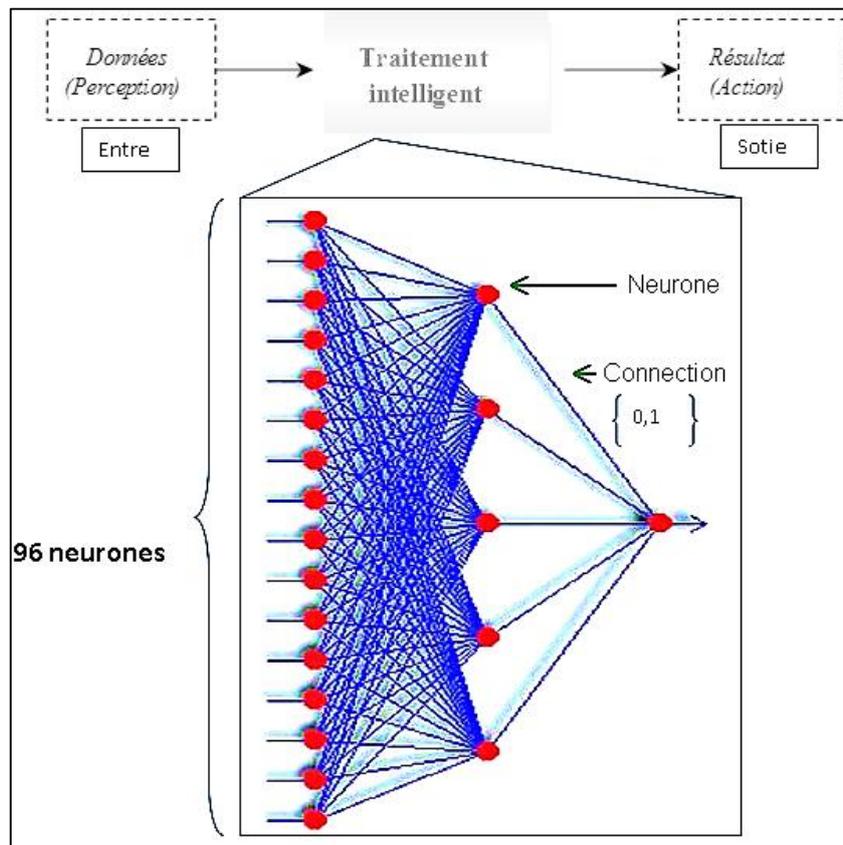


Figure IV. 3: Architecture des descripteurs temporels.

Après la compilation du programme ; nous devons choisir un signal parmi les 13 cas étudiés pour visualiser l’ECG de ce dernier.

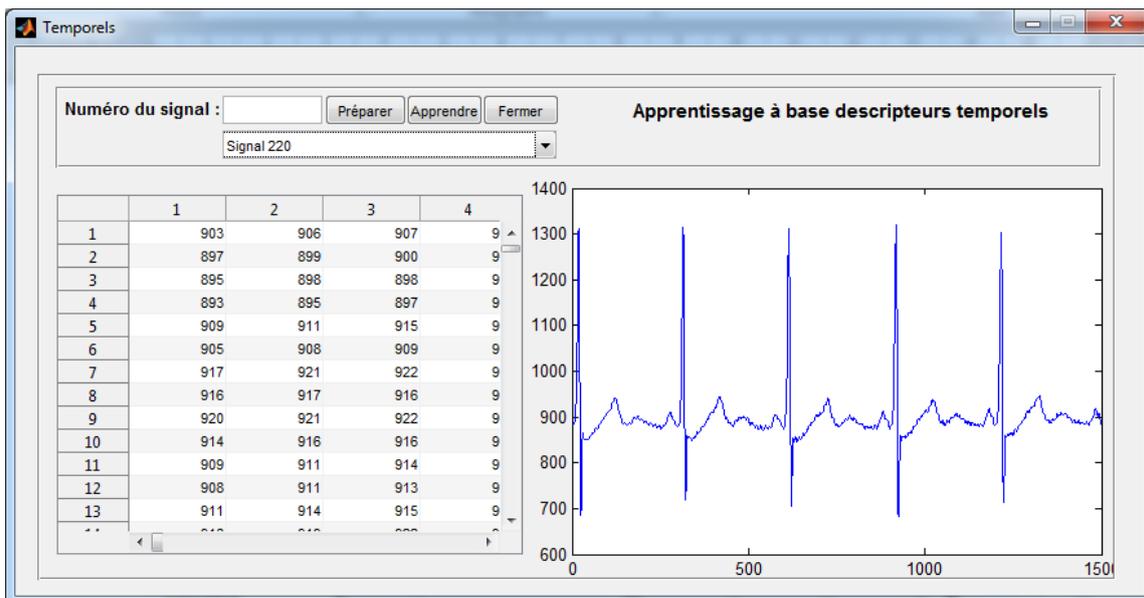


Figure IV. 4: Interface principale du classifieur temporel

Lorsque on prend ce signal ; on va préciser leurs paramètres tels que le nom d'apprentissage et le nombre de couches cachées, le nombre de couches de sortie, la fonction de transfert et la fonction d'activation, et l'erreur.

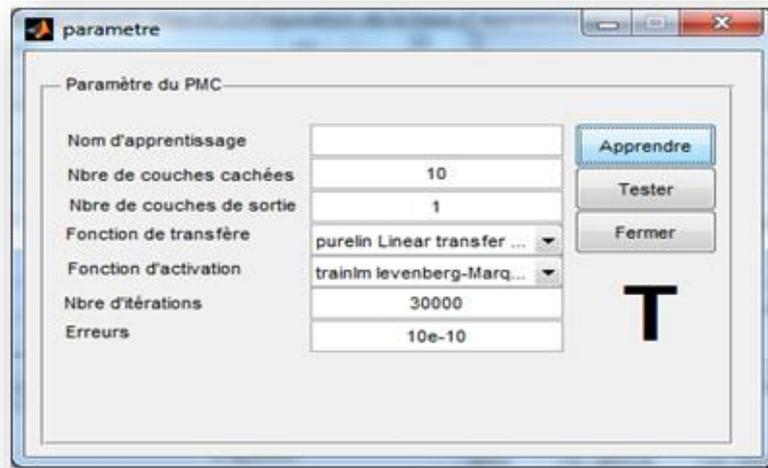


Figure IV. 5: Paramétrage du réseau de neurone

La figure IV.6 va nous montrer le schéma de notre réseau, son algorithme d'apprentissage et la progression du signal ainsi que le lien d'accès aux différents graphes :

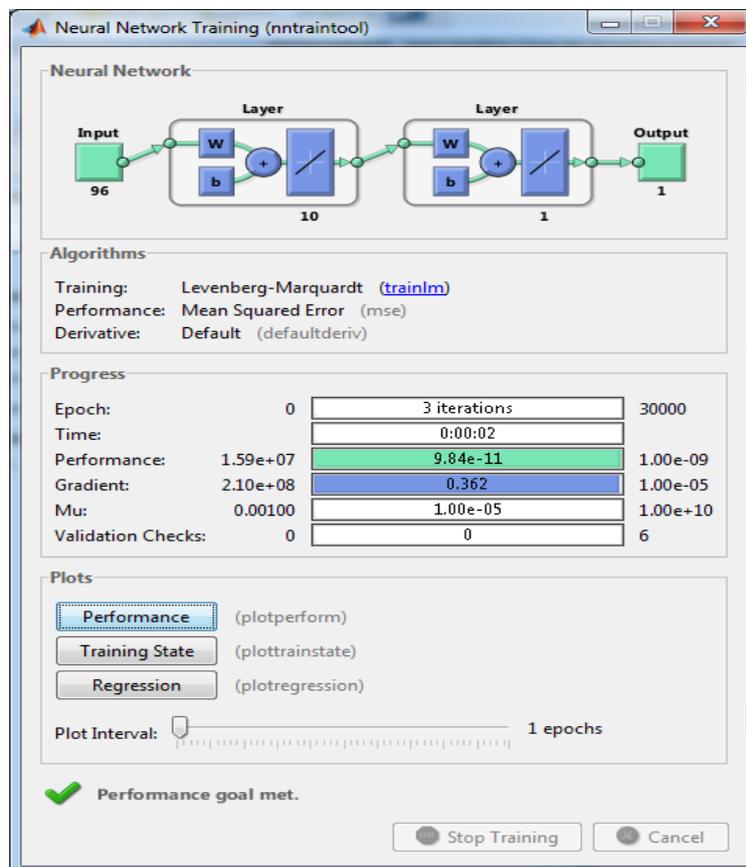


Figure IV. 6: Interface d'apprentissage du réseau de neurone

Dans la figure IV.7, on remarque une convergence de la performance d'apprentissage du réseau pour ce signal vers l'erreur précisée dans le paramétrage en fonction du nombre d'itérations fixées.

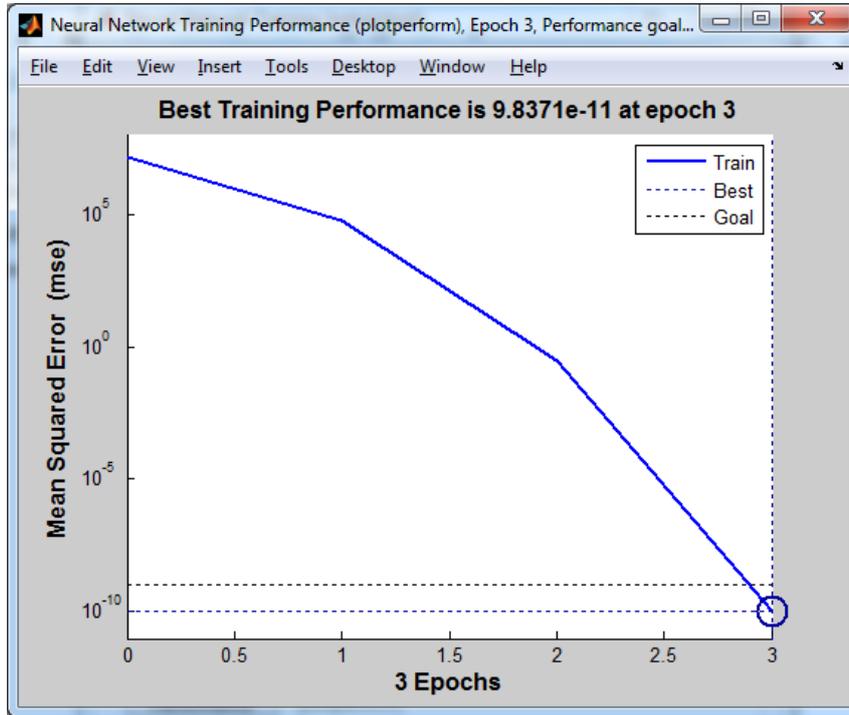


Figure IV. 7: Graphe de performance

La figure présente les graphes du gradient et la moyenne d'erreurs et validation du test par rapport au nombre d'itérations.

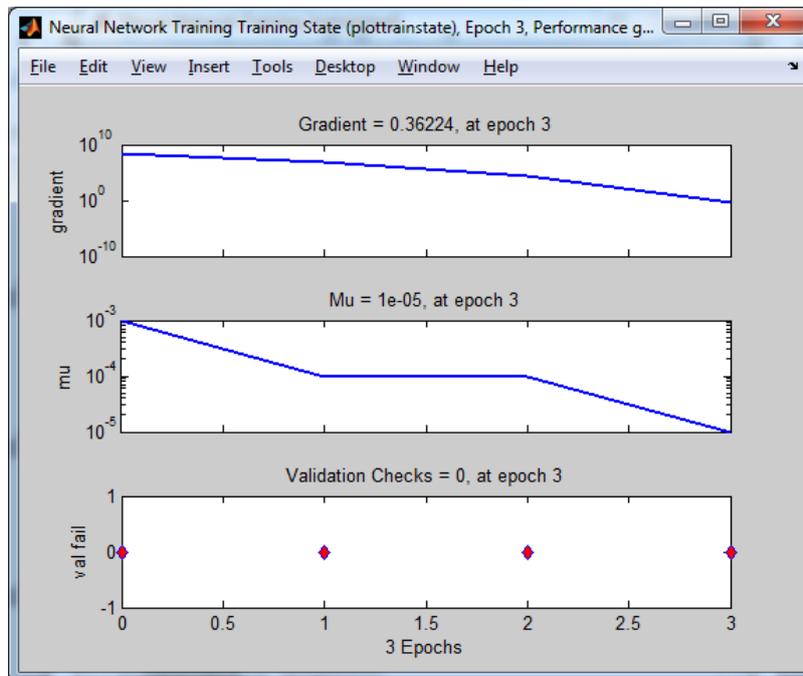


Figure IV. 8: Les graphes des gradient et mu et validation de test.

Après la validation de test . la figure IV.9 va nous montre les résultats du test.

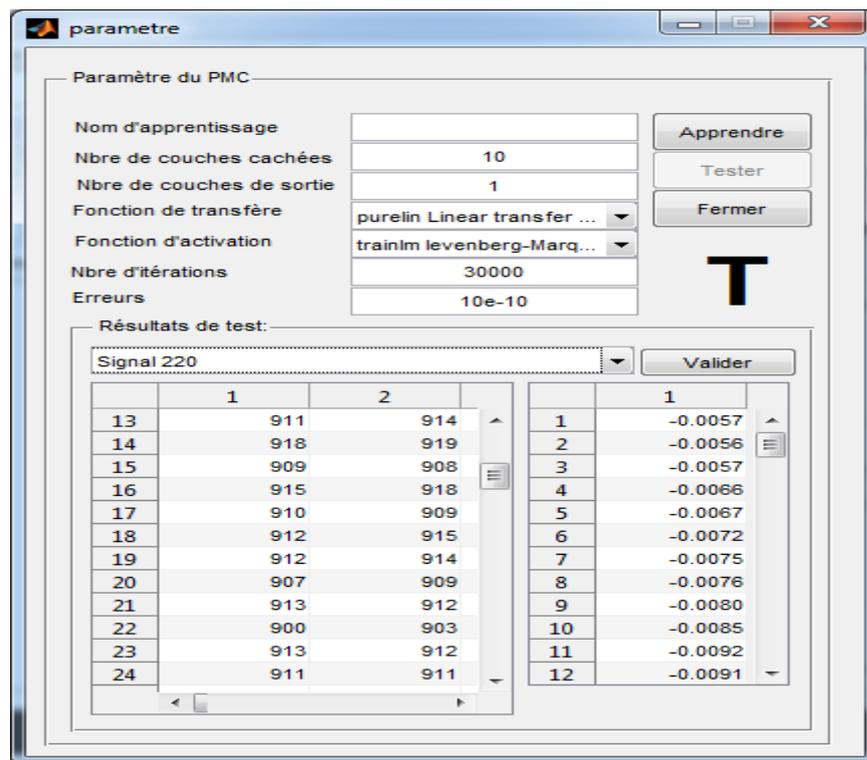


Figure IV. 9: Validation de test.

IV.4. Conclusion

Les signaux physiologiques constituent une source d'information qui prend une part de plus en plus importante dans la démarche diagnostique en médecine et qui devient de plus en plus facilement accessible du fait du développement des systèmes d'acquisition de signaux. Dans cette thèse nous nous sommes intéressés à un signal physiologique particulier, le signal électrocardiogramme (ECG).

Le signal ECG à un potentiel diagnostique important et son utilisation fait partie intégrante de l'examen clinique d'un patient. Son contenu informationnel est reconnu et exploité essentiellement dans le cadre de l'auscultation cardiaque.

Dans cette thèse, nous nous sommes intéressés à la reconnaissance des battements cardiaques normaux, d'une autre façon de classifier les battements cardiaques normaux et battement anormaux, pour pouvoir donner un bon diagnostic et dans les meilleurs délais, afin de protéger la santé des humains. Pour cela, nous avons proposé une méthode pour la classification (classification à base de descripteurs temporels extraits depuis le signal ECG).

CONCLUSION GENERALE

Conclusion générale

L'objectif de ce mémoire était l'application d'une approche neuronale en utilisant des outils de l'intelligence artificielle pour la classification des battements cardiaques, en l'occurrence, les réseaux de neurones artificiels qui ont prouvé leurs performances dans ce domaine (la classification des arythmies cardiaques).

Nous nous sommes intéressés à la classification des battements cardiaques, donc nous avons proposé de créer un modèle de classification, dont ce dernier reçoit les valeurs descriptives temporelles de l'ECG que nous désirons le classifier, à savoir les distances RRp, RRs, le rapport entre ces deux derniers, PP, la durée du complexe QRS, ce classifieur va faire un apprentissage supervisé en utilisant un réseau de neurones multi-Couches, qui donne une seule sortie (classe de classification, soit « N » pour normal, soit « A », pour autres).

Cette approche nous a permis de profiter des avantages des réseaux de neurones artificiels pour faire cette classification qui va servir sans aucun doute de sauver la vie des autres.

Cette petite recherche sera, dans le proche futur, améliorée afin de donner des solutions efficaces, opérationnelles et moins coûteuses permettant d'identifier les patients à risque de mort subite rythmique et de prendre des mesures préventives selon le cas rencontré et donc des millions de décès pourraient donc être évités.

Perspectives

Nous avons la classification des deux classes d'ECG (Normal/Anormal).

Donc nous proposons comme futur travail d'enrichir la base d'ECG à classer c'est-à-dire non seulement deux classes, nous irons vers le traitement des classes (Extrasystole Ventriculaire(ESV), Bradycardie, Tachycardie...).

RÉFÉRENCES
BIBLIOGRAPHIQUES

References bibliographies

- [1] W. Einthoven, "The galvanometric registration of the human electrocardiogram, likewise a review of the use of capillary-electrometer in physiology," *Cardiac classics: Willius FW-CV* Mosby., St Louis, 1941.
- [2] W. Zareba, P. Maison-Blanche and H. L. Locat, , *Noninvasive electrocardiology in clinical practice*. New-York: Futura Publishing Compagny, Inc., 2001.
- [3] Hernández, A, *Fusion de signaux et de modèle pour la caractérisation d'arythmies cardiaques*. Phd, Université de Rennes 1, Rennes, France., 2000.
- [4] Blondeau, M. et Hiltgen, M, "Electrocardiographie clinique. Masson," 1980.
- [5] Fromont Elisa, "thèse de doctorat « Apprentissage multisource par programmation logique inductive : application _a la caract_erisation d'arythmies cardiaques," 2005.
- [6] Stein, E, *Rapid analysis of arrhythmias*. Lippincott Williams ans Wilkins. 1999.
- [7] SZOLOVITS P., *Arti_cial Intelligence in Medicine*, Westview Press,. Inc., Boulder, Colorado (<http://medg.lcs.mit.edu/ftp/psz/AIM82/>), 1982.
- [8] STEIMANN F., , « *Fuzzy set theory in medicine* », *Arti_cial Intelligence in Medicine*, vol. , vol. 11, no 1, vols. 1997.
- [9] Belaid, A, Belaid, Y, « *Reconnaissance des formes. Méthodes et application* », InterEditions,. Paris, 1992.
- [10] CHEIKH Mohamed Amine, "Analyse du signal ECG par les réseaux neurones et la logique floue : Application à la reconnaissance des battements ventriculaires prématurés," thèse de doctorat, 2005.
- [11] Leschi C, "Aspects fondamentaux et évolutifs de la reconnaissance de formes. Approche pluridisciplinaire, notion d'approximation," », Thèse de doctorat en informatique et automatique appliquée, INSA de Lyon, 1991.
- [12] Milgram M, *Reconnaissance des formes : Méthodes numériques et connexionniste*. Armand Collin, Paris, 1993.
- [13] Guyon I, et Al, *What size test set gives good error rate estimates?*. 1998.

- [14] Ali MESSAOUD, Mohamed BEN MESSAOUD, Abdennaceur KACHOURI et Faïçal SELLAMI, « *Classification des arythmies cardiaques par logique floue à partir de signaux ECG* », Laboratoire d'Electronique et de Technologies de l'Information, : LETI,.
- [15] CHU Thanh-Quang, *Rapport de stage « Application des méthodes d'intelligence artificielle à l'analyse des données concernant les canalisations de gaz »*. Institut de la francophonie pour l'informatique, 2006.
- [16] V. Vapnik, "*Statistical Learning Theory*," New York, 1998.
- [17] Bouron T, *Structure de communication et d'organisation pour la coopération dans un univers multi-agents*,. Université Paris VI: (Thèse 3ème cycle, Informatique), 1992.
- [18] E. Marey, "*Des variations électriques des muscles et du cœur en particulier, étudiées au moyen de l'électromètre de M. Lippmann*", vol. ,vol. 82. CR Acad Sci (Paris), 1876.
- [19] Thirumalaiah, K., et Deo, . *River stage forecasting using artificial neural networks. ASCE Journal of Hydrologic Engineering 3(1) :. 1997.*
- [20] S. J. Rosenschein and G. Zlotkin, . *Rules of Encounter: Designing Conventions for Automated Negotiation Among Computer*. Boston, MA: . MIT Press, 1992.
- [21] T.J. Haykin, , *Neural network, a comprehensive foundation*,. 1994.
- [22] Hinton, G.E, *How neural networks learn from experience*. Scientific American, 1992.
- [23] Kohonen, T., *Self-organization and associative memory*,. 2e éd. Berlin.: Springer-Verlag, 1987.
- [24] Grossberg, S., *Birth of a learning law. International Neural Network Society Newsletter*,. 1–4., 1998.
- [25] BIS, *Neural networks for pattern recognition*. Oxford University Press, 1995.
- [26] Nash J.C, *Compact Numerical Methods for Computers: linear algebra and function minimisation*. Ed, Adam Hilger, 2000.
- [27] Hornik K, *Approximation capabilities of multilayer feedforward networks, Neural, Networks*, vol.4 vols. 1991.

ANNEXES

Annexe I : Base de données MIT-BIH

Depuis 1975, les laboratoires de l'hôpital de Beth Israël à Boston et MIT ont réalisé une base de données MIT-BIH qui a été commencée à être distribuée en 1980. Cette base de données contient 48 enregistrements extraits d'une demi-heure des enregistrements ambulatoires à deux voies d'ECG, obtenus à partir de 47 sujets étudiés par le laboratoire d'arythmie de BIH entre 1975 et 1979. 23 enregistrements ont été choisis au hasard d'un ensemble de 4000 enregistrements ambulatoires de 24 heures d'ECG rassemblés d'une population mélangée des patients hospitalisés (60%) et des patients non hospitalisés (40%) à l'hôpital de Beth Israël à Boston (les séries '100'), les 25 enregistrements restants ont été choisis parmi les mêmes enregistrements mais qui en considération des arythmies rarement observés qui ont une signification clinique (les séries '200').

Les enregistrements ont été échantillonnés à une fréquence $f_e = 360\text{Hz}$ avec une résolution de 11 bits sur une gamme de 10 mV. Deux cardiologues ou plus ont indépendamment annoté chaque enregistrement, environ 110,000 annotations ont été incluses avec la base de données. Le tableau B.1 présente les 48 enregistrements de la base de données MIT-BIH.

N° d'enregistrement Arythmies	N	A	V	F	P	Q
100	2239	33	1	-	-	-
101	1860	3	-	-	-	2
102	99	-	4	-	2028	-
103	2082	2	-	-	-	-
104	163	-	2	-	1380	18
105	2526	-	41	-	-	5
106	1507	-	520	-	-	-
107	-	-	59	-	2078	-
108	1739	4	17	2	-	-
109	-	-	38	2	-	-

Annexes

111	-	-	1	-	-	-
112	2537	2	-	-	-	-
113	1789	-	-	-	-	-
114	1820	10	43	4	-	-
115	1953	-	-	-	-	-
116	2302	1	109	-	-	-
117	1534	1	-	-	-	-
118	-	96	16	-	-	-
119	1543	-	444	-	-	-
121	1861	1	1	-	-	-
122	2476	-	-	-	-	-
123	1515	-	3	-	-	-
124	-	2	47	5	-	-
200	1743	30	826	2	-	-
201	1625	30	198	2	-	-
202	2061	36	19	1	-	-
203	2529	-	444	1	-	4
205	2571	3	71	11	-	-
207	-	107	105	-	-	-
208	1586	-	992	373	-	2
209	2621	383	1	-	-	-
210	2423	-	194	10	-	-
212	923	-	-	-	-	-

213	2641	25	220	362	-	-
214	-	-	256	1	-	2
215	3195	3	164	1	-	-
217	244	-	162	-	1542	-
219	2082	7	64	1	-	-
220	1954	94	-	-	-	-
221	2031	-	396	-	-	-
222	2062	208	-	-	-	-
223	2029	72	473	14	-	-
228	1688	3	362	-	-	-
230	2255	-	1	-	-	-
231	314	1	2	-	-	-
232	-	1382	-	-	-	-
233	2230	7	831	11	-	-
234	2700	-	3	-	-	-

Tab B.1 : Les 48 enregistrements de la base de données MIT-BIH

N : Battement normal (*Normal beat*).

A : Battement auriculaire prématuré (*Atrial premature beat*).

V : Contraction ventriculaire prématurée (*Premature ventricular contraction*).

F : Mélange entre les battements ventriculaires et les battements normaux (*Fusion of ventricular and normal beat*).

P : Pause.

Q : Battement non classifiable (*Unclassifiable beat*).

Résumé :

Les maladies cardiovasculaires constituent un problème majeur de santé public, elles sont les premières causes de mortalités dans le monde.

En raison de l'ampleur du problème, divers travaux ont été mis en place afin de réduire le risque, préconisant notamment l'éducation, la prévention, et le suivi des patients à risque. Dans ce mémoire, nous avons appliqué une approche neuronale pour la classification temporelle du signal ECG où nous avons proposé de mettre au point un système de classification des battements cardiaques. Ce dernier est basé principalement sur l'utilisation des réseaux de neurones, dont, le classifieur neuronal reçoit en entrée 96 paramètres temporels caractérisant un battement ECG. Notre système a pour objectif de faire la classification neuronale en se basant sur les données temporelles d'ECG selon deux classes, « normale » et « anormale », qui a été validé sur des signaux de la base de données MIT_BIH. Les résultats qualitatifs et quantitatifs obtenus démontrent l'efficacité de cette approche.

Mots clés : ECG, les maladies cardio-vasculaires, Arythmies cardiaques, classification, réseaux de neurones, apprentissage artificielle, descripteur temporel, Apprentissage supervisé.

المخلص:

الأمراض القلبية تشكل خطرا كبيرا على الصحة العمومية وهي أولى اسباب الوفيات في العالم.

وللحد من هذه المشكلة وأخطارها بما في ذلك التعليم والوقاية ورصد المرضى المعرضين لهذا الخطر. في هذه المذكرة قمنا بتطبيق نظام زمني لمعاينة نبضات القلب ويعتمد هذا الأخير على الشبكة العصبية حيث يتلقى النظام 96 ميزة زمنية لتمييز نبضة قلبية واحدة وتصنيفها الى فئتين عادية وغير عادية معتمدا على قاعدة بيانات MIT_BIH وتبيين النتائج الكمية والنوعية على فعالية هذا النظام.

الكلمات الأساسية: مخطط نبضات القلب، أمراض الأوعية الدموية، عدم انتظام نبضات القلب، التصنيف، الشبكة العصبية، التعلم الاصطناعي، الواصف الزمني، التعلم الاشرافي.