



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mohamed Khider – BISKRA
Faculté des Sciences Exactes, des Sciences de la Nature et de la Vie
Département d'informatique

N° d'ordre : SIOD 2/M2/2019

Mémoire

Présenté pour obtenir le diplôme de master académique en

Informatique

Parcours : **Système d'information optimisation décision (SIOD)**

La détection précoce des anomalies du rythme cardiaque basée sur le signal ECG par l'arbre de décision automatisée

Par :

KERDOUDI Saouab

Soutenu le 7 juillet 2019, devant le jury composé de :

ABABSA Tarek

MCB

Président

BITAM Salim

Professeur

Rapporteur

ABDELLI Belkacem

MCB

Examineur

Remerciements

*Je tiens tout d'abord, à remercier **ALLAH** le tout puissant qui m'a donné la force, la volonté, et la patience d'accomplir ce modeste travail.*

*Ensuite, je tiens à exprimer mes vifs remerciements à monsieur le professeur **BITAM Salim**, mon encadreur de recherche, pour ses efforts, ses conseils, et sa disponibilité*

Mes vifs remerciements vont également aux membres du jury qui ont accepté d'examiner et de juger mon travail de fin d'étude et de l'enrichir par leurs propositions.

*Ma gratitude va également à **Dr. OKBI Ridha** pour ses propositions qui ont enrichi ce modeste travail, pour son aide, ses conseils et ses encouragements.*

Je tiens à remercier aussi tous les enseignants du département de l'informatique.

*Je tiens à remercier aussi **DR. AGLI Nasir** qui m'a beaucoup aidé par la documentation nécessaire pour réaliser ce travail.*

Finalement, un grand Merci à tous ceux qui m'ont aidée de près ou de loin à la réalisation de ce travail de recherche.

MERCI, à tous et à toutes

Dédicace

Ce mémoire n'aurait jamais pu voir le jour sans le soutien actif des membres de ma famille que je tien tant à remercier :

*À la flamme de ma vie, mon exemple éternel, mon soutien moral et la source de ma joie ainsi que de mon bonheur ; celle qui s'est toujours sacrifiée pour me voir réussir. **À ma mère .***

*À celui à qui je dois tous le respect et l'amour, pour son soutien, sa tendresse, sa patience et surtout son aide dans les moments les plus difficiles de ma vie. **À mon père.***

***À mes sœurs**, qu'elles occupent une place unique dans ma vie, j'offre cette dédicace pour leurs confirmer l'importance, l'amour, l'amitié qu'elles ont à mes yeux.*

*Je dédier ce modeste travail à mes très chers frères, **A mes beaux frères, A mes belles-sœurs** chacun à son nom.*

Pour le soutien qu'ils m'ont apporté et pour leurs contributions à ma formation.

A mes anges, mes nièces et mes neveux.

J'exprime également à traves ce travail mes vifs sentiments de fraternité et d'amitié...

A mes amies IBTISSEM, NESSEMA, AMINA, SAMAH...

A ceux que j'aime

A ceux qui m'aime

Je dédie ce travail.

Résumé

Les maladies du rythme cardiaque (les arythmies) sont l'une des principales causes de mortalité dans le monde. Par conséquent, la détection et la classification des arythmies cardiaques sont essentielles au diagnostic des patients souffrant d'anomalies cardiaques. Dans la littérature, il existe de nombreuses approches de classification efficaces, telles que l'arbre de décision, conçue pour l'analyse du signal Électrocardiogramme (ECG). Cependant, le taux de classification atteint reste modéré et que la sauvegarde permanente n'est pas encore offerte. De plus, il n'existe pas des dispositifs de suivi du rythme cardiaque à la portée des citoyens en termes de prix.

Pour faire face à ces problèmes, nous proposons d'améliorer de la méthode d'arbre de décision en suggérant un nouvel algorithme de recherche tabou (RT) pour trouver le nombre optimal d'arbres où la précision de la classification du signal ECG et la taille d'arbre sont traitées comme des critères de la fonction objectif. De plus, nous proposons une conception de ce système pour qu'il soit capable à enregistrer le signal ECG de manière permanente. Le système proposé comporte quatre étapes principales, à savoir la collecte de données du signal ECG, le prétraitement et le débruitage de ces données, l'extraction de caractéristiques et la classification de ce signal à l'aide de l'approche d'arbre de décision améliorée. Pour valider cette proposition, un ensemble d'expériences a été mené sur les bases de données provenant du site web Physionet et d'une plate-forme réaliste de santé connectée de marque Cooking hacks. Les résultats obtenus ont montré que l'arbre de décision améliorée peut atteindre 97,8% de la précision de la classification en fonction d'un nombre optimal d'arbres qui est 6.

Table des matières

Remerciements	I
Dédicace	II
Résumé	III
Table des matières	VI
Table des figures	IX
Table des tableaux	X
Introduction générale	1
1 Le rythme cardiaque et l'électrocardiographie	4
1.1 Introduction	4
1.2 Le système cardiovasculaire et le rythme cardiaque	4
1.2.1 Le système cardiovasculaire	4
1.2.2 L'électrocardiogramme ECG	5
1.2.3 Le cycle cardiaque	5
1.2.4 Les dérivations cardiaques	6
1.2.5 Le rythme cardiaque et la fréquence	9
1.2.6 Rythme sinusal régulier	10
1.2.7 Le battement cardiaque standard et ses caractéristiques	10
1.2.8 Symptômes d'arythmies	10
1.2.9 L'examen diagnostique pour l'arythmie	11
1.3 L'interprétation de l'ECG	12
1.4 Conclusion	17

2	L'étude de l'arythmie : classification et état de l'art	19
2.1	Introduction	19
2.2	Les anomalies cardiaques : étude de l'arythmie	20
2.2.1	Les anomalies sinusales	22
2.2.2	Les anomalies auriculaires	25
2.2.3	Les anomalies du nœud A.V	27
2.2.4	Les anomalies ventriculaires	29
2.3	Travaux généraux sur les signaux ECG	32
2.4	Revue de littérature	34
2.4.1	Classification	34
2.4.2	Les applications connectées pour détecter les arythmies cardiaques : y-a-il un dispositif à la portée des citoyens ?	36
2.4.3	Travaux similaires et enregistrement permanent des signaux ECG	36
2.5	Conclusion	37
3	Conception du système de détection de l'arythmie : par l'arbre de déci- sion améliorée	39
3.1	Introduction	39
3.2	Conception globale du système	39
3.2.1	Collecte de données de santé	41
3.2.2	Sauvegarder les données	41
3.2.3	Prétraitement et l'extraction des caractéristiques	41
3.2.4	La sélection des caractéristiques	42
3.2.5	Entraînement avec le nombre optimal des nœuds dans l'arbre de décision	42
3.3	Conception détaillée du système	43
3.3.1	Collecte de données de l'ECG	45
3.3.2	Sauvegarde des données	46
3.3.3	Prétraitement et extraction des caractéristiques	48
3.3.4	Sélection des caractéristiques	49

3.3.5	Entraînement avec le nombre optimal des d'arbres dans l'arbre de décision	49
3.3.6	L'évaluation des solutions obtenues par RT	53
3.3.7	Visualisation, Interprétation, Utilisation	54
3.4	Conclusion	56
4	Etude expérimentale et résultats	58
4.1	Introduction	58
4.2	Outils et langages de développement	58
4.2.1	Java	59
4.2.2	Eclipse	59
4.2.3	Le langage R	59
4.2.4	RStudio	60
4.2.5	Xampp	60
4.3	Le système développé	61
4.3.1	L'ensemble des données utilisées	61
4.3.2	Prétraitement et extraction des caractéristiques	76
4.3.3	Sélection et calcul des caractéristiques	78
4.3.4	Entraînement avec l'arbre de décision optimisé	80
4.3.5	Visualisation, Interprétation, Utilisation	82
4.4	Résultats obtenus et discussion	86
4.5	Conclusion	88
	Conclusion générale	89

Table des figures

1.1	L'électrocardiogramme ECG [3].	5
1.2	Cycle cardiaque : mouvements des parois et des valves (flèches) [4].	5
1.3	Système de conduction cardiaque [5].	6
1.4	Six dérivations des membres [6].	7
1.5	Vue schématique de l'électrode précordiale [8].	8
1.6	Vue schématique de l'électrode postérieure [10].	9
1.7	Rythme sinusal régulier [1].	10
1.8	Symptômes d'Arythmies [13].	11
1.9	Tableau récapitulatif [7].	16
2.1	Schéma récapitule les maladies qu'un ECG peut détecter.	21
2.2	Bradycardie sinusale : le rythme est de l'ordre de 40 BPM [17].	22
2.3	Tachycardie sinusale : le rythme est de 120 BPM [17].	23
2.4	Rythme supra-ventriculaire multifocal [17].	25
2.5	Tachycardie supra-ventriculaire [18].	25
2.6	La tachycardie auriculaire [18].	26
2.7	La fibrillation auriculaire [17].	26
2.8	Le flutter auriculaire [17].	27
2.9	B.A.V de 1er degré [17].	27
2.10	B.A.V de 2ème degré Mobitz typeI [17].	28
2.11	B.A.V de 2ème degré Mobitz typeII [17].	28
2.12	B.A.V de 3ème degré [17].	28
2.13	L'extrasystole ventriculaire (ESV).	30

2.14	Tachycardie ventriculaire (TV). On observe sur l'enregistrement une succession d'extrasystoles ventriculaires à une fréquence de 150 bpm ; après la phase de TV, le cœur reprend ici spontanément un rythme sinusal normal [1].	31
2.15	Flutter ventriculaire [18].	31
2.16	La fibrillation ventriculaire [1].	32
2.17	Les trois tâches de data mining en relation avec les signes vitaux de santé [19].	33
2.18	Méthodes d'exploration de données utilisées dans la littérature pour analyser le signal ECG [19].	33
2.19	Les applications connectées pour détecter les arythmies cardiaques ECG [25].	36
3.1	L'architecture générale du système.	40
3.2	Prétraitement et extraction des caractéristiques.	42
3.3	L'architecture détaillée du système.	44
3.4	Le modèle multidimensionnel proposé du système de sauvegarde de données ECG.	46
3.5	Signal ECG avec les bruits [38].	48
3.6	Signal ECG sans les bruits [38].	48
3.7	L'algorithme Recherche Tabou pour l'optimisation de l'arbre de décision.	53
3.8	Utilisation du modèle.	55
3.9	Le système proposé en cas d'urgence suite à une détection d'anomalie cardiaque.	56
4.1	R logo.	59
4.2	a.Raspberry Pi3 B, b.Pont de connexion entre la e-HP et le raspberry Pi, c. la e-HP.	61
4.3	Capteurs ECG connectés à la plate-forme.	62
4.4	Regroupement de système.	62
4.5	La plate-forme et les électrodes ECG sont placé sur le patient.	63

4.6	Le code illustratif d’envoi des données depuis le RaspBerry vers le serveur.	64
4.7	Création d’un compte pour le PC1.	66
4.8	La connexion à distance à la base de données.	66
4.9	La connexion à notre base de données.	67
4.10	Réception des données du client.	68
4.11	Enregistrement des données acquises depuis la plate-forme.	69
4.12	Les types et les valeurs des champs de la table de cookie hack.	70
4.13	La fonction de numérisation	71
4.14	Exemple de fichier (.dat).	73
4.15	Exemple de fichier (.hea).	73
4.16	Exemple de fichier (.atr).	74
4.17	Prétraitement et extraction des caractéristiques.	77
4.18	Le fichier des caractéristiques extraites (.txt).	77
4.19	Le processus de correspondance [52].	79
4.20	La matrice de données des caractéristiques sélectionnées et calculées.	79
4.21	Une fonction de sélection d’un nombre d’arbre au hasard.	80
4.22	Interface de l’application médecin.	83
4.23	Interface pour l’identification de médecin.	83
4.24	Interface pour ajouter des commentaires et des rendez-vous.	84
4.25	Interface de profil de médecin.	84
4.26	Interface ajouter un patient.	85
4.27	Interface consulter les signaux avec des dérivations différentes.	85
4.28	Interface les signaux avec des dérivations différentes d’un patient.	86
4.29	les valeurs de taux de précision en fonction de nombre des arbres avec algorithme d’optimisation Recherche Tabou.	87
4.30	les valeurs de taux de précision en fonction de nombre des arbres sans algorithme d’optimisation RT.	87

Liste des tableaux

1.1	Valeurs habituelles des différents paramètres caractérisant un battement cardiaque [1].	10
1.2	Les valeurs normales et des cas anormaux avec ses diagnostics probables des différents paramètres caractérisant un battement cardiaque.	13
4.1	Valeurs physiques en représentation numérique	71
4.2	Les annotations des battements.	75
4.3	Les commandes WFDB utilisés [28].	76
4.4	Les caractéristiques détectées et extraites [52].	78

Introduction générale

D'après l'organisation mondiale de la Santé (OMS), les maladies cardiovasculaires (MCV) coûtent la vie à 17,9 millions de personnes chaque année, soit 31% de tous les décès dans le monde. Une de ces maladies cardiaques est connue sous l'appellation de l'arythmie ou aussi les anomalies du rythme cardiaque.

Parmi les outils les plus utilisés pour détecter de telles anomalies, les médecins spécialistes en Cardiologie utilisent le signal Électrocardiogramme (abrégé ECG) qui est un outil de diagnostic utile et efficace pour constater les diverses maladies cardiovasculaires telles que l'arythmie cardiaque. L'ECG enregistre l'activité électrique du cœur et ces signaux peuvent refléter l'activité anormale du cœur. Cependant, il est difficile d'interpréter visuellement les signaux ECG en raison de sa faible amplitude et de sa faible durée. Au cours des dernières années, de nombreuses entreprises ont manifesté leur intérêt pour la détection des maladies cardiaques telles que l'arythmie cardiaque via des appareils intelligents à savoir l'Android Wear, Apple Watch, Kardia, et CRONOVO... Des millions de personnes portent à tout moment des capteurs de fréquence cardiaque aux poignets. Ces appareils portables génèrent des ordres de grandeur plus nombreux que les ECG (électrocardiogrammes). Cependant, ces appareils ne permettent pas un apprentissage automatique pour détecter les personnes souffrant d'arythmie grave mettant la vie des humains en danger en utilisant ces données de fréquence cardiaque. De plus, une simple acquisition de l'ECG est insuffisante car plusieurs anomalies liées au rythme cardiaque exigent un enregistrement permanent de ce signal, ce qui permettra aux médecins spécialistes de bien diagnostiquer l'état des patients qui subissent des attaques dans des périodes imprévisibles telles que la nuit, durant la conduite, ou autres. Par conséquent, le praticien spécialiste peut faire référence à de telles données pour bien savoir le type d'anomalie et donc pour prescrire le traitement ou le processus que le patient doit suivre. Il est à noter qu'une absence de l'historique oblige le médecin de prescrire un traitement maximal pour faire face à l'incertitude du diagnostic, ce qui peut fatiguer de façon sévère le patient par une prise de médicament à effet indésirable ou à effectuer des analyses quotidiennes épuisantes.

Pour éviter ces risques ou pour contribuer à la réduction du nombre de décès causés

par ce type d'anomalie de rythme cardiaque, ce projet de Master propose la conception et le développement d'un système d'aide à la détection précoce des anomalies de rythme cardiaque et d'un suivi permanence de ce rythme pour des patients ou même pour des personnes normales en se basant sur un enregistrement en continu de signal ECG dans une infrastructure de données massive (Big Data) et aussi par une analyse de ce signal à l'aide de l'approche de l'arbre de décision permettant une aide au diagnostic au profit du médecin spécialiste. Par ailleurs, nous proposons une optimisation de la méthode l'arbre de décision en intégrant une méthode d'optimisation avancées qui est la recherche Tabou afin d'améliorer le taux de précision lors de l'analyse et de la classification du signal acquis.

Afin de valider ces propositions, nous allons développer une plate-forme de santé connectée (Connected Health) qui utilise le dispositif d'acquisition de l'ECG Cooking hacks, et qui est équipé d'un nano-ordinateur de type Raspberry PI 3.

Ce mémoire est composé de quatre chapitres en plus de cette introduction et une conclusion générale, il est organisé de la manière suivante : le premier chapitre est consacré aux concepts du rythme cardiaque et de l'électrocardiographie afin que les lecteurs connaissent bien ce domaine.

Le deuxième chapitre expose une étude sur l'arythmie ainsi qu'un état de l'art en présentant quelques méthodes similaires proposées dans la littérature.

Dans le troisième chapitre, nous présenterons nos propositions à savoir la conception générale et détaillée du notre système d'enregistrement, d'analyse et de classification du signal ECG. On présente aussi notre proposition pour améliorer la méthode : l'Arbre de décision par l'algorithme : la recherche Tabou (abrégé RT).

Dans le dernier chapitre, nous illustrerons l'étude expérimentation en expliquant les différentes étapes permettant la mise en œuvre de notre projet, les tests réalisés et les résultats obtenues.

Le mémoire se termine par une conclusion générale contenant les perspectives envisagées.

Chapitre 1 :

Le rythme cardiaque
et l'électrocardiographie

Chapitre 1

Le rythme cardiaque et l'électrocardiographie

1.1 Introduction

Nous présentons dans ce chapitre les notions de base en relation avec les maladies cardiovasculaires associées à une arythmie, qui dépendent du fonctionnement général du système cardiovasculaire. Nous présentons aussi les motivations d'une détection précoce et d'un suivi d'une maladie ou une anomalie de rythme cardiaque qui peut empêcher de graves conséquences telles que les morts subites ou les séquelles physiques ou psychologiques postérieures. Par la suite, nous allons concentrer en particulier sur un signal qui représente l'activité électrique du cœur qui s'appelle l'électrocardiogramme (ECG),

1.2 Le système cardiovasculaire et le rythme cardiaque

1.2.1 Le système cardiovasculaire

Le système cardiovasculaire assure la circulation du sang dans l'organisme et permet ainsi son alimentation en oxygène et en nutriments. Il est composé du cœur, sorte de double pompe, qui assure la circulation dans deux réseaux complémentaires : celui des artères et celui des veines [1].

1.2.2 L'électrocardiogramme ECG

Un ECG (électrocardiogramme), également connu sous le nom d'EKG, est un examen rapide ne prenant que quelques minutes, indolore et non invasive, dénué de tout danger. Pour enregistrer les impulsions électriques qui déclenchent la contraction cardiaque, il donne également des informations valables sur le cœur pendant les phases de repos et de récupération [2].

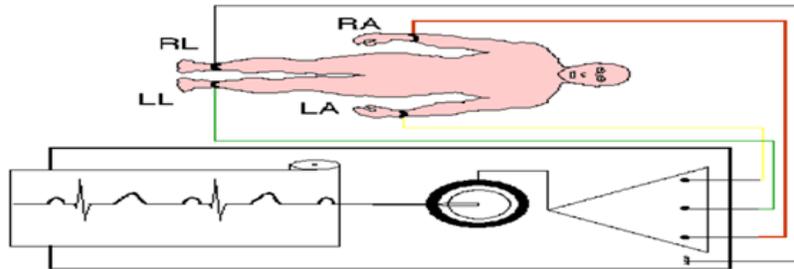


FIGURE 1.1 – L'électrocardiogramme ECG [3].

1.2.3 Le cycle cardiaque

L'activité électrique périodique engendre l'activité mécanique du cœur, dominée par la contraction puis la relaxation des ventricules. La succession d'une systole ventriculaire (figure 2a) et d'une diastole ventriculaire (figure 2b) forme un cycle cardiaque [4].

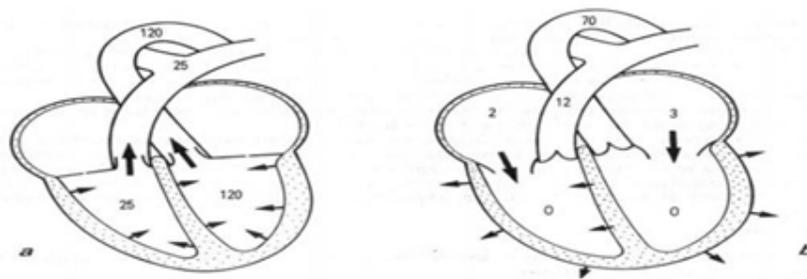


FIGURE 1.2 – Cycle cardiaque : mouvements des parois et des valves (flèches) [4].

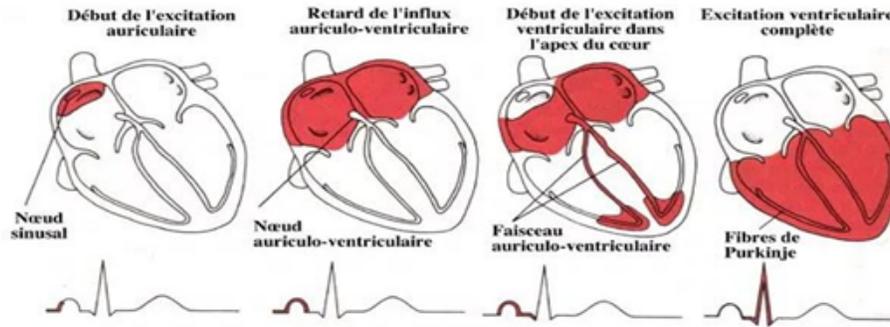


FIGURE 1.3 – Système de conduction cardiaque [5].

Le nœud SA est l'origine de l'impulsion électrique, ces impulsions se lancent dans les oreillettes sous forme d'une onde appelée onde P, ainsi l'onde P représente l'activité électrique de la contraction des deux oreillettes. L'impulsion atteint le **nœud AV**, où il se produit une pause de 1/10 seconde, permettant au sang de pénétrer dans les ventricules. Après la pause, le nœud AV est stimulé, déclenchant une impulsion électrique qui descend le long du **faisceau AV** vers les branches du faisceau. Le complexe QRS représente l'impulsion électrique qui se déplace du nœud AV vers **les fibres de Purkinje** et les cellules myocardiques (l'activité électrique de la contraction ventriculaire), les fibres de Purkinje transmettent l'impulsion électrique aux cellules myocardiques, entraînant la contraction simultanée des ventricules [2].

L'onde Q est la première déflexion vers le bas du complexe QRS. Elle est suivie d'une déflexion vers le haut de l'onde R, est suivie par une onde S dirigée vers le bas, après le complexe QRS une pause est produite, puis l'onde T est apparue, qui représente la repolarisation des ventricules qui les rend de nouveau stimulables [2].

Un cycle cardiaque est représenté par l'onde P, le complexe QRS et l'onde T, ce cycle se répète indéfiniment [2].

1.2.4 Les dérivations cardiaques

L'ECG est composé de 12 dérivations standards :

a **Six dérivations des membres (le plan frontal)**

Ils peuvent être divisés en deux parties, les dérivation bipolaires DI, DII et DIII et les dérivation monopolaires aVR, aVL et aVF. Pour obtenir les dérivation des membres, les électrodes sont placées sur les bras droit et gauche et sur la jambe gauche pour former un triangle (triangle d'Einthoven), et pour les dérivation monopolaires la dérivation AVR utilise le bras droit comme positif et toutes les autres électrodes des membres comme une terre (commune) négative. Les deux autres dérivation des membres, AVL et AVF, sont obtenues de façon analogue [2].

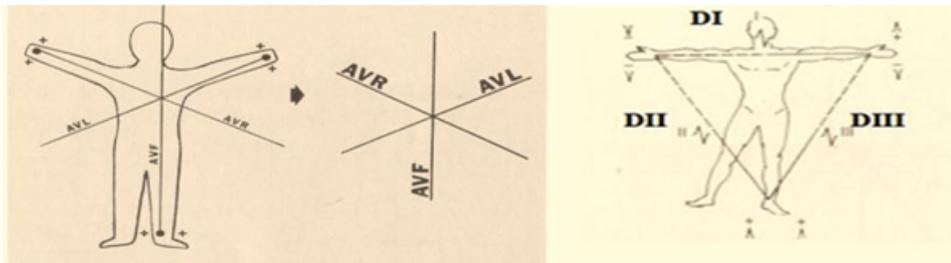


FIGURE 1.4 – Six dérivation des membres [6].

b Six dérivation des précordiales

Pour obtenir les six dérivation précordiales (thoraciques), une électrode positive est placée en six positions différentes tout autour de la poitrine [2].

- V1 : est placée sur le 4ème espace intercostal droit, au bord droit du sternum (ligne parasternale) [7].
- V2 : est placée sur le 4ème espace intercostal gauche, au bord gauche du sternum (ligne parasternale) [7].
- V3 : à mi-distance entre V2 et V4 [7].
- V4 : est placée sur le 5ème espace intercostal gauche, sur la ligne médioclaviculaire [7].
- V5 : est placée sur le 5ème espace intercostal gauche, sur la ligne axillaire antérieure [7].
- V6 : est placée sur le 5ème espace intercostal gauche, sur la ligne axillaire moyenne [7].

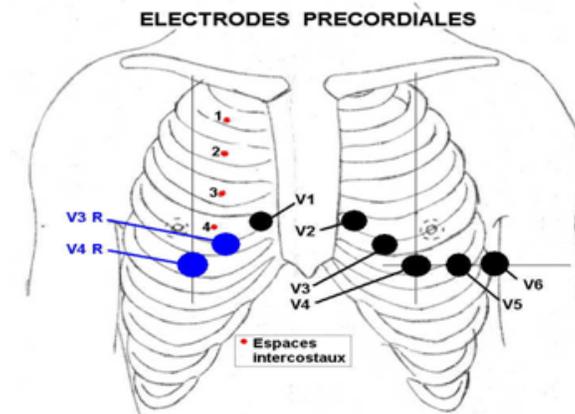


FIGURE 1.5 – Vue schématique de l'électrode précordiale [8].

c Autres dérivations [7]

Il est possible de réaliser d'autres dérivations selon les besoins diagnostiques. En effet, si un infarctus inférieur droit ou postérieur est suspecté, on réalise en plus des dérivations «standard», des dérivations «complémentaires» droites et postérieures. Il est à noter qu'un infarctus du myocarde parfois appelé crise cardiaque, correspond à la destruction d'une partie du muscle cardiaque appelé le myocarde. Il se produit lorsque par exemple, un caillot empêche le sang de circuler normalement dans l'artère coronaire, l'artère qui irrigue le cœur. Ce dernier est alors mal irrigué et le muscle cardiaque abîmé [9]. Les dérivations droites sont en miroir des dérivations V3 et V4.

- V3r : à mi-distance entre V1 et V4
- V4r : 5ème espace intercostal, ligne médio-claviculaire

Des dérivations postérieures sont quant à elles, sur le prolongement de la même horizontale que V4.

- V7 : même niveau que V4, sur la ligne axillaire postérieure.
- V8 : même niveau que V4, sur la pointe de l'omoplate.
- V9 : même niveau que V4, à mi-distance entre V8 et le rachis.

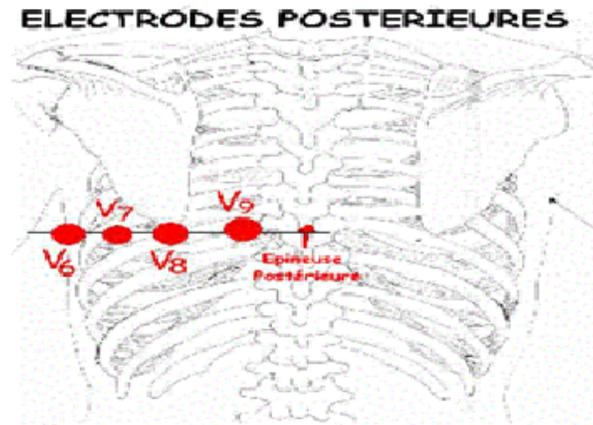


FIGURE 1.6 – Vue schématique de l'électrode postérieure [10].

1.2.5 Le rythme cardiaque et la fréquence

Pour **étudier le rythme**, il faut examiner une dérivation ayant un tracé avec au moins 12 complexes (DII souvent). Le rythme cardiaque normal est le rythme sinusal dans lequel le nœud sinusal initie la contraction cardiaque en délivrant des impulsions à la fréquence de 60 à 100 battements par minute. Les principales caractéristiques du rythme sinusal sont :

- Une fréquence cardiaque comprise entre 60 et 100 bpm.
- Chaque onde P est suivie d'un complexe QRS et inversement.
- Une onde P positive en DII et négative en aVR.

Une arythmie est un problème dû à un fonctionnement anormal de l'électricité cardiaque. L'arythmie peut être vue si le cœur bat trop vite (plus de 100 battements à la minute), ou à l'inverse, si il bat trop lentement (moins de 60 battements à la minute). Également, sans être ni trop rapide ni trop lent, le cœur peut avoir un rythme dit irrégulier. Ce sont tous des exemples d'arythmies.

Quand on parle de **la fréquence cardiaque**, on parle en fait de la fréquence ventriculaire. Elle se mesure donc en comptant le nombre de complexes QRS par minute [11].

1.2.6 Rythme sinusal régulier

Ce rythme est représenté ci-dessous, il s'agit d'un rythme cardiaque (Rythme sinusal régulier) de un sujet adulte en bonne santé.

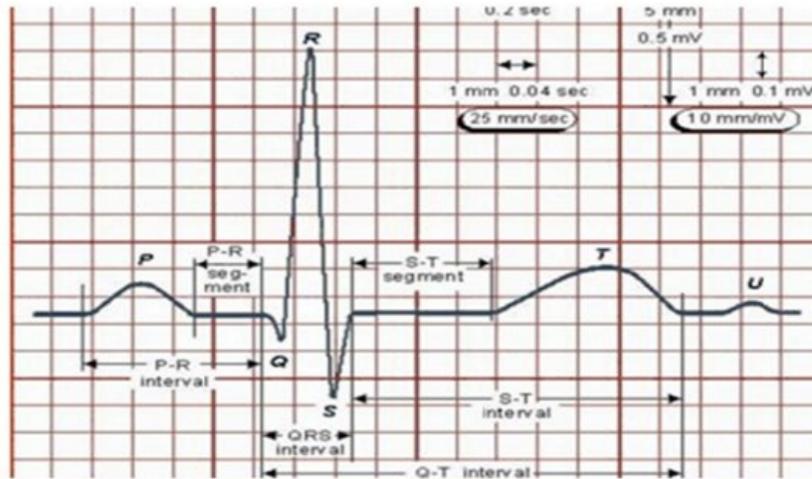


FIGURE 1.7 – Rythme sinusal régulier [1].

1.2.7 Le battement cardiaque standard et ses caractéristiques

Les valeurs des paramètres de la Figure 1.7 couramment constatées chez l'adulte en bonne santé sont présentées dans le tableau 1 [1].

	Onde P	Intervalle PQ	Complexe QRS	Intervalle ST	Intervalle QT	Onde T
Durée (s)	0.08-0.10	0.12-0.20	0.08	0.20	0.36	0.2
Amplitude (mV)	0.25	Isoélectrique:0	$Q < 0, R > 0, S < 0$	Isoélectrique: 0	-	$(T) > 0$

TABLE 1.1 – Valeurs habituelles des différents paramètres caractérisant un battement cardiaque [1].

1.2.8 Symptômes d'arythmies

Parmi les symptômes associés aux arythmies, mentionnons [12] :

- Les palpitations (battements de cœur rapides ou irréguliers).
- Les étourdissements ou vertiges.
- L'évanouissement ou quasi-évanouissement.

- L'essoufflement.
- La douleur ou l'inconfort au niveau de la poitrine.

Un grand nombre de ces symptômes risquent d'être confondus avec ceux d'autres maladies du cœur ou avec les effets de l'âge ou de l'inactivité physique.

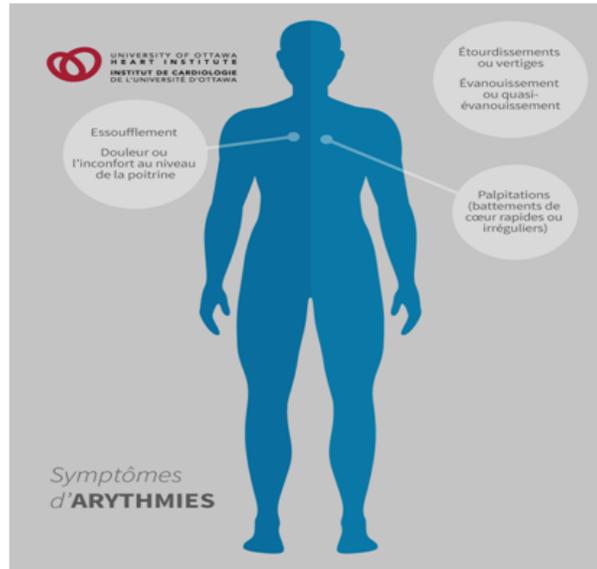


FIGURE 1.8 – Symptômes d'Arythmies [13].

1.2.9 L'examen diagnostique pour l'arythmie

Après une recherche dans des ouvrages médicaux et des conversations avec des médecins, nous avons confirmé qu'en cas de doute que la personne a une maladie cardiaque l'examen standard effectué dans ce cas est l'électrocardiographie abrégé ECG qui est accompagné d'une analyse de sang. Pour plus d'investigation, le médecin peut demander une échographie cardiaque.

Il y a aussi Échocardiogramme, Cathétérisme cardiaque, Tomodensitométrie (TDM) cardiaque, La tomodensitométrie / IRM (imagerie par résonance magnétique). Si on parle de l'arythmie, il peut être difficile de la diagnostiquer car l'apparition de l'arythmie peut survenir au hasard. Il existe toutefois plusieurs méthodes pour y diagnostiquer par :

- **L'électrocardiographie** : pour déceler les problèmes de rythme cardiaque [12].
- **Le moniteur Holter** : c'est un appareil d'électrocardiographie portatif utilisé pour surveiller en permanence la fréquence cardiaque pendant un jour ou deux [12].

- **Le moniteur d'événement** : c'est un appareil similaire au moniteur Holter, mais qu'on peut utiliser pendant plusieurs semaines ou plusieurs mois. Le moniteur d'événements est activé par le patient dès l'apparition des symptômes [12].
- **Les examens électro physiologiques** : pour suivre et stimuler les impulsions électriques dans le cœur [12].
- **L'épreuve d'effort** : permet de déterminer si l'épisode d'arythmie est déclenché quand le cœur fournit un effort plus important que d'habitude [12].

1.3 L'interprétation de l'ECG

On introduit dans cette partie les paramètres d'intérêt pour l'étude de l'ECG, où nous pouvons répondre à la question « Que faut-il observer sur un l'électrocardiogramme ? », et identifier les éléments de base sur lesquels le médecin se concentre lors de son analyse de l'ECG pour l'identification d'une maladie cardiaque, quelles sont les conséquences de tout changement dans un élément de l'ECG.

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 : **le rythme et la fréquence cardiaque** [1], **la durée** et **l'amplitude des ondes** et **l'axe du cœur**. Après une recherche dans des ouvrages médicaux [2] [14] [15], nous avons pu résumer les valeurs normales de ces paramètres caractéristiques couramment constatées chez l'adulte en bonne santé, et des cas anormaux avec ses diagnostics probables comme indiqué dans le tableau ci-dessous :

	Cas normal		Cas anormal		Diagnostic Probable
Rythme	Complexe QRS après chaque onde P				Le rythme est normal 'sinusal'
			L'intervalle R-R entre chaque cycle est constant		Le rythme est 'régulier'
Fréquence	Entre 60 et 100 battements par minute par journée, et entre 40 et 80 battements par minute pendant la nuit.		Inférieure à 60 par minute (trop lent)		Bradycardie
			Supérieure à 100 par minute (trop rapide)		Tachycardie
Les ondes	Cas normal		Cas anormal		Diagnostic Probable
	Durée ms	Amplitude mV	Durée Ms	Amplitude mV	
L'onde P	80	0.25	Plus de 120		hypertrophie auriculaire gauche
				Plus de 0,25 en II et 0,2 en III	hypertrophie auriculaire droite
L'intervalle PR	120<PR< 200	-	Plus de 200ms	-	bloc atrio-ventriculaire est du 1 ^{er} degré
Le complexe QRS	Moins de 100	Q<0 R>0 S<0	100<QRS<120		bloc de branche incomplet
			Plus de 120		bloc de branche est complet
L'onde T	200	T>0 dans la plupart des dérivations (exceptées aVR, V1, V2, aVF) et asymétrique.		T<0 (négative asymétrique)	*Bloc de branche (en V1 pour le BBD, V6 pour le BBG) *Hypertrophie VG (en V6).
L'intervalle QT	Moins de 440	-			
L'onde U		0.2 dans (V2-V3)		U<0 (négative)	*Ischémie myocardique. *Cardiopathies gauches. *Facteurs pharmacologique.

TABLE 1.2 – Les valeurs normales et des cas anormaux avec ses diagnostics probables des différents paramètres caractérisant un battement cardiaque.

Pour l'axe du cœur, l'axe se réfère à la direction de la dépolarisation qui diffuse à travers le cœur pour stimuler la contraction des fibres musculaires [1]. Les valeurs normales sont : [14]

- axe P : entre +30 et +70.
- axe T : entre 0 et +70.
- axe du complexe QRS : est le plus souvent déterminé, car le plus utile du point de

vue clinique. On utilise les dérivations du plan frontal pour déterminer l'axe du cœur (axe QRS).

Les valeurs normales sont : entre 0 et +90 le plus souvent autour de +60.

- Axe $>+90$: déviation axiale droite du QRS (en l'absence de BBG, correspondant à un héli bloc postérieur gauche ou une hypertrophie ventriculaire D).
- Axe < -30 : déviation axiale G du QRS (en l'absence de BBG, correspondant à un héli bloc antérieur gauche).

Le tableau suivant récapitule d'interprétation de l'ECG.

RYTHME	FREQUENCE	ONDES P	QRS	CONDUCTION	REMARQUES	DIAGNOSTIC PROBABLE	
REGULIER	LENT	Oui	Fin ou large	P & QRS sans relation	Avec bloc possible	BAV III *	
		Oui	n P= n QRS	PR>0,2s Espace PR constant		BAV I*	
		Oui	n P> n QRS QRS normal ou élargi	Espace PR constant	Avec possible bocage conduction (2/1, 3/1...)	BAV II MOBILIZ II (bloc de type 2/1,....)	
			Oui	< 0,12 s	PR : 0,1 s - 0,2s		RYTHME SINUSAL
		NORMAL	Oui	Large (>0,12s)		Aspect RSR' en V1-V2 Négative en V6	BLOC DE BRANCHE DROIT
			Oui	Large (>0,12s)		RR' Forme de M positive : V5-V6, AVL, DI Négative V1	BLOC DE BRANCHE GAUCHE
	RAPIDE	Non : ondes dents de scie = onde F (freq 300/min)	Fin		Possible bloc (2/1,3/1,..) Freq : 300, 150,100,.... Visible en AVF, DII, DIII, VI Intervalles F-F régulières	FLUTTER AURICULAIRE (bloc de type 2/1,3/1...)	
		φ	Large (>0,12s)	φ	Même aspect sur la dérivation Fusion, capture, dissociation AV	TV MONOMORPHE	
		φ	Large (>0,12s)	φ	Aspect sinusôide Amplitude variable	TV POLYMORPHE = Torsade de pointe	
		φ ou onde P rétrogrades	Fin		Fréquence : 150-250/min Si 180-200/min : penser à un Bouveret	TACHYCARDIE JONCTIONNELLE PAROXYSTIQUE	

RYTHME	FREQUENCE	ONDES P	QRS	CONDUCTION	REMARQUE	DIAGNOSTIC PROBABLE
IRREGULIER	LENT	oui	Fin	↑PR puis 1P/φ QRS		BAV II MOBILIZ I
	NORMAL					
	RAPIDE	Non = trémulation ligne de base	Fin	φ	Peut-être lent si prise Traitements +visible en V1, DII, DIII	FIBRILLATION AURICULAIRE
		φ	φ	φ	Rythme Anarchique	FIBRILLATION VENTRICULAIRE
		φ	Large >0,12s		Peut-être de fréquence variable Onde T négative Polarité QRS inversée	ESV
	φ	φ	φ	φ	ABSENCE PQRST	ASYSTOLIE COMPLETE
Lent			possible		Signes d'activité électrique Sans pouls	AESP

FIGURE 1.9 – Tableau récapitulatif [7].

1.4 Conclusion

La détection précoce des anomalies cardiaques est un domaine intéressant, qui est devenu le centre d'attention pour beaucoup de chercheurs en raison de la gravité de ses conséquences telles que les morts subites ou les séquelles physiques ou psychologiques postérieures.

La détection précoce des anomalies cardiaque est une tâche complexe et a besoin de grands efforts. Dans l'optique de traiter cette problématique dans ce projet, nous avons commencé par expliquer dans ce chapitre les concepts de base liés aux anomalies cardiaques ainsi que les détails d'un signal ECG.

Dans le chapitre suivant, nous parlerons des anomalies cardiaques qui concernent l'arythmie en plus d'un l'état de l'art en présentant les principaux travaux de recherche dans ce domaine.

Chapitre 2 :
L'étude de l'arythmie :
classification et état de l'art

Chapitre 2

L'étude de l'arythmie : classification et état de l'art

2.1 Introduction

L'apprentissage automatique pour l'ECG est un domaine qui a particulièrement évolué ces dernières années où il a réalisé un développement important dans les résultats, non seulement grâce à une meilleure connaissance du mécanisme physiopathologique, mais surtout par le développement de nouveaux moyens diagnostiques et thérapeutiques, mais il convient de mentionner que l'application dans le monde réel manque d'appareils intelligents simples qu'il est utilisable par toutes les personnes, où qu'elles se trouvent, comme un téléphone portable, en termes de résultats et de prix élevés.

La littérature sur l'apprentissage automatique ECG a présenté de nombreux travaux tels que les méthodes : support vector machine (abrègé SVM), l'arbre de décision (abrègé AD), clustering, réseau de neurone (abrègé RN) avec la diversité des cas comme nous l'avons vu dans le chapitre précédent.

Dans ce chapitre, on va présenter les anomalies cardiaques qui concernent l'arythmie ainsi qu'un état de l'art exposant quelques travaux de recherche pour la classification des signaux ECG afin de détecter les anomalies du rythme de cœur.

2.2 Les anomalies cardiaques : étude de l'arythmie

Si on parle des maladies cardiaques, qu'ils surviennent sur un cœur sain ou pathologique, ou bien de risque de mort subite, c'est peut-être les troubles du rythme, la première chose qui vient à l'esprit de médecin. Ils posent d'importants problèmes d'ordre diagnostique et thérapeutique. C'est un domaine qui a particulièrement évolué ces dernières années, à cause de développement de nouveaux moyens thérapeutiques et de diagnostics. Le schéma suivant récapitule les maladies qu'un ECG peut détecter où la plupart de ces maladies sont des anomalies de rythme cardiaque.

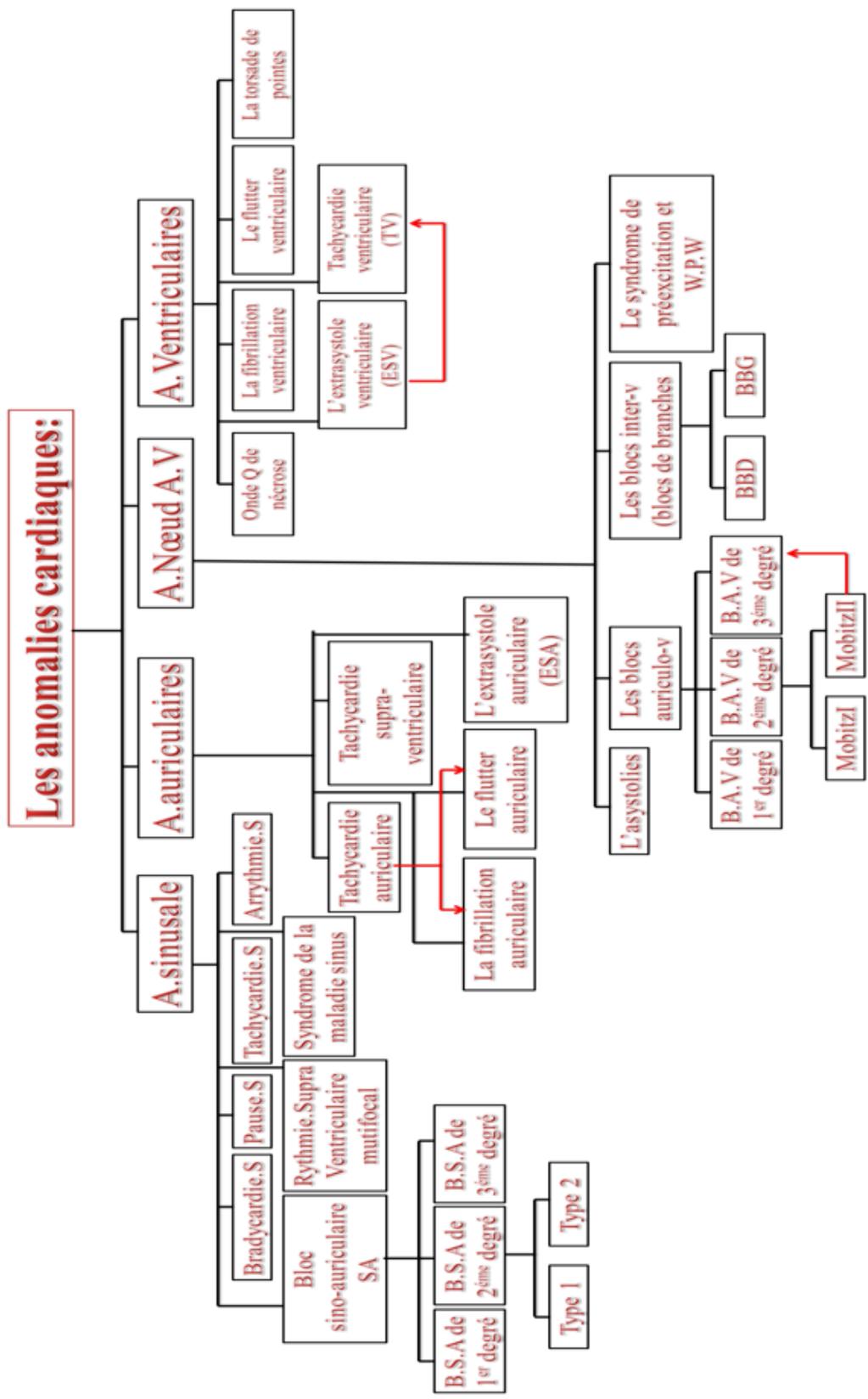


FIGURE 2.1 – Schéma récapitule les maladies qu'un ECG peut détecter.

Comme nous notons dans le schéma précédent, les anomalies de rythme cardiaque, peuvent être divisées en quatre catégories, expliquées comme suit :

2.2.1 Les anomalies sinusales

2.2.1.1 Arythmie sinusale

Le rythme sinusal irrégulier est lié aux modifications du tonus sympathique et parasympathique au cours du cycle respiratoire. Dans ce cas, on remarque que la fréquence augmente pendant l'inspiration et diminue pendant l'expiration [16].

Sur l'ECG, l'irrégularité du tracé, généralement visible à l'œil, se caractérise par une variation de plus de 0,12 secondes dans la durée des intervalles P-P successifs [9], de plus l'intervalle R-R se raccourcit au cours de l'inspiration et s'allonge durant l'expiration [16].

2.2.1.2 Bradycardie sinusale

La bradycardie sinusale est caractérisée par la présence systématique d'une onde P avant les complexes QRS (Figure 2.2) car l'origine de la contraction ventriculaire reste la dépolarisation du sinus et des oreillettes, comme lors de battements normaux, et de fréquence inférieure à 60/min. Les causes d'une telle arythmie sont multiples et souvent extérieures au système cardiovasculaire : traitement médicamenteux (bêtabloquant, digoxine, ...), hypothermie, urémie...etc [1].

Si la bradycardie sinusale est importante, elle peut s'accompagner de complexes d'échappement, jonctionnels ou ventriculaires [17].



FIGURE 2.2 – Bradycardie sinusale : le rythme est de l'ordre de 40 BPM [17].

2.2.1.3 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 [1].

Une tachycardie sinusale est considérée comme pathologique si elle est de longue durée et indépendante du contexte (Figure 2.3). Les causes de cette pathologie sont habituellement extra-cardiaques, et incluent tous les facteurs de stimulations du système nerveux : surmenage, anxiété, ... et l'influence de différentes substances comme l'adrénaline ou la caféine [1].



FIGURE 2.3 – Tachycardie sinusale : le rythme est de 120 BPM [17].

2.2.1.4 Bloc sino-auriculaire SA

C'est une anomalie de la conduction où les impulsions du nœud sinusal sont bloquées à leur sortie ("bloc"), ce qui empêche l'influx d'atteindre les oreillettes et les ventricules [17].

Le tracé se caractérise par un intervalle PP régulier, avec des pauses (absence complète d'un cycle) dont la durée est un multiple d'un intervalle PP normal [17].

Il en existe 3 sortes :

- B. S. A. de 1er degré : il est invisible à l'ECG.
- B. S. A. de 2ème degré :
 - Type 1 : (Wenckebach) on observe des intervalles P-P croissants ou décroissants mais l'intervalle P-P de la pause est toujours inférieur à deux intervalles P-P normaux [16].
 - Type 2 : Interruption intermittente de la conduction sino-auriculaire, avec des pauses égales au double de PP. De haut degré : avec au moins deux pauses successives, avec des pauses égales à un multiple de PP [17].
- B. S. A. de 3ème degré : Blocage complet : Le tracé reste sur la ligne isoélectrique jusqu'à échappement jonctionnels [17].

2.2.1.5 Pause sinusale

La pause sinusale se produit lorsque la zone de commande du nœud SA est soudain "arrêtée" et n'envoie pas les stimuli de commande. Après la pause due à l'arrêt sinusal, une nouvelle zone de pacemaker prend la commande mais elle ne tombe pas en phase avec le rythme précédent.

2.2.1.6 Rythme supra-ventriculaire multifocal [18]

Wandering Pacemaker ce terme anglais désigne un changement physiologique - progressif, temporaire et de courte durée - du site de contrôle de l'activation cardiaque. Il survient à l'occasion d'un changement d'équilibre entre les tonus sympathique et parasympathique (lié par exemple au stress ou à la relaxation lors de l'ECG), sans que la fréquence cardiaque en soit beaucoup modifiée. Il est totalement bénin et aucun traitement n'est nécessaire.

Ce changement progressif entre deux pacemakers physiologiques se produit :

- au sein du nœud sinusal : Dans ce cas, on observe une modification de l'onde P sinusale qui change d'un battement à l'autre.
- entre le nœud sinusal et un autre pacemaker para-sinusal dont le nœud du sinus coronaire : Dans ce cas, on observe une modification de l'onde P qui change d'un battement à l'autre.
- plus rarement, entre le nœud sinusal et le nœud AV : Il s'agit d'une compétition transitoire entre un rythme sinusal et un rythme jonctionnels accéléré. Dans ce cas, on observe un raccourcissement du PR, puis l'onde P se superpose progressivement avec le QRS, apparaît derrière et enfin se négative si l'activation rétrograde depuis le nœud AV vers l'oreillette recycle le sinus. Ultérieurement, l'onde P repasse devant le QRS et le nœud sinusal reprend la commande du rythme.

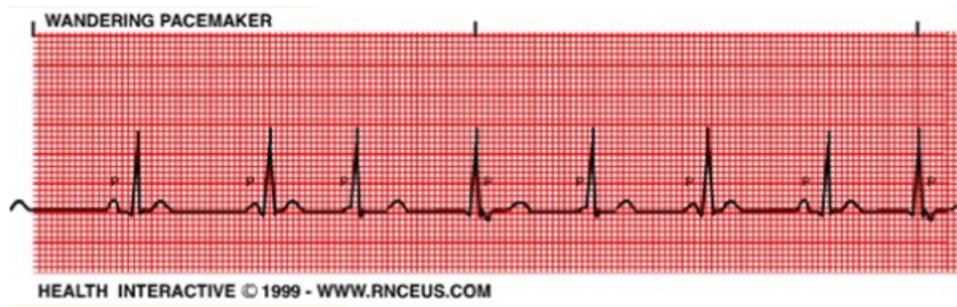


FIGURE 2.4 – Rythme supra-ventriculaire multifocal [17].

2.2.1.7 Syndrome de la maladie du sinus

Syndrome de la maladie du sinus. Condition chronique représentée par une variété d'arythmies supra-ventriculaires accompagnées de syncopes secondaires à des épisodes de bloc sino-auriculaire du troisième degré ou d'arrêt sinusal [17].

2.2.2 Les anomalies auriculaires

2.2.2.1 Tachycardie supra-ventriculaire

La tachycardie auriculaire paroxystique et la tachycardie paroxystique nodale prennent toutes deux naissance au dessus des ventricules et sont dénommées tachycardies supra-ventriculaires [2].

L'expression tachycardie supra-ventriculaire englobe tous les rythmes à QRS fins dont la fréquence, habituellement rapide, masque la morphologie des ondes P et/ou la provenance de l'arythmie [17].



FIGURE 2.5 – Tachycardie supra-ventriculaire [18].

2.2.2.2 Tachycardie auriculaire

Tachycardie auriculaire paroxystique est une cadence cardiaque rapide prenant naissance à partir d'un foyer ectopique situé dans une oreillette. la fréquence est habituelle-

ment entre 150 et 250. [2] Plus de trois extrasystoles auriculaires consécutives entraînant un rythme cardiaque à une fréquence supérieur à 100 par minute [17].

Comme le foyer est ectopique les ondes P de la tachycardie auriculaire ne ressemblent habituellement pas aux autres ondes p (celles d'avant la tachycardie) dans la même dérivation [2].

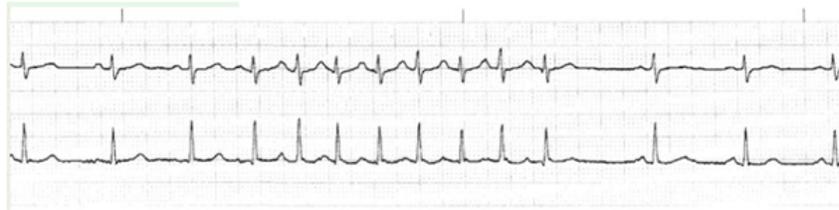


FIGURE 2.6 – La tachycardie auriculaire [18].

2.2.2.3 Fibrillation auriculaire

La fibrillation auriculaire est due à de nombreux foyers auriculaires ectopiques déchargeant à des fréquences différentes, ce qui entraîne un rythme auriculaire chaotique et irrégulier [2].

La fibrillation auriculaire se produit lorsque de nombreux foyers ectopiques auriculaires déchargent de façon continue. [2].

La fibrillation auriculaire apparaît souvent comme une ligne de base irrégulière sans ondes P. la réponse QRS n'est pas régulière et peut être rapide ou lente [2]. Visible surtout en VI, DII, DIII [7].

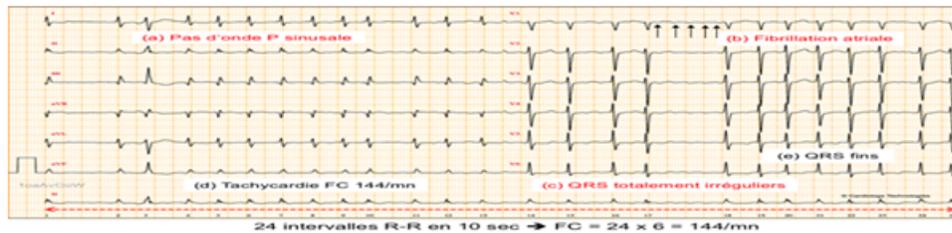


FIGURE 2.7 – La fibrillation auriculaire [17].

2.2.2.4 Flutter auriculaire [17]

Il provoque une accélération régulière de la fréquence auriculaire à environ 300/mn, souvent accompagnée d'une réponse ventriculaire à 150/min.

Il est lié à l'activité d'un foyer ectopique qui commande l'activité auriculaire. L'influx progresse à partir de ce foyer vers l'oreillette gauche, le septum et la paroi postérieure de l'oreillette droite, pour revenir plus lentement vers la paroi antérieure de l'oreillette droite. Le foyer initial est alors réactivé à son tour, maintenant la tachycardie à la même fréquence.

Sur l'ECG, il est décrit par des ondes auriculaires anormales, dites ondes F, qui sont parfaitement régulières, toutes identiques, en dents de scie, négatives en D2. La fréquence est autour de 300/mn.

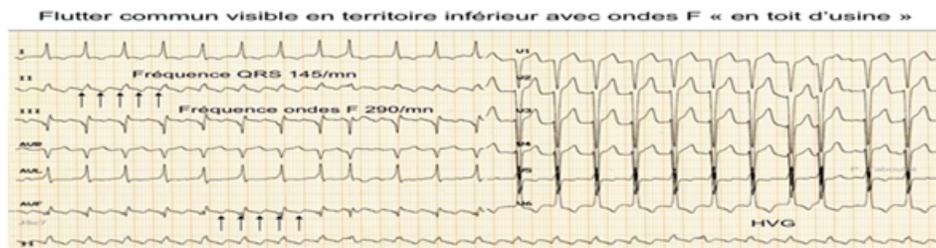


FIGURE 2.8 – Le flutter auriculaire [17].

2.2.3 Les anomalies du nœud A.V

2.2.3.1 Les blocs auriculo-ventriculaires

Le bloc du nœud AV entraîne un retard de l'impulsion auriculaire au niveau du nœud AV : la pause qui précède la stimulation des ventricules est plus longue que la normale [2].

Il en existe 3 sortes :

- B.A.V de 1er degré : se caractérise par un intervalle P-R supérieur à 0.20 seconde (un grand carré) [2].



FIGURE 2.9 – B.A.V de 1er degré [17].

- B.A.V de 2ème degré : Ils se définissent par l'interruption intermittente de la conduction auriculo-ventriculaire, ce qui se traduit par la survenue d'onde P non suivie de complexe QRS (onde P bloquée). Le nombre d'onde P est ainsi supérieur au nombre des complexes QRS [2].

Parmi les B.A.V. du 2ème degré, on distingue deux types :

- Mobitz type I (Wenckebach) : Sur le tracé, cela se traduit par un allongement progressif du PR, jusqu'à la survenue d'une onde P bloquée. Cette séquence se répète : ce sont les périodes de Wenckebach [17].

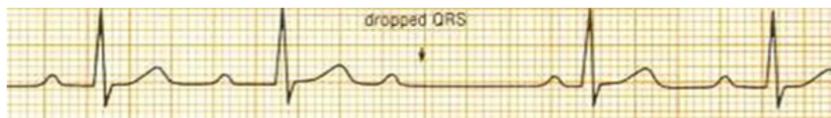


FIGURE 2.10 – B.A.V de 2ème degré Mobitz typeI [17].

- - Mobitz type II : Sur le tracé, on observe fréquemment et régulièrement une onde P bloquée. L'intervalle PR reste constant [17]. il manque complexe QRS [2].



FIGURE 2.11 – B.A.V de 2ème degré Mobitz typeII [17].

- B.A.V de 3ème degré : Il y a alors une absence complète de conduction entre les oreillettes et les ventricules. On obtient une totale dissociation auriculo-ventriculaire [17].



FIGURE 2.12 – B.A.V de 3ème degré [17].

2.2.3.2 Les blocs inter ventriculaires (Blocs de branches) [7].

il y a ralentissement ou interruption de la conduction sur les fibres de la branche droite ou gauche de Purkinje. La conduction des ventricules ne se fait pas en même temps, la conduction se fait par proximité ce qui implique un QRS élargi.

- Le bloc de branche droit (BBD), sur le tracé on retrouve :
 - un rythme supra-ventriculaire.
 - QRS large, positive en V1 V2 et négative en V6.
 - L'onde T est contraire à l'onde terminale du complexe QRS.
- Le bloc de branche gauche (BBG), sur le tracé on retrouve :
 - C'est également un rythme supra-ventriculaire.
 - QRS large avec un sommet en creux positive en V5, V6, AVL et DI et négative en V1.

2.2.4 Les anomalies ventriculaires

2.2.4.1 L'extrasystole ventriculaire (ESV)

Extrasystole ventriculaire (ESV), Ce type de battement (anormal) a pour origine la dépolarisation spontanée d'un groupe de cellules des ventricules [1], appelé alors foyer ectopique Ventriculaire. Il est donc caractérisé par les propriétés suivant :

- l'onde R n'est pas précédée d'une onde P (il n'est donc pas précédé d'une onde P), puisqu'il n'y a pas eu d'activité auriculaire préalable [1].
- Si une ESV tombe sur une onde T, elle survient pendant une période vulnérable et des troubles du rythme dangereux peuvent en résulter [12] Onde T anormale (négative).
- la durée du complexe est supérieure à la durée d'un complexe QRS normal.
- En générale suivi d'une pause [7]

Comptez les ESV : [7], [2].

- Si 3 ESV successive = Tachycardie ventriculaire.

- Si il y a plus de 6 ESV/min = risque de Tachycardie ventriculaire.
- Donc il est facilement reconnu sur le tracé électrocardiographique [2].

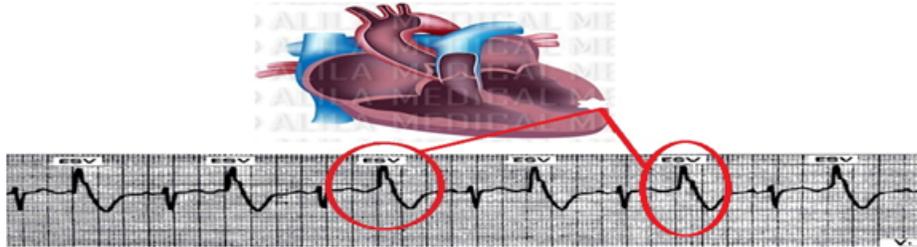


FIGURE 2.13 – L’extrasystole ventriculaire (ESV).

2.2.4.2 La tachycardie ventriculaire

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 (Figure 2.14). 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 [1].

Les accès de tachycardie ventriculaire peuvent signifier l’existence d’une maladie des artères coronaires [2].

Sur l’ECG, on voit un tracé qui ressemble à une succession de plus de 3 ESV successives [17].

3 critères sont nécessaires [17].

- Une succession de QRS larges supérieurs à 0,14 seconde, de fréquence régulière supérieure à 100 par minute, avec dépression ou une élévation du ST et des ondes T inversées.
- Il y a une dissociation auriculo-ventriculaire.
- Il doit exister quelques complexes QRS fins d’origine sinusale, appelés complexes de capture.

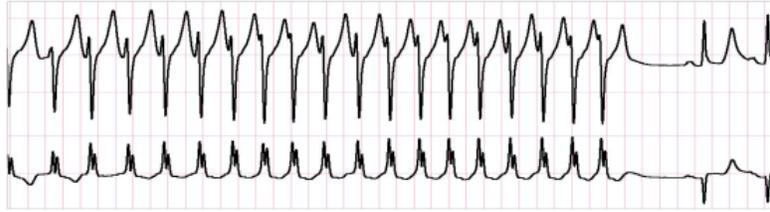


FIGURE 2.14 – Tachycardie ventriculaire (TV). On observe sur l'enregistrement une succession d'extrasystoles ventriculaires à une fréquence de 150 bpm ; après la phase de TV, le cœur reprend ici spontanément un rythme sinusal normal [1].

2.2.4.3 La torsade de pointes

La torsade de pointes Ce trouble du rythme ventriculaire individualisé par dessertenne se distingue par des QRS d'amplitudes et de polarités variables. Il débute par une extrasystole ventriculaire qui tombe sur l'onde T ou l'onde U puis survient immédiatement une succession rapide de complexes ventriculaires élargis. Sur le tracé, on note des QRS tantôt positifs, tantôt négatifs, une ESV prématurée survenant tardivement sur un QT supérieur à 0,50 seconde, une fréquence supérieure à 160-280/mn [17].

2.2.4.4 Le flutter ventriculaire [18]

Tachycardie ventriculaire monomorphe très rapide où les complexes ventriculaires sont relativement similaires et ondulent sous forme sinusoïdale à plus de 250/mn.

Ce trouble du rythme en rapport avec une réentrée est observé généralement sur cœur ischémique. Il est mal toléré sur le plan hémodynamique et dégénère rapidement en fibrillation ventriculaire.

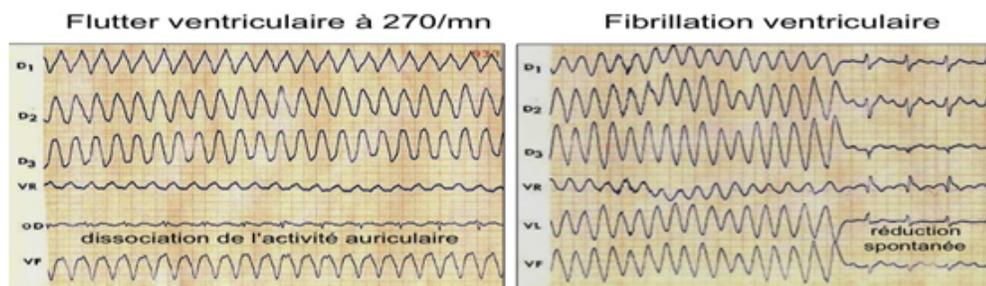


FIGURE 2.15 – Flutter ventriculaire [18].

2.2.4.5 La fibrillation ventriculaire [18]

Activité électrique anarchique du myocarde ventriculaire. La désynchronisation totale qui en résulte est responsable d'un arrêt circulatoire (asystolie). Une FV démarre généralement comme une salve de tachycardie ventriculaire rapide et régulière qui se fragmente ensuite en de multiples ondelettes de réentrée.

Cliniquement, le patient se présente en état de mort apparente quelques secondes après le début. Ce trouble du rythme fatal complique généralement une extrasystolie maligne ou une tachycardie ventriculaire.

Sur l'ECG, on observe des ondulations anarchiques, d'amplitude et de fréquence variables qui se raréfient et disparaissent jusqu'à un tracé plat. La mort est inéluctable en l'absence de défibrillation immédiate.

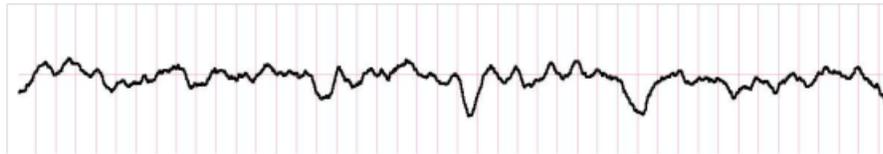


FIGURE 2.16 – La fibrillation ventriculaire [1].

2.3 Travaux généraux sur les signaux ECG

La figure 2.17 représente les trois principales tâches de data mining dans le secteur de la santé en relation avec certains signes vitaux dans la littérature consultée par [19]. Nous pouvons observer que l'ECG bénéficie de la plus grande attention dans la littérature comparée aux autres signes vitaux. Cela est dû à l'importance du diagnostic ECG et à la sensibilité de ses résultats.

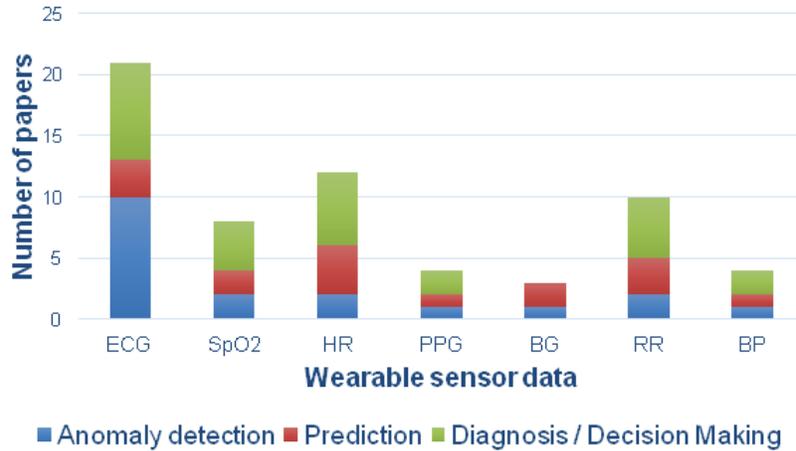


FIGURE 2.17 – Les trois tâches de data mining en relation avec les signes vitaux de santé [19].

De nombreux travaux ont été réalisés autour de l’analyse du signal ECG, et en raison de la complexité et de l’importance de l’étude, les chercheurs se concentrent chaque fois sur une partie spécifique du processus d’analyse de l’ECG, de l’acquisition du signal à la prédiction de la maladie, en passant par le filtrage et la réduction du bruit, extraction et classification des fonctionnalités, et utilisation de nombreuses méthodes, techniques et algorithmes. Selon [19], la figure 2.18 illustre l’application des méthodes de data mining à l’analyse du signal ECG :

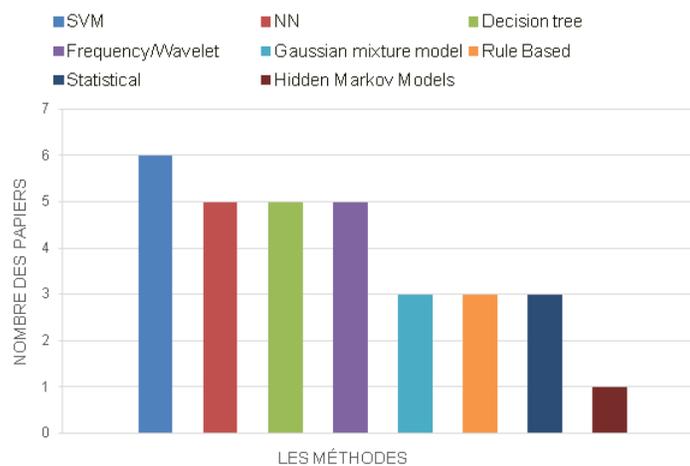


FIGURE 2.18 – Méthodes d’exploration de données utilisées dans la littérature pour analyser le signal ECG [19].

2.4 Revue de littérature

Nous exposons dans cette section, quelques travaux qui ont été réalisés autour de l'analyse du signal ECG par une approche de classification

2.4.1 Classification

2.4.1.1 Arbre de décision

Les auteurs de [20] ont étudié l'anomalie de rythme cardiaque par une analyse de l'électrocardiogramme (ECG) à l'aide de la méthode : l'Arbre de décision, l'objectif est d'évaluer l'effet des méthodes d'apprentissage automatique dans la création du modèle de classification concernant l'insuffisance cardiaque normale et l'insuffisance cardiaque congestive sur la série chronologique de l'ECG à long terme

Cette étude a été réalisée en deux phases : l'extraction des caractéristiques et la classification. En phase d'extraction de caractéristiques, la méthode de Burg autorégressive est appliquée pour extraire les caractéristiques. Au cours de la phase de classification, cinq classificateurs différents sont examinés, à savoir l'arbre décisionnel C4.5, k - le plus proche voisin, la machine à vecteurs de support, les réseaux de neurones artificiels et le classificateur de l'arbre de décision. Les signaux ECG ont été acquis à partir des bases de données ECG BIDMC [21] sur l'insuffisance cardiaque congestive et les bases de données PTB Diagnostic [22].

Les résultats expérimentaux de cette étude ont évalué plusieurs mesures statistiques telles que la sensibilité, la spécificité, la précision, en montrant que la méthode Arbre de Décision donne une précision de classification de 100%.

Cependant, on n'est pas certain d'avoir ce taux qui dépend du nombre d'arbres qui est choisi de manière aléatoire.

2.4.1.2 Cluster K-moyennes

Dans la littérature, de diverses techniques de cluster ont été utilisées pour l'étude de l'arythmie. Le cluster est une technique recommandée pour l'analyse et l'interprétation des enregistrements Holter ECG à long terme [23].

Dans [23], le cluster K-moyennes a été utilisé avec une distance euclidienne au carré, pour l'analyse et l'interprétation des enregistrements Holter ECG. Les ensembles de données comportant quatre types d'arythmie ont été créés à l'aide des bases de données MIT-BIH. Les données sont classées en cinq types d'arythmie, à savoir Normal (N), Contraction ventriculaire prématurée (PVC), Battements stimulés (P), Bloc de branche gauche (LBBB) et Bloc de branche droite (RBBB).

[23] a donné de bons résultats mais les imperfections demeurent néanmoins là, où on a observé un faible taux d'évaluation qui avoisine les 75%.

2.4.1.3 Les réseaux de neurones convolutionnels (CNN)

En raison de la complexité des signaux ECG, il est difficile d'analyser manuellement ces signaux. De plus, l'interprétation des signaux ECG est subjective et peut varier d'un expert à l'autre. Par conséquent, les auteurs de [24] ont présenté une technique de réseau de neurones convolutionnels (CNN) pour détecter automatiquement les différents segments de l'ECG. Le système proposé consiste en un CNN profond de 11 couches avec la couche de sortie de quatre neurones, chacun représentant la classe ECG normale (Nsr), Fibrillation auriculaire (Afib), flutter auriculaire (Afl) et fibrillation ventriculaire (Vfib). Dans ce travail, les auteurs ont utilisé des signaux ECG d'une durée de deux à cinq secondes sans détection du complexe QRS. Ils ont atteint une précision de 92,50% pour deux secondes de segments ECG, ils ont aussi obtenu une précision de 94,90% pendant cinq secondes de la durée de l'ECG. Il est à noter que les périodes évaluées sont limitées et peuvent ne pas refléter le cas général d'une telle étude.

On note que les CNN sont applicables de façon efficace pour l'analyse de l'ECG mais à condition que ce signal soit traité en tant qu'image, cependant, le traitement précis et cohérent d'un tel signal sensible exige que l'ECG soit analysé de sa forme numérique qui est une suite de valeurs réelles.

2.4.2 Les applications connectées pour détecter les arythmies cardiaques : y-a-il un dispositif à la portée des citoyens ?

Avec Android Wear et Apple Watch, des millions de personnes portent à tout moment des capteurs de fréquence cardiaque aux poignets. Ces appareils portables génèrent des ordres de grandeur plus nombreux que les ECG [25].



FIGURE 2.19 – Les applications connectées pour détecter les arythmies cardiaques ECG [25].

La question spécifique qui se pose avec la présence de ces montres est la suivante : pouvons-nous diagnostiquer les personnes à travers des données de fréquence cardiaque mesurées par des montres intelligentes qui assure une analyse à un prix à la portée des citoyens ?

2.4.3 Travaux similaires et enregistrement permanent des signaux ECG

Suite à notre recherche bibliographique et d'après nos modestes connaissances, aucun travail de recherche ou technologie ne permet un enregistrement permanent des signaux ECG qui est exigence des spécialistes en Cardiologie.

A cause de prix élevé des appareils et l'absence des systèmes de sauvegarde les données

ECG, nous proposons dans ce projet de Master la réalisation d'un prototype d'acquisition, de sauvegarde et d'analyse des signaux ECG pour la détection précoce des anomalies du rythme cardiaque. La réalisation de ce prototype s'appuie aussi sur la proposition d'une optimisation de la méthode L'arbre de décision avec l'algorithme recherche tabou afin d'améliorer la précision de la détection et de la classification de l'ECG.

2.5 Conclusion

Dans ce chapitre, nous avons présenté les anomalies cardiaques qui concernent l'arythmie en plus des quelques travaux visant la classification des signaux ECG pour la détection précoce des problèmes du rythme cardiaque. Nous avons constaté que le taux de classification donné par ces travaux était modéré et que la sauvegarde permanente de l'ECG est à projeter. De plus, il est nécessaire de penser à la démocratisation des dispositifs de suivi du rythme cardiaque par une production à bas prix. Par conséquent, nous présentons dans le chapitre suivant la conception de notre système proposé dans l'espoir de contribuer à la résolution des ces problématiques

Chapitre 3 :

Conception du système
de détection de l'arythmie :
par l'arbre de décision améliorée

Chapitre 3

Conception du système de détection de l'arythmie : par l'arbre de décision améliorée

3.1 Introduction

Dans ce chapitre, nous présentons notre vision conceptuelle de notre système de classification qui concerne l'analyse des battements ECG pour l'arythmie. On va expliquer les étapes et les modules composant notre système. De plus, nous illustrons les principales étapes de la conception détaillée en expliquant les différents éléments du système et précisant leurs fonctionnements.

3.2 Conception globale du système

Dans cette section, le nouveau système de classification automatique pour l'analyse des battements ECG est illustré. Notre proposition est basée sur la sauvegarde des données de la santé relative au rythme cardiaque car le médecin a besoin de connaître les divers changements que subit le patient, pour qu'il puisse prendre la décision en thérapie. Cet historique enregistré est primordial qui aide à consulter le rythme cardiaque dans des périodes antérieures dans lesquelles une attaque peut survenir subitement en dehors de la

présence d'un médecin. La détection de cette attaque peut aider à diagnostiquer et prescrire un traitement convenable au patient sans faire référence à des traitements préventifs et généraux issus d'une ignorance du type d'attaque. Dans notre système, nous introduisons aussi un algorithme d'optimisation avancé qui l'algorithme de la recherche tabou (RT) dans le but de trouver le nombre optimal des arbres exigé par l'approche de classification « l'Arbre de décision », conduisant à une grande précision de la classification du signal ECG. Ce choix est fait sur la base que RT est considérée comme une des méta-heuristiques qui a montré son efficacité sur de nombreux problèmes difficiles au cours des 20 dernières années [27].

Globalement, on peut représenter l'architecture de notre système de détection précoce et d'un suivi d'une maladie ou d'une anomalie de rythme cardiaque comme suit :

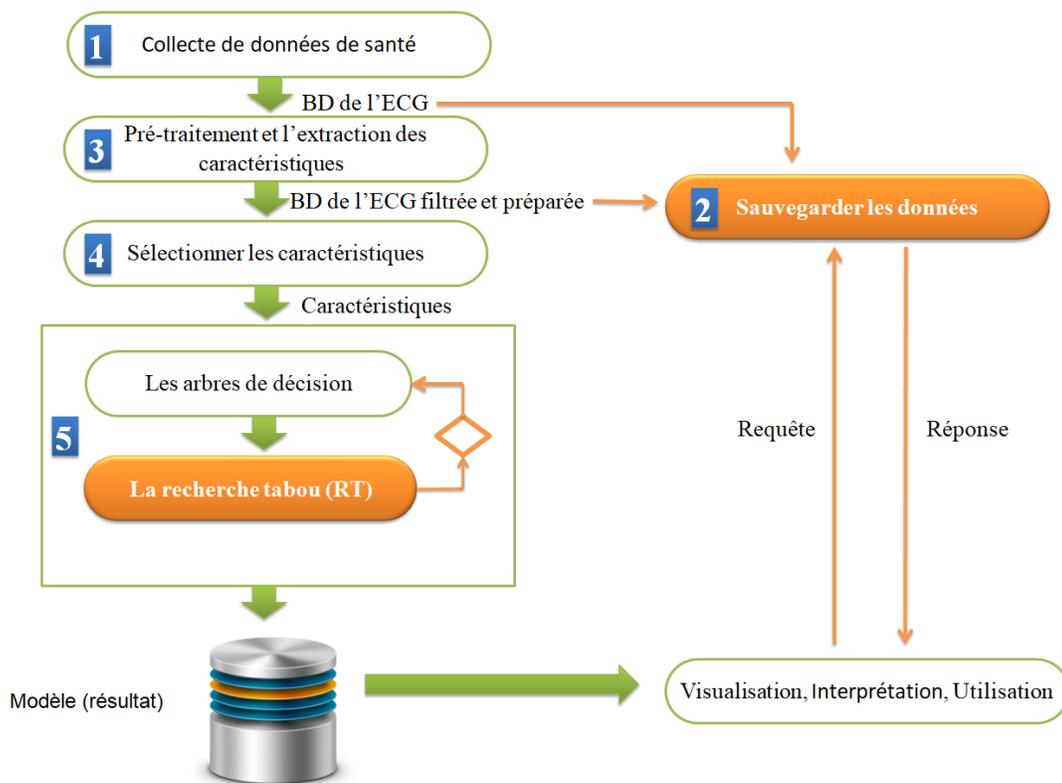


FIGURE 3.1 – L'architecture générale du système.

Comme illustré dans l'architecture précédente, nous pouvons diviser notre système en cinq parties :

3.2.1 Collecte de données de santé

Dans ce module, on collecte deux corpus de données, le premier est ramassé à partir d'un site web pour les données de sante (PhysioNet) [28] pour cela, on a choisi les bases de données des maladies cardiaques liées à l'arythmie pour aider notre système à un apprentissage plus précis. Le deuxième type de données est les signaux ECG des patients issues de la plate-forme cooking hacks.

3.2.2 Sauvegarder les données

Suite aux recommandations du **Docteur OKBI Ridha**, Médecin spécialiste en Cardiologie, qui nous a suggéré d'enregistrer du signal ECG 24h/24h et de l'envoyer au médecin en cas de changement sur anormal du rythme (ECG Naturel ou normale), avec l'utilisation d'une capture de l'ECG permanent. C'est une nécessité exigée par les médecins cardiologues pour qu'ils puissent prendre la décision de thérapie adéquate; il s'agit de connaître les divers changements que subit le patient, à des moments différents, et aussi car Il y a des médicaments qui nécessitent de connaître les enregistrements précédents afin d'être prescrit.

3.2.3 Prétraitement et l'extraction des caractéristiques

3.2.3.1 Prétraitement

Le signal ECG est exposé à plusieurs types de bruits au cours des étapes de génération et d'agrégation. Cela est dû à la nature du corps humain et à la manière dont l'ECG est mesuré. Plusieurs types de bruits de différentes sources sont énumérés, tels que les activités musculaires, les lignes électriques, les étirements de la peau et les mouvements des électrodes, les mouvements du cœur dus à la respiration, etc. Par conséquent, le débruitage ECG vise à éliminer ou au moins à minimiser les signaux indésirables d'un enregistrement ECG, sans entraver les informations cliniques contenues dans le signal lui-même. Dans cette étude, nous considérons que cette phase est exécutée comme une boîte noire par les approches de débruitage proposées dans la littérature du traitement de signal [29] fonctionnant comme le montre la figure 3.2.

3.2.3.2 L'extraction des caractéristiques

Le complexe R-peak et QRS sont les caractéristiques les plus cruciales du cycle cardiaque. Leurs emplacements sont utilisés pour obtenir et comprendre le relevé de rythme de l'ECG. De nombreux travaux de recherche sont effectués par des scientifiques dans ce domaine dans le but d'améliorer le débit du pic R et d'autres caractéristiques telles que l'onde T, l'onde P et le complexe QRS [30].

Ces caractéristiques sont fondamentales dans l'interprétation de l'ECG et constituent la principale caractéristique utilisée dans le système de classification des battements [30].



FIGURE 3.2 – Prétraitement et extraction des caractéristiques.

3.2.4 La sélection des caractéristiques

Dans cette étape, après une analyse numérique du signal ECG en découvrant le début et la fin de chaque onde (Prétraitement et extraction des caractéristiques), nous avons besoin de calculer toutes les caractéristiques considérées selon les besoins de la classification en visant par exemple le calcul des amplitudes des ondes p, T... etc.

3.2.5 Entraînement avec le nombre optimal des nœuds dans l'arbre de décision

Dans cette partie, nous appliquons une méthode d'entraînement et d'apprentissage, qui est la méthode : l'arbres de décision pour une aide à la décision à une détection précoce et à un suivi d'une anomalie de rythme cardiaque, ce qui peut empêcher de graves conséquences telles que les morts subites, les séquelles physiques ou psychologiques postérieures. Cependant, l'appel de la méthode l'arbre de décision exige l'introduction d'un paramètre choisi par l'utilisateur de manière aléatoire ; il s'agit du nombre des arbres. Ce choix aléatoire peut affecter négativement la précision de l'apprentissage. Pour cela, nous

proposons de ce projet de Master de renforcer la méthode de classification par l'intégration d'un nouvel algorithme d'optimisation qui est l'algorithme : la recherche tabou (RT ou Tabu Search en Anglais). Cet algorithme a pour objectif de trouver le nombre optimal des arbres à introduire dans l'arbre de décision, ce qui peut garantir un entraînement de notre système et qui va essayer d'apprendre et de créer un modèle donnant une classification de l'ECG avec un taux de prédiction très élevé.

3.3 Conception détaillée du système

Dans cette section, on peut représenter et détailler l'architecture de notre système de détection précoce et de suivi d'une anomalie de rythme cardiaque, chaque partie du système proposé est illustrée en détail en mentionnant son principe de fonctionnement.

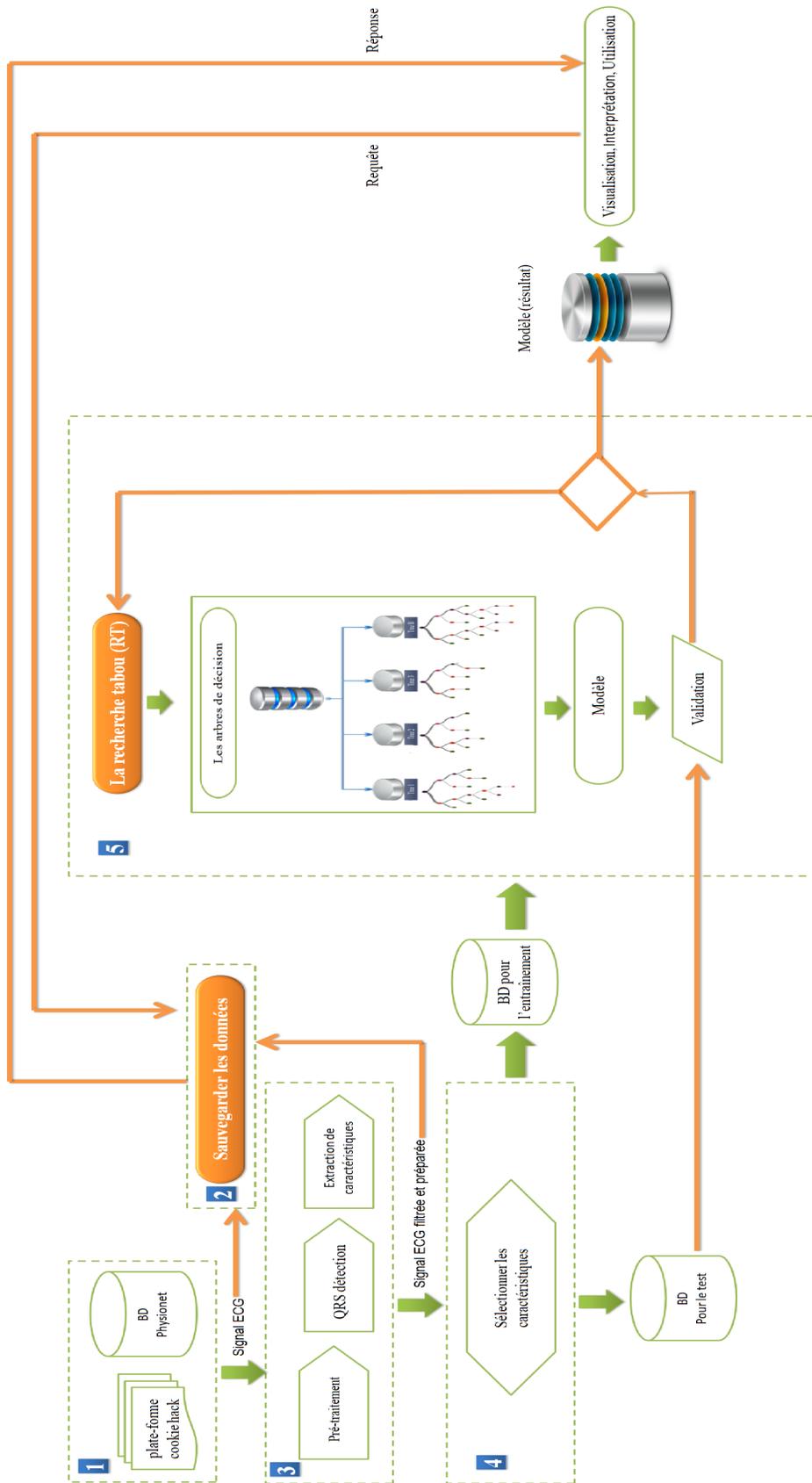


FIGURE 3.3 – L'architecture détaillée du système.

3.3.1 Collecte de données de l'ECG

Puisque l'ECG est le plus représentatif de l'état rythmique du cœur et le plus largement utilisé dans la surveillance des maladies cardiaques, et en raison de d'avoir un apprentissage plus précis qui se base sur des données diversifiées et volumineuses, nous avons choisi de collecter les données issues étudiée de deux sources :

3.3.1.1 Site Web PhysioNet

PhysioNet [28] offre un accès Web gratuit à de grandes collections de signaux physiologiques enregistrés (Physio Bank) [31] et au logiciel libre associé (PhysioToolkit) [32].

3.3.1.2 plate-forme E-Health

E-Health Sensor Shield permet aux utilisateurs d'Arduino et de Raspberry Pi de réaliser des applications biométriques et médicales nécessitant une surveillance corporelle à l'aide de plusieurs capteurs. Ces informations peuvent être utilisées pour surveiller en temps réel l'état d'un patient ou pour obtenir des données sensibles en vue d'une analyse ultérieure et en vue d'un diagnostic médical. Les informations biométriques recueillies peuvent être envoyées sans fil en utilisant l'une des 6 options de connectivité disponibles, telles que les interfaces Wi-Fi, 3G, GPRS, Bluetooth, 802.15.4 et ZigBee, selon l'application [33].

Il est intéressant, de se référer aux types de fichiers qui sont acquis, à savoir :

- Pour la plate-forme E-Health : un fichier contenant des lignes de valeurs réelles, représentant l'activité électrique du cœur,
- Pour le Site Web PhysioNet : trois types de fichiers sont acquis :
 - un fichier .dat : contenant des lignes de valeurs réelles numérisées, un fichier.
 - un fichier ann : contenant l'annotation de tous les temps du fichier .dat.
 - un fichier .hea : contenant des informations complémentaires sur le patient et le signal ECG.

3.3.2 Sauvegarde des données

Suivant une recommandation stricte donnée par le **Docteur OKBI, Spécialiste en cardiologie**, en soulignant l'importance de la conservation durable et non interrompue des données de l'ECG des patients, nous proposons le développement d'une infrastructure d'enregistrement des données de type massive (Big data), basée sur le modèle multidimensionnel [34], il s'agit d'un entrepôt de données [35] conçu sous forme d'un schéma en constellation de faits [36], qui est basé sur deux schémas en flocon [37], voir la figure 3.4.

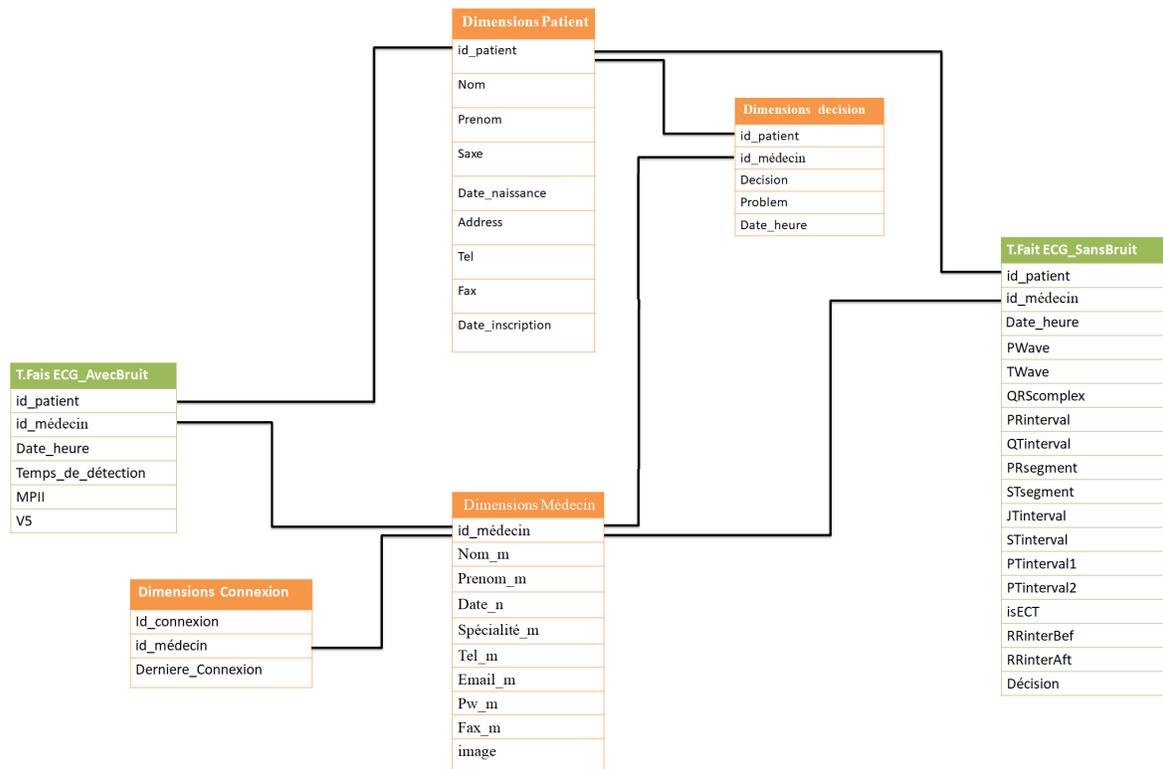


FIGURE 3.4 – Le modèle multidimensionnel proposé du système de sauvegarde de données ECG.

Le premier schéma en flocon est dédié à l'enregistrement de données des signaux ECG qui nous pouvons acquérir à partir d'un dispositif de santé connectée équipé d'un nano-ordinateur Raspberry pi 3 (notre cas) ou d'une montre intelligente. Il contient une table de faits et deux tables de dimension principales, voir la figure 3.4

La table de faits **ECG_AvecBruit** contient les champs **id_patient** pour mémoriser

l'identifiant de patient et **id_médecin** qui représente l'identifiant du médecin traiteur de la situation de ce patient, **Date_heure** représente la date et l'heure d'acquisition de signal ECG, quant au champ **Temps_de_détection**, il représente un fait qui est la seconde d'acquisition de signal ECG. **MPII** et **V5** sont considérés aussi comme des faits dont les valeurs sont des signaux obtenus par des dérivations différentes.

Les tables de dimensions sont proposées pour mémoriser les données personnelles du «**patient**» : son identifiant, nom, prénom, sexe, date de naissance, adresse utilisée en cas d'urgence pour appeler une ambulance, numéro de téléphone, numéro de Fax et la date d'inscription du patient dans le système.

Pour mémoriser les données personnelles du «**médecin** » : on enregistre son identifiant, nom, prénom, date de naissance, spécialité, numéro de téléphone, Email, mot de passe pour se connecter à son compte, Fax, sa photo. On extrait une table de dimension secondaire pour une raison de normalisation de la table médecin car on est dans le cas d'un schéma en flocon, cette table est appelée Dimensions **Connexion** contenant les champs : identifiant de la connexion, identifiant du médecin et le timing du connexion de médecin.

Pour le deuxième schéma en flocon, on propose la table de faits **ECG_SansBruit** qui contient les champs **id_patient** pour mémoriser l'identifiant de patient et **id_médecin** qui représente l'identifiant du médecin traiteur, **Date_heure** représente la date et l'heure d'acquisition de signal ECG, **PWave**, **TWave**, **QRScomplex**, **PRinterval**, **QTinterval**, **PRsegment**, **STsegment** **JTinterval**, **STinterval**, **PTinterval1**, **PTinterval2**, **isECT**, **RRinterBef**, **RRinterAft**, qui sont considérés comme des caractéristiques des signaux extraits et sélectionnés par notre système, **Décision** représente la décision de chaque battement dans le fichier. Les tables de dimensions de cette table de fait sont les tables **Patient** et **Médecin**.

La table **Décision** représente une table de décision secondaire issue de la table **Patient**, aussi pour une raison de normalisation, elle exprime les décisions finales pour chaque fichier des données acquises sur ECG du patient, elle contient les champs **ip_patient**, **id_médecin**, **décision**, **Date_heure**, **Problème**. Le champ **Problème** explique le type de l'anomalie détectée.

3.3.3 Prétraitement et extraction des caractéristiques

3.3.3.1 Pré-traitement

a Entrée

L'entrée de cette étape est les fichiers ECG bruts (avec bruit).

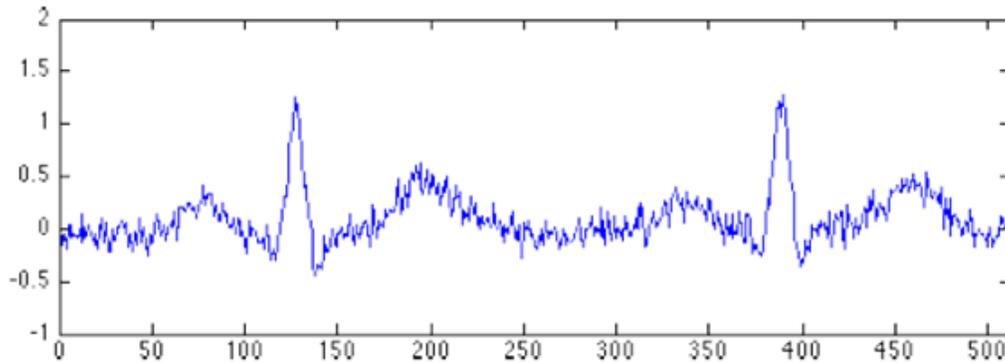


FIGURE 3.5 – Signal ECG avec les bruits [38].

b Sortie

Sortie de cette étape est les fichiers ECG sans bruit.

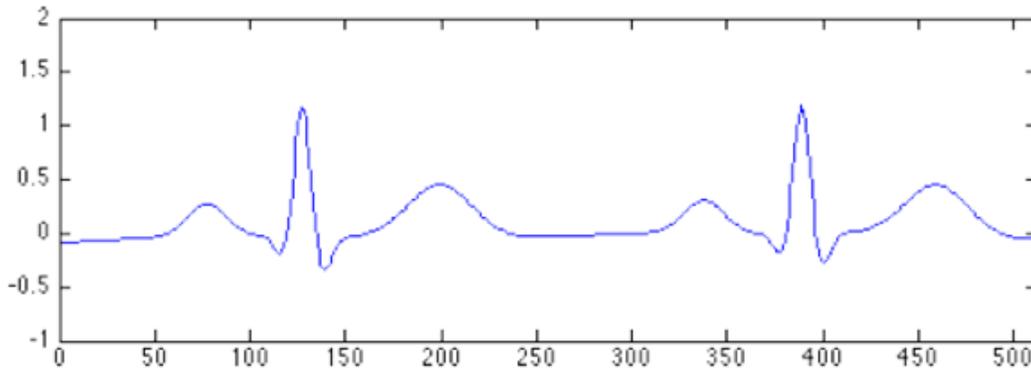


FIGURE 3.6 – Signal ECG sans les bruits [38].

3.3.3.2 L'extraction des caractéristiques

a Entrée

L'entrée de cette étape est représentée par les fichiers ECG sans bruit résultant de l'étape précédente, voir figure 3.6

b Sortie

Le résultat de cette étape est un fichier annoté, qui contient la position temporelle de chaque caractéristique extraite avec son annotation. Ces annotations serviront au calcul des caractéristiques des battements formant l'ECG.

3.3.4 Sélection des caractéristiques

a Entrée

Nous utilisons le fichier annoté résultant de l'étape précédente comme entrée dans cette étape pour calculer les caractéristiques nécessaires.

b Sortie

Après le calcul des caractéristiques souhaitées, nous obtenons une table des caractéristiques de battements, qui contient les valeurs de ses caractéristiques extraites sur la base des battements successifs.

3.3.5 Entraînement avec le nombre optimal des d'arbres dans l'arbre de décision

3.3.5.1 L'arbre de décision

La technique Arbre de décision est une technique de classification très utilisée par les chercheurs, et c'est une méthode très efficace d'apprentissage supervisée, pour cela nous l'avons choisi dans l'étape de modélisation pour trouver notre modèle de décision. Il s'agit de partitionner un ensemble de données en des groupes homogènes le plus possible du point de vue de la variable à prédire [39]. Il consiste à générer de nombreux arbres à partir de sous-échantillons sélectionnés au hasard (sur la base de la méthode bootstrap). On utilise une partie de la base de données, souvent 70% des données pour l'apprentissage et 30% pour le test, afin de former le modèle. La prise de décision finale se fait après avoir connaître les résolutions de chaque arbre généré, elle est obtenue par l'application du vote entre les sous-décisions qu'on aura.

3.3.5.2 Le nombre optimal des d'arbres pour l'exécution de l'arbre de décision

Une fois l'arbre de décision est construit, il peut être d'une taille très importante qui peut épuiser les ressources de calcul et de stockage, pour cela on doit effectuer des opérations d'élagage qui consistent à éliminer de l'arbre les branches les moins significatives. Il peut être aussi d'une taille très petite mais le taux de classification soit très faible.

De cela, nous pouvons déduire l'importance de paramétrer "le nombre des d'arbres dans l'arbre de décision".

C'est pour cela que nous avons proposé d'utiliser **l'algorithme recherche tabou (RT)** comme étant une méthode métaheuristique locale utilisée pour résoudre des problèmes complexes et/ou de très grande taille. On vise à trouver un nombre optimal d'arbres, conduisant à un taux de classification le plus élevé avec la plus petite taille d'arbre possible. RT découvre un nombre optimal des arbres et vérifie la précision du classificateur Arbre de décision, obtenu avec ce nombre d'arbres. Ce processus est répété jusqu'à l'obtention d'un nombre optimal d'arbres, comme le montre l'algorithme suivant.

Algorithme Recherche Tabou (RT) pour trouver le nombre optimal des d'arbres dans l'arbre de décision (AD)

Étape 1: Initialiser la configuration du système

```
Nombre initial d'arbres au hasard num_arb(0)
Meilleur_nbr_arb=num_arb(0)//enregistrer le numéro d'arbre optimal dans
Meilleur_nbr_arb
Appliquer fonction AD(num_arb)
Calculer meilleure_précision=précision(num_arb)//calculer la précision de la
classification ECG en fonction de la solution initiale et la sauvegarder
initialise list_T=∅ // list_T: liste tabou
```

Étape 2: Répéter (critère - vérifié)

```
Générer une solution de voisinage aléatoire nouvelle_num_arb=num_arb + Δ x //
on va choisir un nombre de voisinage aléatoire, pour que ce chiffre ne soit pas
dans la liste tabou
Appliquer fonction AD(nouvelle_num_arb) // évaluer la solution du voisin
Calculer actuelle_précision=précision(nouvelle_num_arb)//calculer la précision
de la classification ECG en fonction de la solution du voisin
Si(actuelle_précision > meilleure_précision) // cette nouvelle haute précision
  Si( longueur(list_T) < max) // Pas plein
    Ajouter l'ancien numéro d'arbre dans liste_T(Meilleur_nbr_arb)
    meilleure_précision=actuelle_précision // sauvegarder la nouvelle haute
    précision
    Meilleur_nbr_arb = nouvelle_num_arb // sauvegarder la nouvelle num
    d'arbre
  Si non // la liste est plein
    vider la liste avec le principe FIFO et Ajouter l'ancien numéro d'arbre
    dans liste_T(Meilleur_nbr_arb)
  Fin si
Si non
  Si(actuelle_précision < meilleure_précision)
    Si( longueur(list_T) < max) // Pas plein
      Ajouter nouvelle numéro d'arbre dans liste_T(nouvelle_num_arb)
    Si non // la liste est plein
      vider la liste avec le principe FIFO et Ajouter nouvelle numéro d'arbre
      dans liste_T(nouvelle_num_arb)
    Fin si
  si non // actuelle_précision == meilleure_précision
    Si(nouvelle_num_arb > Meilleur_nbr_arb) // pour une arbre de plus petit
    taille
      Si( longueur(list_T) < max) // Pas plein
        Ajouter nouvelle numéro d'arbre dans liste_T(nouvelle_num_arb)
      Si non // la liste est plein
        vider la liste avec le principe FIFO et Ajouter nouvelle numéro
        d'arbre dans liste_T(nouvelle_num_arb)
      Fin si
    Si non
      Si( longueur(list_T) < max) // Pas plein
        Ajouter nouvelle numéro d'arbre dans liste_T(Meilleur_nbr_arb)
      Si non // la liste est plein
        vider la liste avec le principe FIFO et Ajouter nouvelle numéro
        d'arbre dans liste_T(Meilleur_nbr_arb)
      Fin si
      Meilleur_nbr_arb = nouvelle_num_arb // mise a jour à le
      Meilleur_nbr_arb
    Fin si
  Fin si n
Fin si
```

Étape 3: Return Meilleur_nbr_arb et meilleure_précision

L'algorithme RT proposé est composé de 3 étapes, l'étape 1 commence par une initialisation, dans laquelle on choisit un nombre d'arbre initial $\text{num_arbre}(0)$ de manière aléatoire. Par la suite, on applique la fonction objectif qui appelée 'appliquer fonction AD' où on évalue la précision de la classification ECG en fonction du nombre d'arbres $\text{num_arbre}(0)$ choisi.

L'algorithme RT est basée son le principe d'une liste tabou (list_T) pour effectuer l'évolution des itérations de RT. On initialise la list_T par un ensemble vide. Dans l'étape 2, on effectue une boucle qui se termine lorsqu'un critère d'arrêt est satisfait.

Par la suite, RT propose de générer une solution de voisinage, qui est un nouveau nombre d'arbres généré depuis le nombre ancien (la solution précédente/initiale), RT exige que ce nouveau nombre (cette nouvelle solution) ne soit pas dans une liste appelée la liste tabou. Cette solution voisine est évaluée en appliquant la fonction objectif qui est un appel à la précision du modèle basée sur le nouveau nombre d'arbres. Si cette solution voisine est meilleure que la précédente, alors elle choisit pour remplacer l'ancienne, et on ajoute l'ancien nombre d'arbres dans la liste tabou $\text{liste_T}(\text{Meilleur_nbr_arb})$, pour qu'il ne soit pas évalué ultérieurement dans le cas où la liste tabou est vide, sinon nous le déchargeons de cette liste avec le principe FIFO.

Au contraire, on ajoute le nouveau nombre d'arbre dans la liste $\text{liste_T}(\text{nouvelle_num_arb})$, si l'ancien nombre d'arbres est considéré comme le meilleur.

Pour un arbre de plus petite taille, l'ancien nombre d'arbre est ajouté a la liste tabou $\text{liste_T}(\text{Meilleur_nbr_arb})$ dans le cas où le taux atteint par le nouveau nombre est identique à celui trouvé par l'ancien nombre et aussi lorsque l'ancien nombre soit supérieur au nombre courant (le nouveau). Comme indiqué dans l'algorithme.

Nous pouvons résumer ces étapes dans le schéma représentatif suivant :

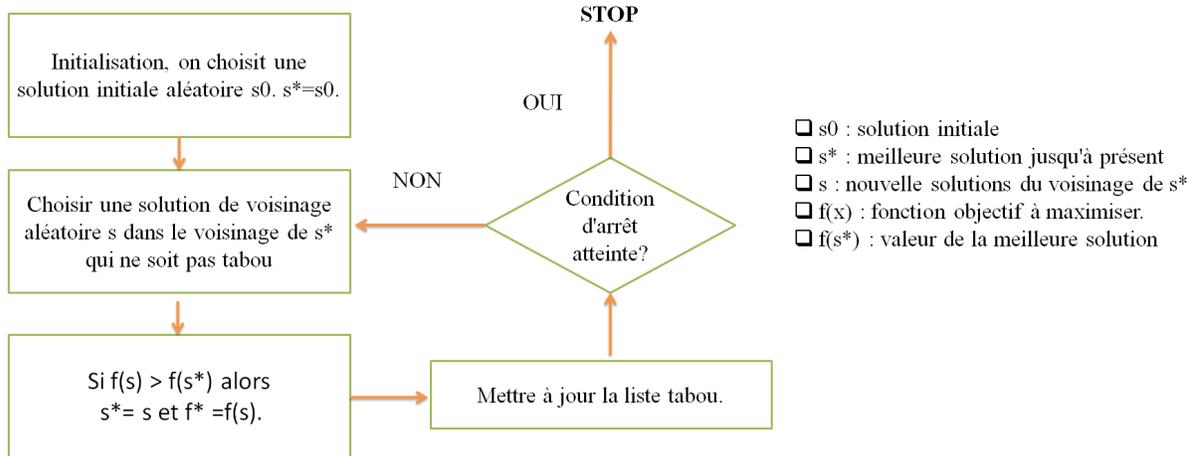


FIGURE 3.7 – L’algorithme Recherche Tabou pour l’optimisation de l’arbre de décision.

3.3.6 L’évaluation des solutions obtenues par RT

Pour évaluer le modèle trouvé, on fait référence aux métriques suivantes :

a La précision du modèle

Appelée aussi le taux de reconnaissance, c’est la métrique utilisée dans notre algorithme pour effectuer l’évaluation des solutions obtenues par RT.

Elle représente le rapport entre le nombre de donnée correctement classées et le nombre total des données testées. L’équation suivante illustre la formule utilisée [40].

$$P = \frac{1}{N} \sum_{i=1}^N L(y_i, \hat{f}(x_i))$$

avec :

$$L = \begin{cases} 1 & \text{si } y_i = \hat{f}(x_i) \\ 0 & \text{sinon} \end{cases}$$

b Matrice de confusion [40]

La mesure précédente donne le taux d’erreurs commises par le modèle appris mais ne donne aucune information sur la nature de ces erreurs. Dans la plus part des cas d’application, il est très important de connaître la nature des erreurs commises, dans notre travail le modèle appris a une finesse médicale, on considère qu’un sujet (un

patient) ne possède pas un problème de rythme cardiaque alors qu'il l'est, est plus grave qu'on considère quelqu'un comme alors qu'il ne l'est pas.

La matrice de confusion est une matrice qui rassemble en lignes les observations (y) et en colonnes les prédictions $f(x)$. Les éléments de la matrice représentent le nombre d'exemples correspondants à chaque cas.

Trois autres mesures sont utilisées dans la littérature dont les valeurs sont calculées dans la matrice de confusion, à savoir

- **La sensibilité S_v** : représente le rapport entre les observations positives correctement prédites et le nombre des observations positives.

$$S_v = \frac{CP}{CP + FN}$$

- **La spécificité S_p** : représente le rapport entre les observations négatives correctement prédites et le nombre total des observations négatives.

$$S_p = \frac{CN}{CN + FP}$$

- **La moyenne harmonique** : calculée à base de la sensibilité et la spécificité.

$$\text{Moyenne harmonique} = \frac{2 \times S_v \times S_p}{S_v + S_p}$$

3.3.7 Visualisation, Interprétation, Utilisation

Nous pouvons utiliser notre modèle qui aide à la décision pour une détection précoce et d'un suivi d'une maladie ou une anomalie de rythme cardiaque, ceci est obtenu après l'apprentissage, dont la manière d'utilisation est la suivante :

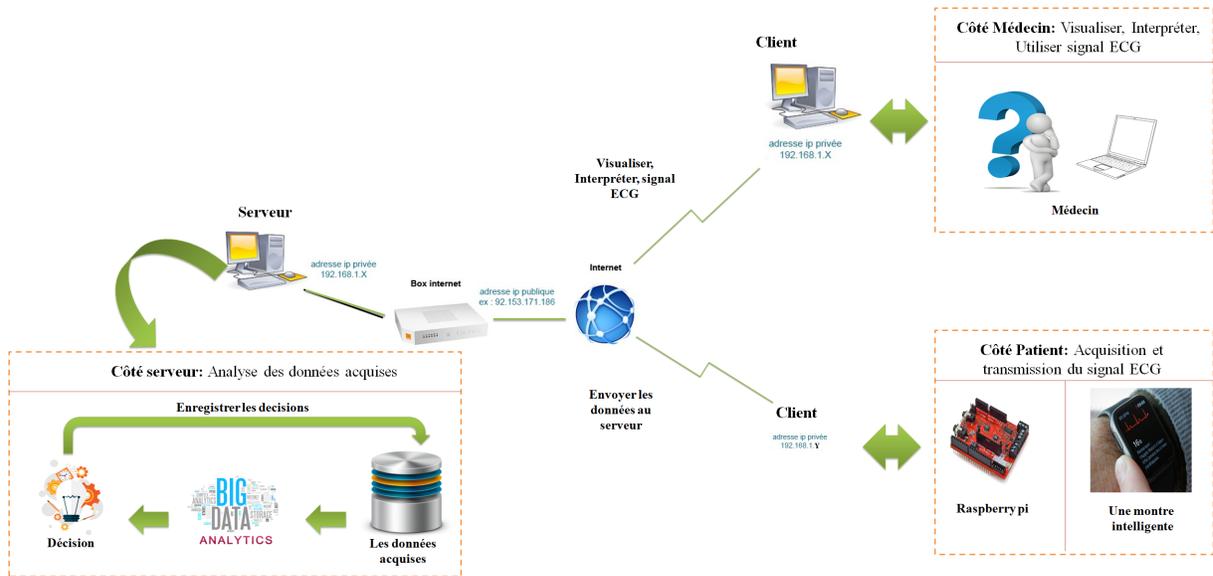


FIGURE 3.8 – Utilisation du modèle.

Nous acquérons le signal ECG, à partir de la plate-forme e-health à base d'un Raspberry PI3 ou bien une montre intelligente qui sera transmis au serveur de traitement par Internet, dans ce serveur toutes les données sont stockées en traitant les signaux et retraitant les décisions qui conviennent. Le médecin peut visualiser et interpréter, les signaux acquis et stockés dans le serveur.

En cas d'urgence suite à une détection d'anomalie cardiaque, nous envoyons une alerte au médecin avec le nom du patient, afin d'intervenir. Une autre alerte est aussi envoyée à la famille du patient afin de surveiller son état, jusqu'à l'arrivée de l'ambulance qui reçoit aussi une alerte indiquant la position géographique du patient, comme indiqué dans la figure suivant :

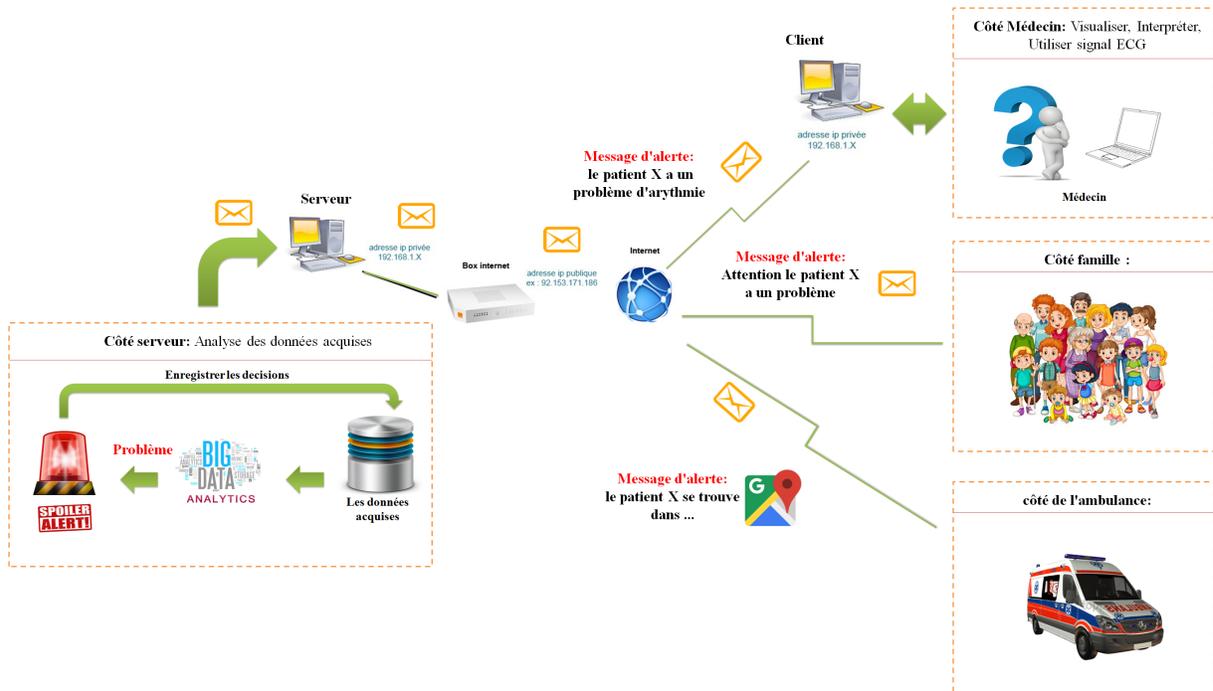


FIGURE 3.9 – Le système proposé en cas d’urgence suite à une détection d’anomalie cardiaque.

3.4 Conclusion

Dans ce chapitre, nous avons présenté la conception de notre système, où nous avons détaillé des différentes étapes de cette conception à savoir la collecte de données de santé, le prétraitement, la sauvegarde des données, la sélection des caractéristiques et l’entraînement du modèle avec l’arbres de décision. Nous avons aussi expliqué notre proposition qui est l’amélioration de la méthode de l’Arbres de décision pour optimiser le choix du nombre des arbres dans le but d’améliorer la précision de la classification.

Dans le chapitre suivant, nous allons décrire l’implémentation de notre système, c’est à dire la mise en œuvre des étapes expliquées dans le chapitre conceptuel, ainsi que les résultats obtenus et leurs discussions.

Chapitre 4 :

Etude expérimentale et résultats

Chapitre 4

Etude expérimentale et résultats

4.1 Introduction

Dans le chapitre précédent, nous avons présenté notre méthode proposée au sein de la conception de notre système afin de classifier les battements ECG basé sur le modèle arbre de décision, avec un nombre optimal d'arbres où nous avons utilisé l'algorithme RT afin de trouver un taux très élevé de classification. Par ailleurs, nous avons détaillé les différentes étapes de cette conception pour atteindre la phase de prise de décision.

Dans ce chapitre, nous allons essayer d'appliquer notre approche en mettant en œuvre plusieurs étapes, nous allons présenter l'environnement de travail, le langage de programmation et les outils que nous avons utilisés pour construire ce système. Par la suite, nous allons expliquer toutes les expérimentations conduites sur la méthode proposée ainsi que les résultats obtenus.

4.2 Outils et langages de développement

Durant l'étape de réalisation de ce projet, il a été nécessaire de recourir à certains domaines de la programmation et nous avons eu l'occasion de nous familiariser avec de diverses techniques et outils logiciels de développement qui sont présentés ci-dessous :

4.2.1 Java



C'est un langage de programmation orienté objet, développé par Sun Microsystems. Il permet de créer des logiciels compatibles avec de nombreux systèmes d'exploitation (Windows, Linux, Macintosh, Solaris). Java donne aussi la possibilité de développer des programmes pour téléphones portables et assistants personnels. Enfin, ce langage peut-être utilisé sur internet pour des petites applications intégrées à la page web (applet) ou encore comme langage serveur (jsp) [41].

4.2.2 Eclipse



L'IDE Eclipse est un environnement de développement intégré libre (le terme Eclipse désigne également le projet correspondant, lancé par IBM) extensible, universel et polyvalent, permettant potentiellement de créer des projets de développement mettant en œuvre n'importe quel langage de programmation. L'IDE Eclipse est principalement écrit en Java (à l'aide de la bibliothèque graphique SWT d'IBM), et ce langage, grâce à des bibliothèques spécifiques, est également utilisé pour écrire des extensions [42]. Nous avons ajouté à l'éclipse la bibliothèque JDBC.

- **JDBC** : est l'acronyme de **Java DataBase Connectivity** qui désigne une API pour permettre un accès aux bases de données avec Java [43]. Normalement, il s'agit d'une base de données relationnelle, et des pilotes JDBC sont disponibles pour tous les systèmes connus de bases de données relationnelles [44].

4.2.3 Le langage R



R est langage et environnement librement «GNU S», disponible pour l'informatique et les graphiques statistiques et fournissant une grande variété de techniques statistiques et graphiques : modélisation linéaire et non linéaire, tests statistiques, analyse de séries chronologiques, classification, regroupement... etc [45].

4.2.4 RStudio



RStudio est un outil apparu récemment et qui vient combler un manque dans la collection des outils associés à R : il s’agit d’un environnement de développement intégré (IDE) fonctionnel, libre, gratuit et multiplate-forme [46].

Pour notre travail, nous avons installé deux paquets (packages) :

- **RandomForest** : Le package "RandomForest" implémente l’algorithme de l’arbre de décision (basé sur le code Fortran original de Breiman et Cutler) utilisé pour créer et analyser les arbres de décisions, la classification et la régression. Il peut également être utilisé en mode non supervisé pour évaluer les proximités entre des points de données [47].
- **Caret** : Le package “caret” (Classification And REgression Training) est une librairie pour R. Elle couvre une large fraction de la pratique de l’analyse prédictive (classement et régression). Un peu à la manière de “scikit-learn” [48] pour Python, il intègre dans un ensemble cohérent les étapes clés de la modélisation : préparation des données, sélection, apprentissage et évaluation. La standardisation des prototypes des fonctions d’apprentissage et de prédiction notamment permet de simplifier notre code, facilitant les tâches d’optimisation et de comparaison des modèles [49].

4.2.5 Xampp



XAMPP est un ensemble de logiciels permettant de mettre en place facilement un serveur Web et un serveur FTP. Il s’agit d’une distribution de logiciels libres (X Apache MySQL Perl PHP) offrant une bonne souplesse d’utilisation, réputée pour son installation simple et rapide. Ainsi, il est à la portée d’un grand nombre de personnes puisqu’il ne requiert pas de connaissances particulières et fonctionne, de plus, sur les systèmes d’exploitation les plus répandus [50].

4.3 Le système développé

4.3.1 L'ensemble des données utilisées

Comme nous avons expliqué dans le chapitre précédent, les données provenaient de deux sources :

4.3.1.1 Plate-forme e-Health

La plate-forme e-Health Sensor Shield (abrégée e-HP) est une plate-forme de capture de données de santé qui peut être utilisée avec un Arduino ou bien avec un Raspberry Pi. Dans notre travail, nous utilisons la version V2.00 de la plate-forme e-Health de marque Cooking hackss équipée d'un Raspberry Pi3 B. Pour brancher la plate-forme e-HP dans le Raspberry, nous devons utiliser le pont de connexion Raspberry Pi vers "Arduino Shields", qui permet d'utiliser l'e-HP avec le Raspberry Pi.

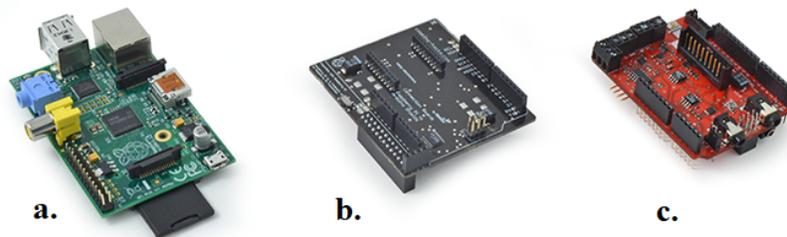


FIGURE 4.2 – a.Raspberry Pi3 B, b.Pont de connexion entre la e-HP et le raspberry Pi, c. la e-HP.

Afin que notre système puisse extraire les signaux ECG, nous suivons les étapes suivantes :

a Connexion du Raspberry Pi à l'Arduino shields

Il est important à configurer le RPi3 pour qu'il soit prêt à communiquer avec Arduino shields avec le Raspberry Pi. Pour cela, premièrement on doit mettre à jour le système d'exploitation avec les derniers correctifs à l'aide de ces commandes :

```
sudo apt-get update
sudo apt-get upgrade
```

Puis, on télécharge et on installe la bibliothèque ArduPi [51] avec la commande suivante :

```
wget http://www.cooking-hacks.com/media/cooking/images/documentation/raspberry_arduino_shield/raspberrypi2.zip && unzip raspberrypi2.zip && cd cooking/arduPi && chmod +x install_arduPi && ./install_arduPi && rm install_arduPi && cd ../
```

b Installation de la bibliothèque e-Health

La bibliothèque e-Health inclut des fonctions de haut niveau pour une gestion facile. La bibliothèque peut être téléchargée à partir de [52]. Il convient de mentionner que la bibliothèque de santé en ligne de Raspberry Pi nécessite la bibliothèque ArduPi et les deux bibliothèques doivent se trouver dans le même chemin.

c Branchement des capteurs ECG

Quand nous arrivons à cette étape, la plate-forme est prête à l'acquisition, alors nous branchons des capteurs ECG sur la plate-forme, puis nous assemblons notre système.



FIGURE 4.3 – Capteurs ECG connectés à la plate-forme.

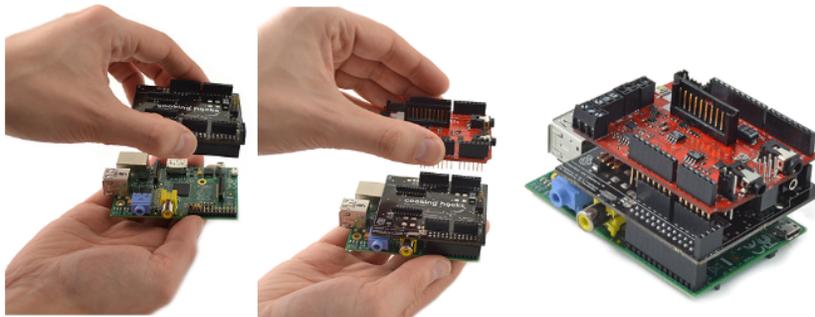


FIGURE 4.4 – Regroupement de système.

d Acquisition de données avec la plate-forme

En atteignant ce stade, notre système est prêt pour l'étape d'acquisition. Le site Web cooking-hacks [51] nous donne un programme en C à cette fin, ce programme est sujet à changement selon le besoin.

La compilation du programme fourni par le site Web cooking-hacks peut être réalisée de la manière suivant :

```
G++ -pthread -lrt ECGToFile.cpp arduPi.cpp eHealth.cpp -o ECGToFile
Sudo ./ECGToFile
```

Où ECGToFile.cpp c'est le nom du fichier qui contient le programme d'acquisition et ECGToFile le fichier résultant.

Il reste à faire le programme d'acquisition en marche de manière automatique dans le Raspberry Pi. Pour cela il faut modifier le fichier /etc/rc.local, en ajoutant le chemin et la commande permettant d'exécuter le programme d'acquisition comme suit :

```
Cd /home/pi/ Healthcare
Sudo ./ECGToFile
```

Ici, tout est considéré comme "prêt", il ne reste que placer les électrodes ECG sur le patient et fonctionner le système, comme indiqué ci-dessous :

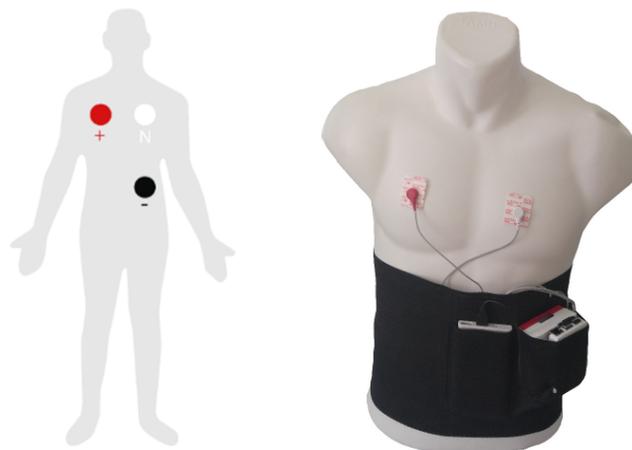


FIGURE 4.5 – La plate-forme et les électrodes ECG sont placés sur le patient.

Le fichier résultant contient des valeurs réelles comprises dans l'intervalle $[-5v, +5v]$ représentant l'activité électrique du cœur.

e Sauvegarde des données

Comme nous l'avons expliqué dans le chapitre précédent, et compte tenu de l'importance de sauvegarder des données pour le médecin, et en raison de la petite taille de Raspberry Pi 3 pour la sauvegarde, il est nécessaire d'avoir un serveur de grande taille qui représente un entrepôt de données afin de sauvegarder ces données de taille massive (de type Big data) et aussi de les analyser ultérieurement.

Pour cette transmission de données (depuis le RaspBerry Pi 3 et le serveur), nous allons utiliser les sockets pour mettre en œuvre cette étape, nous les avons représenté sous forme d'une architecture (Client-Serveur).

a Patient (Client)

A chaque minute, les données acquises qui correspondent aux signaux cardiaques et qui sont enregistrées dans des fichiers sont envoyées au serveur et puis elles sont supprimées du Raspberry. Le programme ci-dessous représente une partie du programme global de cette phase.

```
import java.io.IOException;
import java.net.UnknownHostException;
import java.util.Timer;
import java.util.TimerTask;
public class RepetAction {
    Timer t;

    public RepetAction() {
        t = new Timer();
        // Lancer le programme à chaque 5 min
        t.schedule(new MonAction(), 0, (60)*1000);
    }
    class MonAction extends TimerTask {
        // Contrôle d'arrêter
        int nbrRepetitions = 1000;
        public void run() {
            if (nbrRepetitions > 1000) {
                try {
                    //crée un objet à partir de la classe "Patient" pour envoyer les fichiers
                    patient patient=new patient();
                    System.out.println("Ça marche bien!");
                } catch (UnknownHostException e) {
                    e.printStackTrace();
                } catch (IOException e) {
                    e.printStackTrace();
                }
            }
            nbrRepetitions--;
        } else {
            System.out.println("Terminé!");
            t.cancel();
        }
    }
}

public static void main(String[] args){
    RepetAction repetAction = new RepetAction();
}
```

```
public class patient {
    public patient () throws UnknownHostException, IOException
    {
        //un objet Socket qui tente de se connecter sur l'ip et sur le numéro de port 15123
        Socket socket = new Socket("192.168.1.10",15123);
        File file = new File ("/home/pi/cooking/");
        File[] transferFile = file.listFiles();

        // envoyer le nom de fichier à server
        DataOutputStream dout=new DataOutputStream(socket.getOutputStream());
        String nom_fichier=transferFile[1].getName();
        dout.writeUTF(nom_fichier);
        dout.flush();

        byte [] bytearray = new byte [(int)transferFile[1].length()];
        FileInputStream fin = new FileInputStream(transferFile[1]);
        BufferedInputStream bin = new BufferedInputStream(fin);
        bin.read(bytearray,0,bytearray.length);
        OutputStream os = socket.getOutputStream();
        os.write(bytearray,0,bytearray.length);
        os.flush();
        socket.close();

        input.close();
        transferFile[1].delete();
    }
}
```

FIGURE 4.6 – Le code illustratif d'envoi des données depuis le RaspBerry vers le serveur.

b Serveur

Le serveur est constamment en réception des fichiers du client 24h sur 24h, il les enregistre dans la base de données implémentée par le logiciel « xampp » après avoir assuré une connexion à distance à cette base de données par le protocole JBDC, comme indiqué ci-dessous :

- **Connexion à distance à la base de données**

Pour autoriser l'accès à distance à la base de données même à partir d'un autre ordinateur, on doit d'abord permettre les connexions depuis les machines qui se trouvent dans le réseau local d'adressage 192.168.0.x. Dans le serveur appartient à un autre réseau, son adresse statique doit être introduite afin de permettre la connexion.

Ensuite, nous configurons les ordinateurs :

PC1 (Client):	PC2 (Serveur) :
<p>créer un compte dans xampp avec un nom d'utilisateur et son @IP et lui donner tous les privilèges.</p>	<p>créer un compte dans xampp avec le même nom d'utilisateur et @IP qui nous avons créée dans le PC1 et lui donner tous les privilèges.</p>
<pre>CREATE USER 'nom_d'utilisateur' '@'adress_ip' IDENTIFIED BY 'Mot_de_passe'; GRANT ALL PRIVILEGES ON * . * TO 'nom_d'utilisateur '@'adress_ip';</pre>	<pre>CREATE USER 'nom_d'utilisateur_pc1'@'adress_ip_pc1' IDENTIFIED BY 'Mot_de_passe_pc1'; GRANT ALL PRIVILEGES ON * . * TO 'nom_d'utilisateur_pc1'@'adress_ip_pc1';</pre>

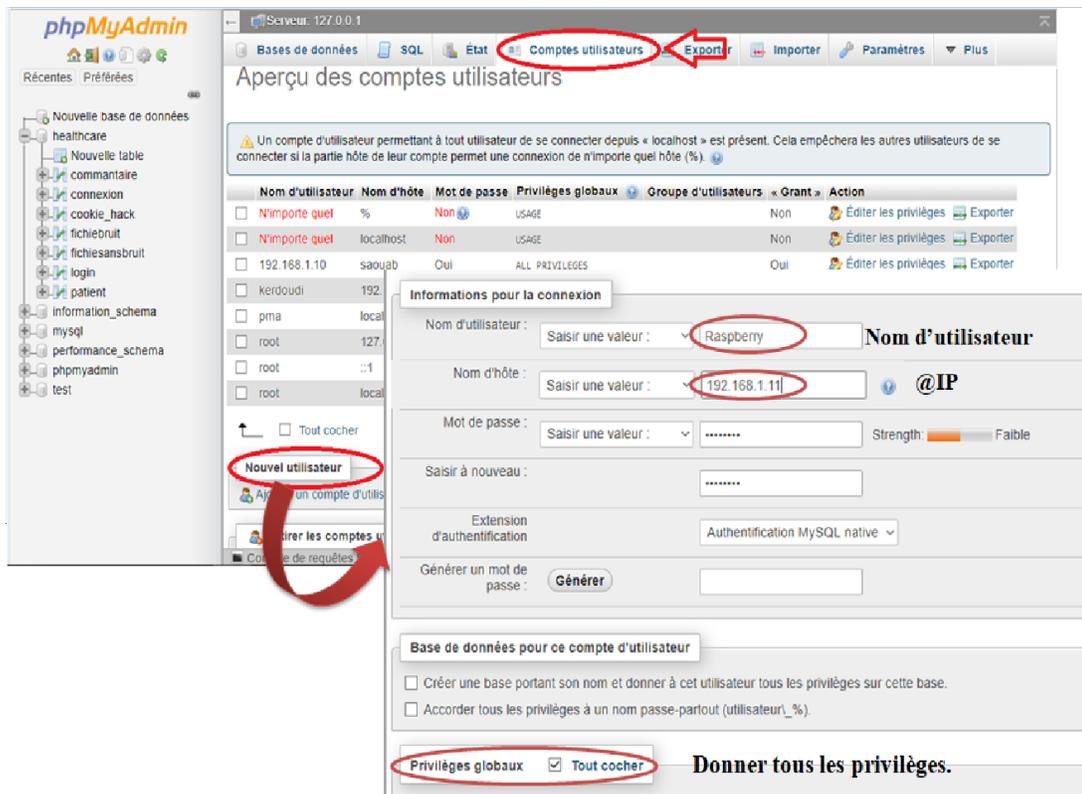


FIGURE 4.7 – Création d'un compte pour le PC1.

Nous répétons les mêmes étapes pour l'ordinateur 2 (le serveur) comme expliquer ci-dessous. Finalement, nous mettons ce code dans java après avoir su @IP de serveur :

```

import java.sql.Connection;
import java.sql.DriverManager;
import java.sql.ResultSet;
import java.sql.Statement;

import org.omg.IOP.CodecFactory;

public class ConnectionavecDB {
    private static final String ConnectionavecDB = null;
    private Connection con;
    private Statement st;
    private ResultSet rst;

    public ConnectionavecDB()
    {
        try {
            Class.forName("com.mysql.jdbc.Driver");
            con=DriverManager.getConnection("jdbc:mysql://192.168.1.10:3306/healthcare","Raspberry","kerdoudi");

            st=con.createStatement();
            System.out.println("la connexion valide");
        }catch(Exception ex){
            System.out.println("La connexion non valide car : "+ex);
        }
    }
}

```

Diagram annotations for Figure 4.8:

- Nom d'utilisateur de Pc1**: Points to the username "Raspberry" in the code.
- Mot de passe de Pc1**: Points to the password "kerdoudi" in the code.
- @IP de Pc2 (serveur)**: Points to the IP address "192.168.1.10" in the code.
- Le nom de la base de données**: Points to the database name "healthcare" in the code.

FIGURE 4.8 – La connexion à distance à la base de données.

Donc, nous pouvons maintenant nous connecter à distance à notre base de données qui est nommée «Healthcare» située dans le serveur.

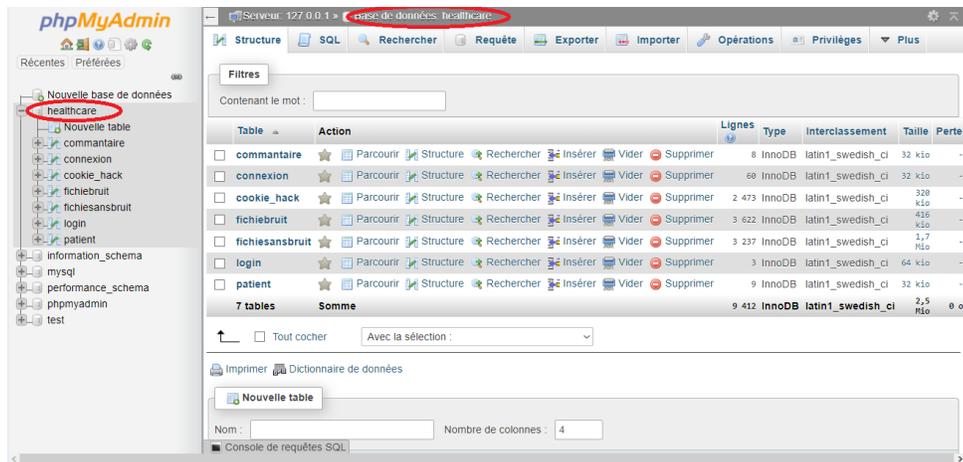


FIGURE 4.9 – La connexion à notre base de données.

- Réception des données du client

```

public server_cookie_hack() throws UnknownHostException, IOException{
    ServerSocket serverSocket = new ServerSocket(15123);
    InetAddress address = InetAddress.getLocalHost();
    String hostIP = address.getHostAddress();
    String hostName = address.getHostName();
    System.out.println("Le nom de serveur est : " + hostName + "\nIP: " + hostIP);
    System.out.println("Le serveur est à l'écoute du port : "+serverSocket.getLocalPort());

    while(true){
        int filesize=2022396;
        //contient les statistiques actuelles des octets lus sur le canal d'entrée, c'est-à-dire inputStream
        int bytesRead;
        //contient le nombre total d'octets lus.
        int currentTot = 0;

        Socket socket = serverSocket.accept();
        // La réception de id_patient
        DataInputStream patient=new DataInputStream(socket.getInputStream());
        int id_patient=(int)patient.readInt();
        System.out.println("Connexion acceptée, raspberry de patient qui à l'identifiant "+id_patient+" est connecté : " + socket);

        // La réception de id_m
        DataInputStream medecin=new DataInputStream(socket.getInputStream());
        int id_m=(int)medecin.readInt();

        // La réception de le nom de fichier
        DataInputStream dis=new DataInputStream(socket.getInputStream());
        String nom_fichier=(String)dis.readUTF();

        //Nous définissons une variable byteArray qui agira comme un tampon pour stocker des données temporaires
        byte [] bytearray = new byte [filesize];
        //un Objet InputStream qui nous aidera à collecter toutes les informations transmises au canal d'entrée du serveur c'est-à-dire aux fichiers ou
        //même aux messages transférés au serveur.
        InputStream is = socket.getInputStream();
        // l'objet FileOutputStream qui pointe vers le fichier qui sera rempli avec les données copiées à partir du fichier du patient.
        FileOutputStream fos = new FileOutputStream("D:/healthcare/cookie_hack/cookie_hack_l'acquisition/"+nom_fichier+".txt");
        // BufferedOutputStream nous aide à écrire des données dans le fichier de sortie via un tableau d'octets.
        BufferedOutputStream bos = new BufferedOutputStream(fos);

        //Nous lisons les données du flux d'entrée(InputStream) en utilisant la méthode de lecture de l'objet(read) Les données lues sur
        //le canal d'entrée sont stockées dans le tableau d'octets(bytearray)
        bytesRead = is.read(bytearray,0,bytearray.length);

        // initialiser currentTot sur le nombre d'octets lus.
        currentTot = bytesRead;
        System.out.println("Envoi de fichiers...");
        do {
            bytesRead = is.read(bytearray, currentTot, (bytearray.length-currentTot));
            if(bytesRead >= 0)
                currentTot += bytesRead;
        } while (bytesRead > -1);
        // Lorsque le paramètre bytesRead est égal à -1, c'est-à-dire qu'il ne reste aucune donnée sur le flux d'entrée

        //Après cela, nous écrivons les octets finalement lus sur le fichier
        bos.write(bytearray, 0 , currentTot);

        // puis fermons le flux.
        bos.flush();
        bos.close();
        socket.close();
        System.out.println("Transfert de fichier terminé");

        //créer un objet de la class manipulation pour enregistrer les donner dans la base de donner
        cookie_hack cookie_hack=new cookie_hack(id_patient,id_m);
    }
}

public static void main(String[] args) throws UnknownHostException, IOException{
    server_cookie_hack s=new server_cookie_hack();
}

```

FIGURE 4.10 – Réception des données du client.

- Sauvegarde de données

La fonction suivante enregistre les données acquises depuis la plateforme :

```

public cookie_hack(int id_patient, int id_m){
// ----- connecter java avec sql server -----
ConnectionavecDB obj=new ConnectionavecDB();
try{
// ----- lire des fichier txt qui contient l'electrocardiogramme ECG pour insérer dans la base de donnée -----
File file = new File("D:/heathcare/cookie_hack/cookie_hack_1/acquisition");

File[] files = file.listFiles();
if (files != null) {
int k;
for (int i = 0; i < files.length; i++) {
Scanner input = new Scanner(files[i]);
String s=files[i].getName();
Path source = FileSystems.getDefault().getPath("D:/heathcare/cookie_hack/cookie_hack_1/acquisition/"+s);
Path destination = FileSystems.getDefault().getPath("D:/heathcare/cookie_hack/Analyser_cookie_hack/"+s);
k=i+1;

if (files[i].exists()) {
System.out.println("fichier "+k);

//récuperer la date du jour + heure + min + sec
String format = "yyyy-MM-dd HH:mm:ss";

java.text.SimpleDateFormat formater = new java.text.SimpleDateFormat( format );
java.util.Date date = new java.util.Date();

//System.out.println( formater.format( date ) );
String date_time = formater.format( date );

String line;
// ----- les valeur de l'electrocardiogramme ECG -----
while (input.hasNext()) {
line = input.nextLine();
String[] Splited = line.split("\\s+");
double ECG = Double.parseDouble(Splited[0]);
//----- inserer un element dans la base de donnée sql server -----
obj.setData_cookie_hack(id_patient,id_m,date_time,ECG);
} k++;

input.close();
// couper le fichier manipuler à une autre répertoire
try {
Files.move(source, destination, StandardCopyOption.REPLACE_EXISTING);
} catch (IOException e) {
System.err.println(e);
}

//----- afficher la base de donner la table ecg -----
obj.getData_cookie_hack();
} else {
System.out.println("Le fichier n'existe pas !...");
}
}}
} catch (Exception e) {
e.printStackTrace();
}
}

public static void main(String[] args) throws UnknownHostException, IOException{
cookie_hack s=new cookie_hack(id_patient,id_m);
}
}

```

FIGURE 4.11 – Enregistrement des données acquises depuis la plate-forme.

Les données sont stockées dans la table «**cookie_hack**» comme indiqué ci-dessous :

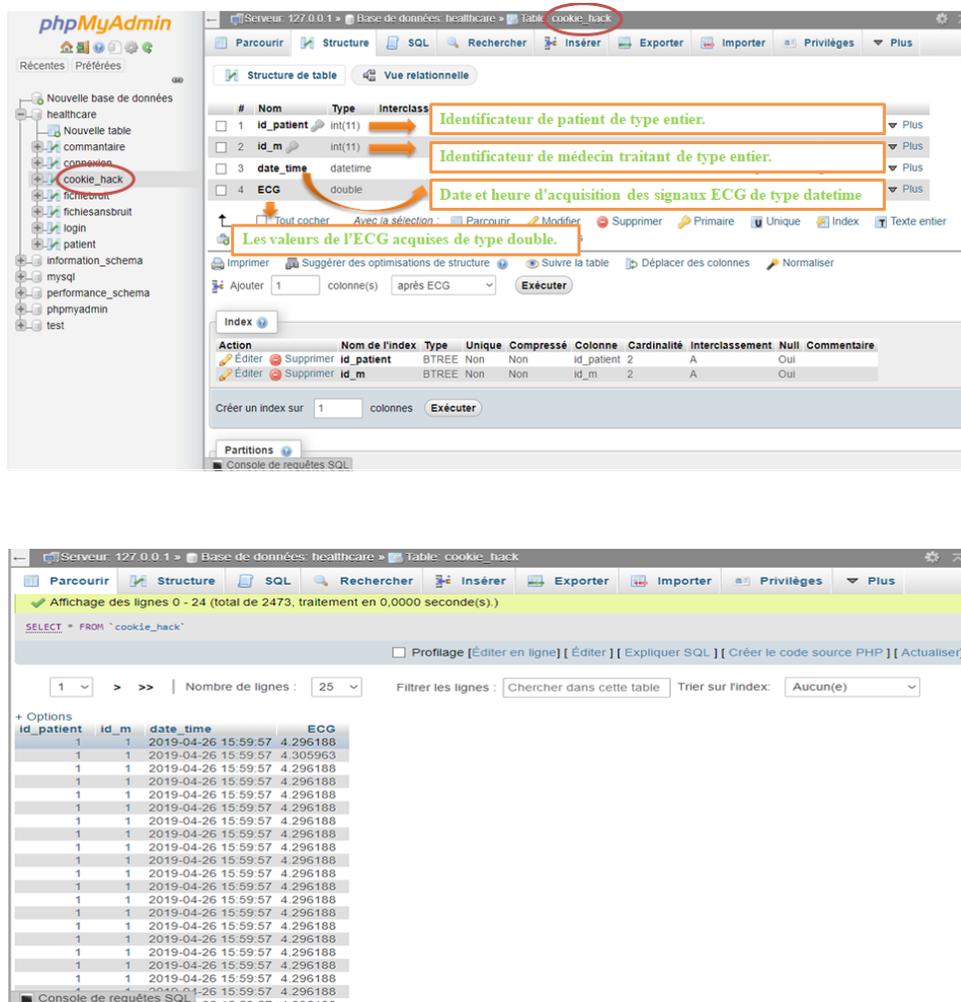


FIGURE 4.12 – Les types et les valeurs des champs de la table de cookie hack.

- **Numérisation des données**

Les données acquises sont exprimées en millivolts, ce qui signifie qu'elles se présentent sous une forme physique et analogique. Toutefois, pour permettre la manipulation de ces données avec les outils appropriés, un processus de numérisation de ces données est nécessaire. Nous utilisons une fonction avec l'outil R permettant la numérisation des données acquises, basée sur la transformation des valeurs en représentation non-flottante avec possibilité de récupérer la représentation d'origine, comme indiqué dans le tableau 4.1 [52].

Valeurs physiques	Représentation numérique
1.388074	2776
1.309873	2620
1.319648	2639
1.466276	2933
1.715543	3431
2.028348	4057
2.380254	4761
2.820137	5640
3.313783	6628
3.602150	7204

TABLE 4.1 – Valeurs physiques en représentation numérique

La fonction utilisée dans la conversion est la suivante [52] :

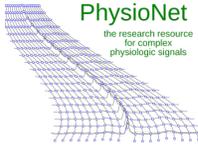
```
#From physical values to digital representation function
Digitization = function(){
  #the original file name and path
  OrigFile = file.choose()
  #Read the original (physical data)
  DigitalECG = read.table(OrigFile, header = FALSE, dec = ".", col.names = c("ECG1"))
  # DigitalECG = myData$ECG1
  #Round to 4 digits
  DigitalECG$ECG1 = round(DigitalECG$ECG1, digits = 3)
  #Eliminate the float
  DigitalECG$ECG1 = DigitalECG$ECG1 * 1000
  #Make the values in the multiplication of 5 form|
  DigitalECG$ECG1 = round(DigitalECG$ECG1/5)*5
  #Divide all values by 5
  DigitalECG$ECG1 = DigitalECG$ECG1 / 5

  #The output digital values file
  OutputFile = paste(dirname(OrigFile),"/Dig_",basename(OrigFile), sep = "")

  #Save the results to external file
  write.table(DigitalECG,OutputFile,row.names = FALSE, quote = FALSE)
}
Digitization()
```

FIGURE 4.13 – La fonction de numérisation

4.3.1.2 Site web Physionet



Comme mentionné au chapitre 3, le site web Physionet offre un accès Web gratuit à de nombreux signaux physiologiques enregistrés dans des bases de données, la plupart d’entre eux étant dédiés à l’étude et à l’analyse du signal ECG [31] [32].

— PhysioBank de Physionet

a Bases de données utilisées

Dans notre projet nous avons utilisé les bases de données [53] suivantes :

- **MIT-BIH Arrhythmia Database** : Cette collection de 48 enregistrements de l’anomalie d’arythmie d’une demi-heure entièrement annotés est disponible dans [54]. Elle est la base de données la plus utilisée dans la littérature.
- **The MIT-BIH Supraventricular Arrhythmia Database [55]** : Cette base de données comprend 78 enregistrements ECG d’une demi-heure choisis pour compléter les exemples de la base de données sur les arythmies MIT-BIH.
- **The Long-Term AF Database** : C’est un ensemble de 84 enregistrements ECG à long terme (sur 24 heures) de sujets atteints de fibrillation auriculaire (FA) paroxystique ou prolongée. Chaque enregistrement contient deux signaux ECG et deux ensembles d’annotations. L’ensemble initial comprend des marqueurs non audités produits par un détecteur QRS automatisé, avec des annotations manuelles des terminaisons d’épisodes de FA avec une durée d’au moins une minute. Le nouvel ensemble contient des annotations de type de battement et de rythme révisées manuellement [56].

b Les fichiers des bases de données

- **Les fichiers de signaux (.dat)** : ce sont des fichiers binaires contenant des échantillons de signaux numérisés. Ceux-ci stockent les formes d’onde, mais ils ne peuvent pas être interprétés correctement sans leurs fichiers d’en-tête correspondants [57].

Elapsed time (seconds)	ECG1 (mV)	ECG2 (mV)
0.000	0.145	-0.635
0.008	0.125	-0.745
0.016	0.145	-0.735
0.023	0.145	-0.695
0.031	0.145	-0.685
0.039	0.145	-0.685
0.047	0.125	-0.675
0.055	0.135	-0.675
0.062	0.145	-0.685
0.070	0.145	-0.685
0.078	0.125	-0.665
0.086	0.125	-0.655
0.094	0.125	-0.655
0.102	0.095	-0.705
0.109	0.095	-0.685
0.117	0.095	-0.675

FIGURE 4.14 – Exemple de fichier (.dat).

- **Les fichiers d'en-tête (.hea)** : sont de courts fichiers texte décrivant le contenu des fichiers de signaux associés [57].

```
14046 2 128 10828800
14046.dat 212 0 12 0 32 -13462 0 ECG1
14046.dat 212 0 12 0 -58 -28804 0 ECG2
#Age: 46 Sex: M
```

FIGURE 4.15 – Exemple de fichier (.hea).

- **Annotation (.atr)** : ces fichiers sont des fichiers binaires contenant des annotations qui sont des étiquettes font généralement référence à des échantillons spécifiques dans les fichiers de signaux associés [57].

L'heure de la crête R.			L'annotation
1:30.183	32466	N	
1:30.928	32734	N	
1:31.694	33010	N	
1:32.492	33297	N	
1:33.319	33595	N	
1:34.119	33883	N	
1:34.606	34058	V	
1:35.044	34216	+	
1:35.486	34375	.	
1:36.297	34667	/	
1:37.083	34950	/	
1:37.883	35238	/	
1:38.711	35536	/	
1:39.578	35848	/	
1:40.425	36153	f	
1:41.256	36452	f	
1:42.097	36755	f	
1:42.897	37043	f	
1:43.683	37326	f	

FIGURE 4.16 – Exemple de fichier (.atr).

La plupart des bases de données PhysioBank utilisent l'ensemble standard de codes d'annotation comme décrit ci-dessous :

Code	Sens	La description
N	NORMAL	Normal beat
L	LBBB	Left bundle branch block beat
R	RBBB	Right bundle branch block beat
B	BBB	Bundle branch block beat (unspecified)
A	APC	Atrial premature beat
a	ABERR	Aberrated atrial premature beat
J	NPC	Nodal (junctional) premature beat
S	SVPB	Supraventricular premature or ectopic beat (atrial or nodal)
V	PVC	Premature ventricular contraction
r	RONT	R-on-T premature ventricular contraction
F	FUSION	Fusion of ventricular and normal beat
e	AESC	Atrial escape beat
j	NESC	Nodal (junctional) escape beat
n	SVESC	Supraventricular escape beat (atrial or nodal) [1]
E	VESC	Ventricular escape beat
/	PACE	Paced beat
f	PFUS	Fusion of paced and normal beat
Q	UNKNOWN	Unclassifiable beat
?	LEARN	Beat not classified during learning

TABLE 4.2 – Les annotations des battements.

— **PhysioToolkit de Physionet**

Les fichiers sur le site web Physionet doivent être convertis en fichiers texte (.txt) pour utilisation, pour cela nous avons téléchargé et installé les packages logiciels WFDB (Wave Form Data Base) de [58].

Le tableau suivant récapitule les commandes utilisées pour la conversion des fichiers de bases de données Physionet, ceci peut être effectué via l’outil cygwin [52] :

La commande	Description	Exemple d'utilisation
rdsamp	Pour lire et convertir les fichiers originaux(.dat) en formulaires physiques numériques et les exporter au format texte(.txt).	rdsamp -r SourceF -p -v >DestinF.txt
rdann	Pour lire le fichier d'annotations et les exporter au format texte (.txt).	rdann -r SourceF -a atr >DestinF.txt
wrsamp	Convertir le signal ECG acquis (par notreplate-forme) en un fichier compatible (.dat).	wrsamp -F freq -i SourceF.csv -o DestinF -s
sqrs	Pour générer l'annotation du signal ECG acquis(par notre plate-forme) et l'exporter au format texte (.txt).	sqrs -r SourceF -s 0
wfdbdesc	Pour lire les notes dans le fichier (.hea).	wfdbdesc SourceF >DestinF.txt

TABLE 4.3 – Les commandes WFDB utilisés [28].

4.3.2 Prétraitement et extraction des caractéristiques

Pour prétraiter le signal ECG et extraire les fonctions nécessaires, nous avons utilisé la «bibliothèque Annotation C ++» de M. Chesnokov Yuriy. La bibliothèque C ++ utilisée est basée sur une analyse d'ondelettes et une application console permettant d'extraire des intervalles vitaux et des ondes à partir de données d'ECG (P, T, QRS, PQ, QT, RR, RRn), de battements ectopiques et de détection de bruit [59].

Nous avons considéré cette étape comme une boîte noire dont l'entrée est un fichier compatible Physionet (.dat) et le résultat est un fichier texte contenant les caractéristiques détectées et la position des ondes en fonction du fichier d'origine.

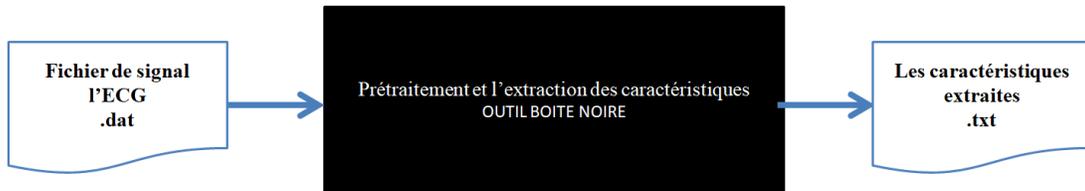


FIGURE 4.17 – Prétraitement et extraction des caractéristiques.

```

1 | leads: 2
2 |   sr: 250.00 Hz
3 |   bits: 12
4 |   UmV: 200
5 | length: 02:00:00.000
6 |
7 | getting QRS complexes... 7281 beats.
8 | getting P, T waves... done.
9 | processing time: 763 ms
10 |
11 |   38 00:00:00.152  N
12 |   42 00:00:00.168  q
13 |   50 00:00:00.200  R
14 |   63 00:00:00.252  )
15 |   85 00:00:00.380  (t
16 |  112 00:00:00.448  T
17 |  128 00:00:00.512  t)
18 |  226 00:00:00.904  (p
19 |  236 00:00:00.944  P
20 |  244 00:00:00.976  p)
21 |  258 00:00:01.032  N
22 |  269 00:00:01.076  R
23 |  280 00:00:01.120  )
24 |  314 00:00:01.256  (t
25 |  340 00:00:01.324  T
26 |  347 00:00:01.388  t)
27 |  461 00:00:01.844  (p
28 |  468 00:00:01.872  P
29 |  475 00:00:01.900  p)
30 |  490 00:00:01.960  N
31 |  502 00:00:02.008  R
32 |  515 00:00:02.060  )
33 |  547 00:00:02.188  (t
34 |  564 00:00:02.256  T
35 |  580 00:00:02.320  t)
  
```

FIGURE 4.18 – Le fichier des caractéristiques extraites (.txt).

Le type des caractéristiques détectées et extraites dans le fichier résultat est comme suivant :

Code	Sens
(p	Début de Pwave
P	Onde de pointe P
p)	offset de Pwave
N	N point
ECT	Ectopique de toute origine battement
Q / q	Q point
R / r	R peak
S / s	S point
)	J point
ARFCT	Artefact
NOISE	Bruit détecté
(t	Début de Twave
T	Onde de pointe T
t)	offset de Twave

TABLE 4.4 – Les caractéristiques détectées et extraites [52].

4.3.3 Sélection et calcul des caractéristiques

Nous avons combiné les fichiers résultants (les fichiers d'acquisition "figure 4.14", le fichier d'annotation « figure 4.16 » et le fichier des caractéristiques extraites « figure 4.18 »), voir la figure 4.19 ». Ensuite, il est nécessaire de calculer les valeurs de ces caractéristiques, par exemple la valeur de $P_{\text{ware}} = \text{le début de Pwave} - \text{l'offset de Pwave}$, et de les stockées dans une matrice implémentée en R, voir figure 4.20. Les lignes de la matrice représentent les battements et chaque colonne représente une caractéristique, la dernière colonne est dédiée au type de battement, qui sert de classe au processus d'apprentissage.

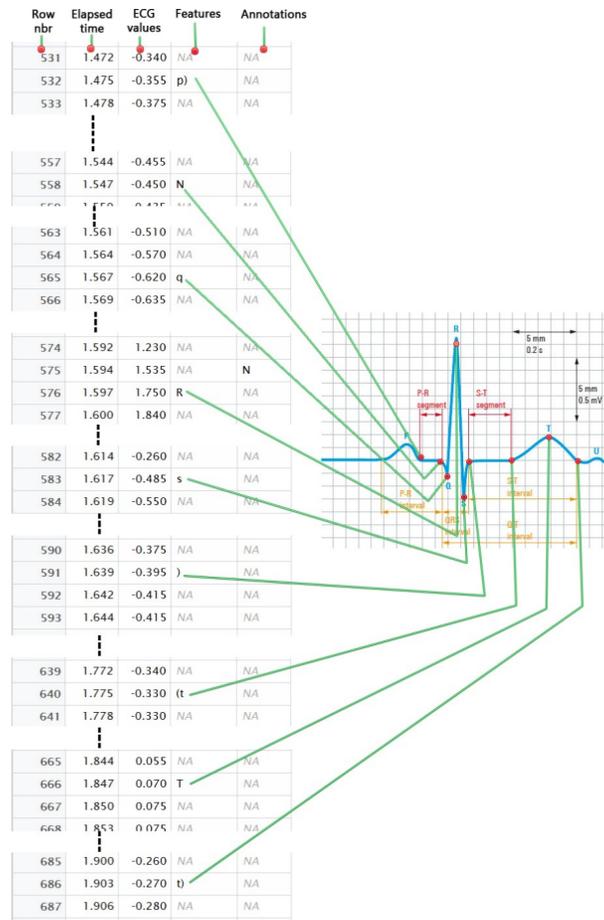


FIGURE 4.19 – Le processus de correspondance [52].

	PWave	TWave	QRSComplex	PRInterval	QTInterval	PRsegment	STsegment	JTInterval	STInterval	PTInterval1	PTInterval2	ISECT	RRinterBef	RRinterAft	BeatClass
20	0.097	0.133	0.069	0.164	0.360	0.067	0.124	0.264	0.290	0.484	0.280	FALSE	0.794	0.797	N
31	0.086	0.133	0.084	0.163	0.360	0.077	0.124	0.264	0.290	0.484	0.280	FALSE	0.814	0.827	N
87	0.086	0.133	0.086	0.156	0.360	0.070	0.124	0.264	0.290	0.484	0.280	FALSE	0.800	0.789	N
126	0.086	0.131	0.086	0.164	0.342	0.078	0.139	0.270	0.289	0.520	0.303	FALSE	0.816	0.820	N
202	0.081	0.133	0.083	0.164	0.360	0.083	0.124	0.264	0.290	0.484	0.280	FALSE	0.809	0.827	N
277	0.089	0.133	0.083	0.173	0.360	0.084	0.124	0.264	0.290	0.484	0.280	FALSE	0.841	0.831	N
347	0.083	0.133	0.083	-0.631	0.360	-0.714	0.124	0.264	0.290	0.484	0.280	FALSE	0.772	0.791	N
352	0.086	0.133	0.083	0.167	0.360	0.081	0.124	0.264	0.290	0.484	0.280	FALSE	0.800	0.814	N
380	0.083	0.134	0.084	0.172	0.364	0.089	0.163	0.297	0.317	0.553	0.336	FALSE	0.845	0.827	N
400	0.075	0.133	0.075	0.159	0.360	0.084	0.124	0.264	0.290	0.484	0.280	FALSE	0.745	0.747	N
461	0.070	0.162	0.075	0.156	0.312	0.086	0.091	0.253	0.273	0.484	0.252	FALSE	0.706	0.691	N
491	0.075	0.155	0.081	0.164	0.319	0.089	0.100	0.255	0.277	0.500	0.270	FALSE	0.750	0.753	N
529	0.089	0.133	0.092	0.158	0.360	0.069	0.124	0.264	0.290	0.484	0.280	FALSE	0.789	0.814	N
593	0.089	0.133	0.086	0.173	0.360	0.084	0.124	0.264	0.290	0.484	0.280	FALSE	0.817	0.833	N
935	0.075	0.155	0.080	0.159	0.328	0.084	0.109	0.264	0.286	0.503	0.273	FALSE	0.797	0.781	N
956	0.081	0.133	0.086	0.173	0.360	0.092	0.124	0.264	0.290	0.484	0.280	FALSE	0.822	0.817	N
1144	0.086	0.139	0.086	0.167	0.347	0.081	0.136	0.275	0.292	0.528	0.303	FALSE	0.822	0.834	N
1170	0.075	0.131	0.081	0.161	0.344	0.086	0.141	0.272	0.292	0.514	0.308	FALSE	0.789	0.808	N
1270	0.091	0.198	0.083	0.172	0.350	0.081	0.083	0.281	0.303	0.536	0.247	FALSE	0.800	0.814	N
1346	0.086	0.133	0.092	0.164	0.360	0.078	0.124	0.264	0.290	0.484	0.280	FALSE	0.794	0.800	N

Showing 1 to 20 of 18,542 entries

FIGURE 4.20 – La matrice de données des caractéristiques sélectionnées et calculées.

4.3.4 Entraînement avec l'arbre de décision optimisé

a Choix aléatoire du nombre d'arbres pour l'arbre de décision : Apprentissage initial

Nous sélectionnons au hasard un nombre d'arbre pour un apprentissage initial

```
recursive.Num_Aliatoire_Arb <- function() {  
  y <- sample(-100:100, 1)  
  nt_new <- NUM_arb +y  
  
  if(length(liste_TS)==0){  
    return(nt_new)  
  }else{  
    id<-1  
    while(id<=length(liste_TS) && nt_new !=liste_TS[id])  
    {  
      id<-id+1  
    }  
    if(id<=length(liste_TS))  
    {  
      nt_new<-recursive.Num_Aliatoire_Arb()  
      return (nt_new)  
    }else{  
      while(nt_new > 1000)  
      {  
        nt_new<-recursive.Num_Aliatoire_Arb()  
      }  
      return (nt_new)  
    }  
    if(nt_new<0)  
    {  
      nt_new <- nt_new*(-1)  
      return (nt_new)  
    }  
  }  
}
```

FIGURE 4.21 – Une fonction de sélection d'un nombre d'arbre au hasard.

b L'arbre de décision

L'ensemble de données assemblées sera divisé en deux parties, la première est consacrée pour l'apprentissage, et la seconde partie sera utilisée comme ensemble de test.

```
ind = sample(2,nrow(FeaturesMatrix),replace=TRUE, prob = c(0.9, 0.1))  
trainData = FeaturesMatrix[ind==1, ]  
testData = FeaturesMatrix[ind==2, ]
```

On note que les entités extraites contiennent des valeurs manquantes, ce qui influe sur les entités calculées. Pour résoudre ce problème, nous avons utilisé la fonction

na.roughfix de la bibliothèque R randomForest, qui impute les valeurs manquantes par une valeur moyenne :

```
trainData = na.roughfix(trainData)
testData = na.roughfix(testData)
```

Le package randomForest donne la possibilité de lancer l'étape de l'apprentissage du modèle, nous donnons le nombre choisi dans l'étape précédente comme paramètre, pour ce faire, on exécute les instructions suivantes :

```
num_arb <- recursive.Num_Aliatoire_Arb <- function()
rf <- randomForest(as.factor(trainData$BeatClass)~., data = trainData, nTree=num_arb)
```

Nous calculons et sauvegardons le taux de précision de modèle obtenu pour le comparer avec d'autres taux calculés durant les itérations de l'algorithme Recherche Tabou (expliqué ci-dessous). Le meilleur taux obtenu sera considéré comme optimal.

```
testData_classReal <- testData$BeatClass
RESULTclassRF <- predict(rf, newdata=testData)
#calculer le taux
Taux_classification <- sum(RESULTclassRF == testData_classReal) / length(RESULTclassRF)
Taux_classification <- Taux_classification*100
cat (" le taux de prédation est := ",Taux_classification," (%) avec nombre d'arbre ",num_arb," \n")

Taux_best <- Taux_classification
nbr_arb_best<-num_arb
```

c Recherche tabou (RT)

L'algorithme recherche Tabou itère par une génération d'une solution voisine qui est un choix d'un nombre d'arbres en se basant sur la solution initiale trouvée précédemment (le nombre d'arbres initial). Par la suite, la meilleure solution (le nombre d'arbres) est choisie suite à une comparaison en termes de taux de précision de classification. Sinon ce nouveau nombre d'arbres est mis dans la liste tabou, pour qu'il ne soit pas évalué ultérieurement. Ce processus est effectué en boucle jusqu'à la vérification d'un critère d'arrêt. Dans notre cas, on a choisi comme critère d'arrêt, un nombre maximum d'itération fixé à 200 comme dans les travaux de recherche présentés dans la littérature.

```

138 if(Taux_new < Taux_best)
139 { if(length(liste_TS)<180){
140   liste_TS[id_liste_TS]<- new_num_arb
141   id_liste_TS<-id_liste_TS+1
142   }else
143   { # FIFO
144     id_fifo<-1
145     while(id_fifo<length(liste_TS))
146     {
147       liste_TS[id_fifo]<-liste_TS[id_fifo+1]
148       id_fifo<-id_fifo+1
149     }
150     id_liste_TS<-id_liste_TS-1
151   }
152 }else{
153   if(Taux_new > Taux_best){
154     if(length(liste_TS)<100){
155       liste_TS[id_liste_TS]<- nbr_arb_best
156       id_liste_TS<-id_liste_TS+1
157     }else
158     { # FIFO
159       id_fifo<-1
160       while(id_fifo<length(liste_TS))
161       {
162         liste_TS[id_fifo]<-liste_TS[id_fifo+1]
163         id_fifo<-id_fifo+1
164       }
165       id_liste_TS<-id_liste_TS-1
166     }
167     Taux_best<-Taux_new
168     nbr_arb_best<- new_num_arb
169   }else # Taux_new = Taux_best
170   { if(nbr_arb_best>new_num_arb){
171     if(length(liste_TS)<100){
172       liste_TS[id_liste_TS]<- nbr_arb_best
173       id_liste_TS<-id_liste_TS+1
174     }else
175     { # FIFO
176       id_fifo<-1
177       while(id_fifo<length(liste_TS))
178       {
179         liste_TS[id_fifo]<-liste_TS[id_fifo+1]
180         id_fifo<-id_fifo+1
181       }
182       id_liste_TS<-id_liste_TS-1
183     }
184     nbr_arb_best<-new_num_arb
185   }
186 }
187 }
188 }
189 }
190 }
191 }
192 cat (" Le taux de prédation est := ",Taux_best," (%) avec nombre d'arbre ",nbr_arb_best," \n")

```

4.3.5 Visualisation, Interprétation, Utilisation

Comme la sauvegarde de données est une nécessité exigée par les médecins cardiologues pour qu'ils puissent prendre la décision de thérapie adéquate, nous avons développé une application au profit du médecin pour qu'il puisse inscrire, contrôler et surveiller les patients à distance.



FIGURE 4.22 – Interface de l'application médecin.

Chaque médecin peut s'inscrire pour obtenir son propre compte :

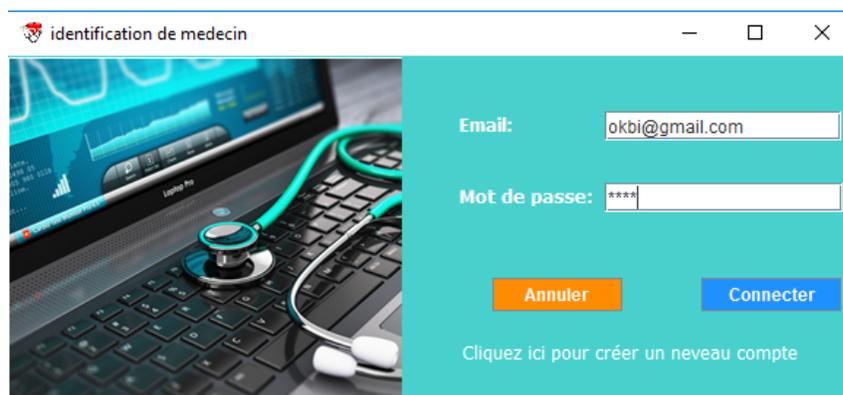


FIGURE 4.23 – Interface pour l'identification de médecin.

Le médecin peut ajouter des commentaires et des rendez-vous à retenir pour des usages plus tard.

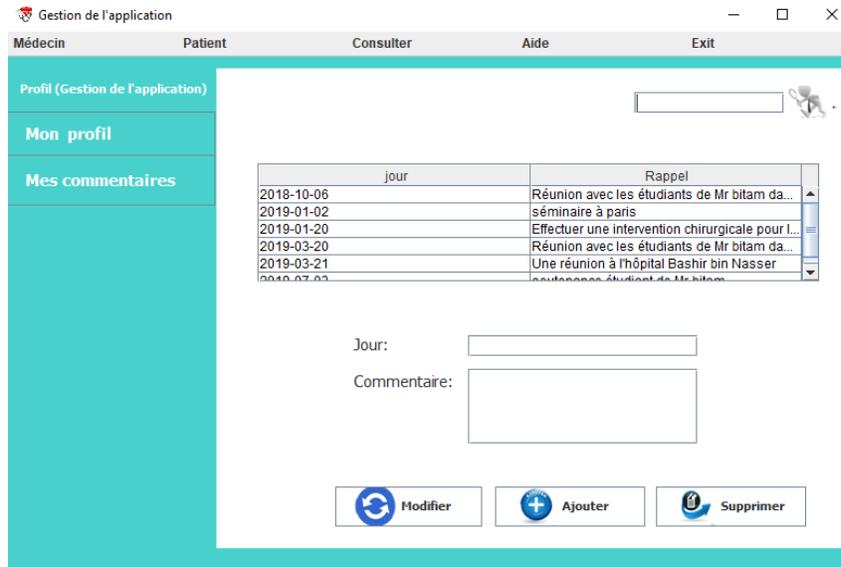


FIGURE 4.24 – Interface pour ajouter des commentaires et des rendez-vous.



FIGURE 4.25 – Interface de profil de médecin.

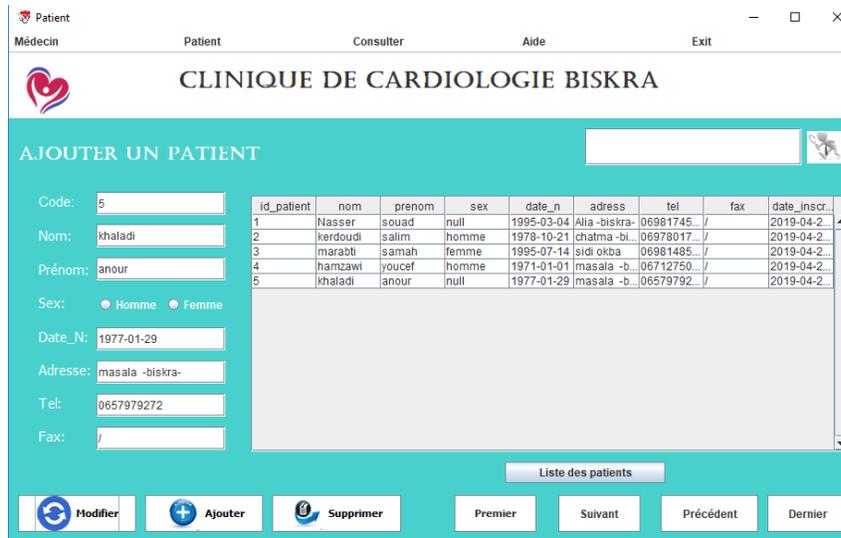


FIGURE 4.26 – Interface ajouter un patient.

Le médecin peut consulter les signaux avec des différentes dérivations et pour des intervalles de temps désirés :



FIGURE 4.27 – Interface consulter les signaux avec des dérivations différentes.

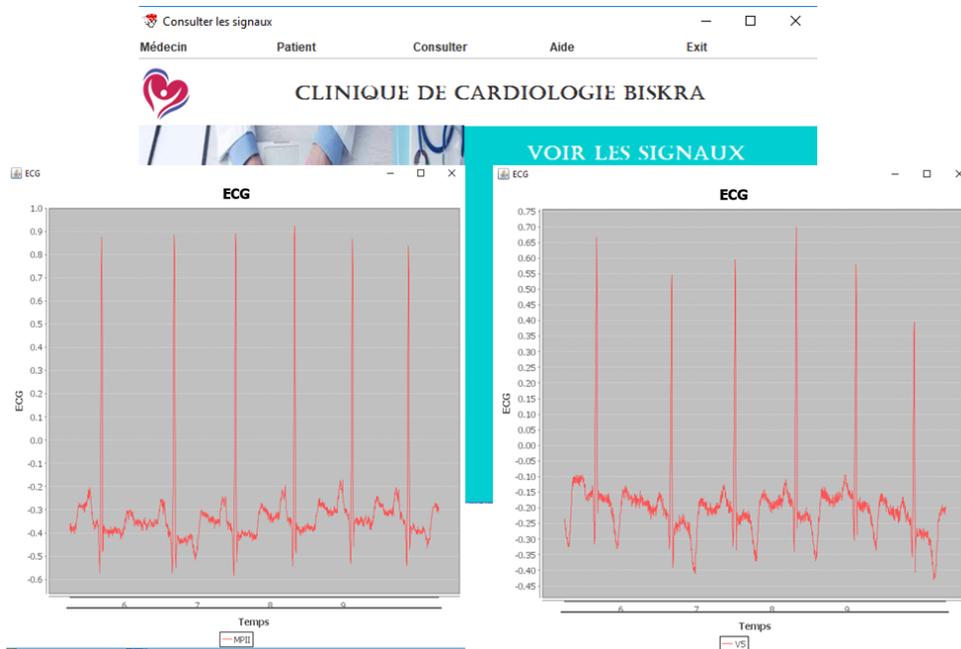


FIGURE 4.28 – Interface les signaux avec des dérivations différentes d’un patient.

4.4 Résultats obtenus et discussion

Les résultats de la classification de notre proposition sont illustrés aux figures suivant. Nous effectuons une étude comparative entre les arbres de décisions améliorés (OAD) et les arbres de décision classiques.

Nous mentionnons clairement l’impact du nombre trouvé d’arbres OAD sur la précision de la classification afin de prédire les battements des signaux ECG.

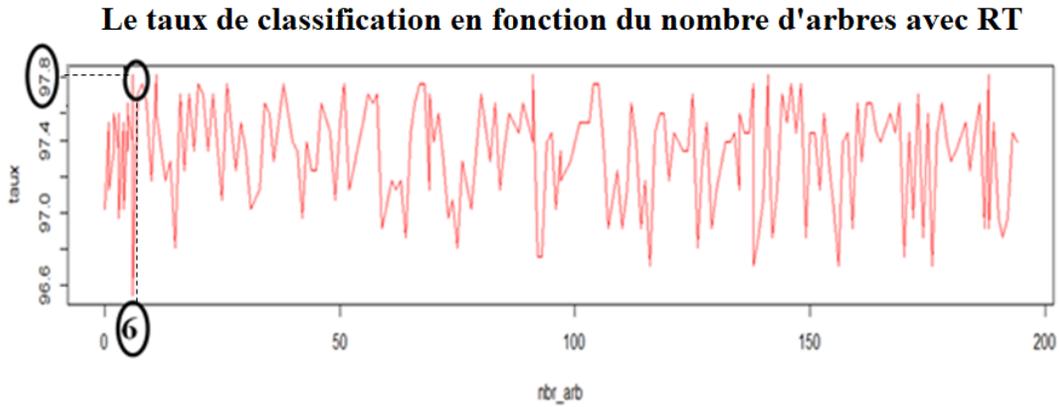


FIGURE 4.29 – les valeurs de taux de précision en fonction de nombre des arbres avec algorithme d’optimisation Recherche Tabou.

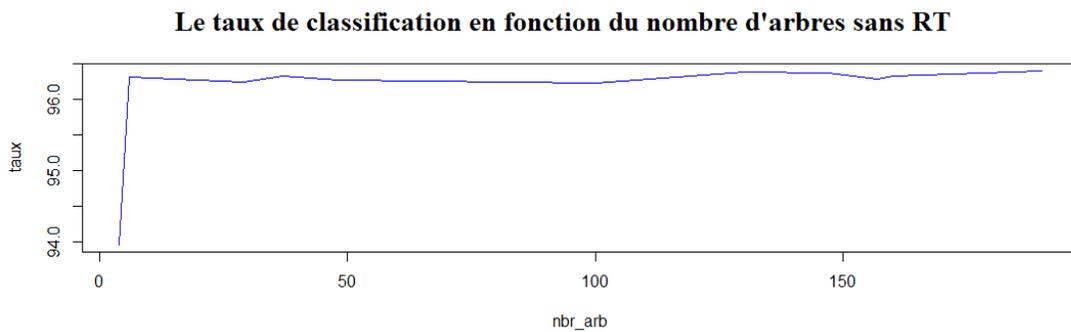


FIGURE 4.30 – les valeurs de taux de précision en fonction de nombre des arbres sans algorithme d’optimisation RT.

Les résultats trouvés illustrés dans figures montrent que les arbres de décision améliorés surpassent les arbres de décision classiques en termes de précision de la classification ECG par un taux égal à 97,8% après avoir atteint le nombre optimal d’arbres, et en terme de taille d’arbre qui est de 6 arbres (voir la figure 4.32). Par contre, l’arbre de décision l’originale ne dépasse pas le taux d’évaluation 96.5% pour un nombre d’arbre égal à 198 choisi de façon aléatoire.

4.5 Conclusion

Dans ce chapitre, nous avons présenté l'environnement et les outils mis en œuvre dans notre projet et nous avons exposé l'implémentation de notre système. Nous avons commencé par l'acquisition et la préparation de la plate-forme e-HP et par le traitement des fichiers de la base de données disponible sur le site web de Physionet. Puis, nous avons présenté les étapes nécessaires à l'entraînement et au réglage du classifieur AD utilisé avec le nombre optimal des d'arbres pour AD obtenu par l'algorithme d'optimisation Recherche Tabou. À la fin, nous avons discuté des résultats trouvés par notre classificateur qui ont été satisfaisants en termes de précision de classification du signal ECG.

Conclusion générale

Le diagnostic précoce d'une maladie cardiovasculaire telle que l'arythmie cardiaque peut sauver des vies humaines et aider à fournir un traitement rapide pour éviter des séquelles physiques, psychologiques et financières lourdes. L'ECG est le principal outil permettant de diagnostiquer l'activité électrique du cœur car toute anomalie présente dans l'activité cardiaque est reflétée dans les signaux ECG. Cependant, l'évaluation visuelle des signaux ECG est une tâche difficile et qui prend beaucoup de temps. Par conséquent, la mise en œuvre d'un système assure une surveillance permanente du patient garantira un diagnostic objectif et rapide de l'arythmie cardiaque.

Pour faire face à ce type d'anomalie, nous avons proposé dans ce travail de recherche une optimisation de la méthode de diagnostic automatique de l'arythmie cardiaque qu'est l'arbre de décision, afin d'obtenir le meilleur nombre d'arbres, conduisant à un taux de classification le plus élevé avec la plus petite taille d'arbre possible. Pour ce faire nous avons utilisé l'algorithme recherche tabou (RT).

La conservation durable et non interrompue des données de l'ECG des patients est très importante. Pour cette raison, nous avons proposé l'idée de créer un système qui permet l'enregistrement permanent des données de type massive (Big data), basé sur le modèle multidimensionnel.

Les résultats obtenus ont été satisfaisants en termes de taux de classification, nous avons atteint une précision de 97,8% en utilisant seulement 6 d'arbres. Cela montre que les performances globales de notre système sont acceptables et peuvent les médecins spécialistes dans leur diagnostic

Comme perspectives, on peut envisager de minimiser la taille des appareils utilisés pour le suivi afin de faciliter leur utilisation et d'assurer une démocratisation de tels outils. La précision de la classification peut être aussi renforcée en augmentant le nombre de la base de données, et les caractéristiques utilisées. Cette vision futuriste vise à garantir une surveillance des patients en temps réel et à offrir de meilleures tâches de suivi au profit des citoyens, des patients, et des médecins cardiologues.

On peut penser aussi à étendre cette étude à d'autres signaux biologiques tels que

l'électroencéphalographie EEG de sauver des vies humaines qui peuvent subir d'autres types d'anomalies biologiques

Bibliographie

- [1] DUBOIS, R,(2004). Application des nouvelles méthodes d'apprentissage à la détection précoce d'anomalies en électrocardiographie. L'université paris 6.
- [2] Dubin, M. D (1970,1974). Lecture accélérée de l'ECG, maloine S.A éditeur PARIS.
- [3] Yves Benisty. Électrocardiogramme (ÉCG) 1, Définition. Extrait le 6 octobre 2018, https://sofia.medicalistes.fr/spip/IMG/pdf/Electrocardiogramme_Yves_Benisty_.pdf
- [4] Collège National des Enseignants de Cardiologie et Maladies Vasculaires. Généralités, Université Médicale Virtuelle Francophone.
- [5] Le cycle cardiaque. Thanatofrance écoles et préparation au diplôme national de Thanatopracteur enregistrée à la préfecture des Alpes-Maritimes. <https://thanatofrance.wordpress.com/2015/07/06/le-cycle-cardiaque/>
- [6] Sztajzel, J. (n.d.). Introduction à l'ECG. Service de Cardiologie Hôpitaux Universitaires Genève. Extrait le 6 octobre 2018, http://medweb1.unige.ch/enseignement/apprentissage/module2/circ/apprentissage/intranet/cours/intro_ecg_dias.pdf
- [7] Sophie, P. Sabrina, V.Aide à l'interprétation des ECG Ã l'attention des étudiants de 3 ème et 4 ème SIAMU.CHU Brugmann UVC, 2013. Extrait le 6 octobre 2018, <https://www.chu-brugmann.be/fr/news/20131203-criticalday-volpe.pdf>
- [8] Extrait le 03 novembre 2018, https://www.google.com/search?q=les+d%C3%A9rivations+monopolaires&source=lnms&tbm=isch&sa=X&ved=0ahUKEwj1jZrvhrneAhVi_SoKHdFnBdkQ_AUIDigBbiw=1366bih=657imgrc=bceMA1Bd4vBSaM
- [9] <https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=infarctus-du-myocarde>

- [10] <http://m.20-bal.com/pravo/1369/index.html>
- [11] L'électrocardiogramme de repos normal, Université médicale virtuelle francophone. Extrait le 08 juin 2019, <http://campus.cerimes.fr/semiologie-cardiologique/enseignement/electrocardiogramme/site/html/5.html5a>
- [12] INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA, Extrait le 03 May 2019, <https://www.ottawaheart.ca/fr/maladie-du-c%C5%93ur/les-arythmies-anomalies-du-rythme-cardiaque>
- [13] <http://m.20-bal.com/pravo/1369/index.html>
- [14] Pr JP Camous, CHU NICE.ECG InterpétationInterprétation.DocPlayer,. October 6, 2018 <https://docplayer.fr/20811159-Ecg-pr-jp-camous-chu-nice-pr-p-ambrosi-chu-marseille.ht>
- [15] Fifi Salmon. Principes d'électrocardiographie Électrocardiogramme normal. Extrait le 13 octobre 2018, <https://slideplayer.fr/slide/2580548>
- [16] CABASSON A. Estimation et analyse des intervalles cardiaques, Université de Nice - Sophia Antipolis Extrait le 13 octobre 2018, <https://tel.archives-ouvertes.fr/tel-00357389/document>
- [17] Gaétan T. Initiation à l'ECG et aux arythmies cardiaques. Extrait le 13 octobre 2018, <https://slideplayer.fr/slide/1144542/>
- [18] Pierre T. Formation à « l'ECG de A à Z » par Pierre Taboulet, Extrait le 13 octobre 2018, <https://www.e-cardiogram.com>
- [19] H. Banae , M. U. Ahmed and A. Loutfi, "Data Mining for Wearable Sensors in Health Monitoring Systems : A Review of Recent Trends and Challenges", Sensors 2013, 13, 17472-17500 ; doi :10.3390/s131217472.
- [20] Z. Masetic, and A. Subasi., "Congestive heart failure detection using random forest classifier", Computer methods and programs in biomedicine, vol. 130, p. 54-64, 2016.
- [21] Baim DS, Colucci WS, Monrad ES, Smith HS, Wright RF, Lanoue A, Gauthier DF, Ransil BJ, Grossman W, Braunwald E. Survival of patients with severe congestive heart failure treated with oral milrinone. J American College of Cardiology 1986 Mar ; 7(3) :661-670. [View Abstract]

- [22] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23) :e215-e220 [Circulation Electronic Pages ; <http://circ.ahajournals.org/content/101/23/e215.full>] ; 2000 (June 13).
- [23] Kaur, M., Arora, A. S. (2010). Unsupervised Analysis of Arrhythmias using K-means Clustering. *International Journal of Computer Science and Information Technologies*, 1(5), 417-419.
- [24] U.R. Acharya , H. Fujita , S.L. Oh , Y. Hagiwara , J.H. Tan , A. Muhammad , Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network, *Inf. Sci.* 405 (2017) 81–90 .
- [25] Detecting heart arrhythmias using machine learning and Apple Watch data(2016), Extrait le 3 Juin 2019, <https://blog.insightdatascience.com/detecting-heart-arrhythmias-using-machine-learning-and-apple-watch-data-84bc97f30d3e>
- [26] Pal, M., 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing* 26 (1), 217–222.
- [27] Gendreau .M, Introduction à la recherche avec tabous, CIRRELT et MAGI, École Polytechnique de Montréal
- [28] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23) :e215-e220 [Circulation Electronic Pages ; <http://circ.ahajournals.org/content/101/23/e215.full>] ; 2000 (June 13). Extrait le 29 Mai 2019, <https://physionet.org>
- [29] M.H. Rai, A. Trivedi, and S. Shukla., "ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier", *Measurement*, vol. 46, no 9, p. 3238-3246, 2013
- [30] NITA S, BITAM S, MALLOUK A. An Enhanced Random Forest for Cardiac Diseases Identification based on ECG signal.

- [31] <https://physionet.org/physiobank/>
- [32] <https://physionet.org/physiotools/>
- [33] <https://www.cooking-hacks.com/ehealth-sensors-complete-kit-biometric-medical-arduino-raspberry-pi>
- [34] Microsoft, Extrait 20 juin 2019 <https://docs.microsoft.com/fr-fr/sql/analysis-services/multidimensional-models/multidimensional-models-ssas?view=sql-server-2017>
- [35] J.-F. Desnos, Entrepôt de données – Introduction. . Extrait le 29 Mai 2019, <http://imss-www.upmf-grenoble.fr/prevert/SpecialiteIHS/ED/IntroductionED.pdf>
- [36] Richard K. L(2001) , Extraction des connaissances à partir d'un datamart à l'aide de l'arbre de décision application aux données médicales. Extrait le 29 Mai 2019, https://www.memoireonline.com/04/14/8861/m_Extraction-des-connaissances-partir-d-un-datamart-l-aide-de-l-arbre-de-decision-application-au17.html
- [37] developpez, Conception d'un entrepôt de données (Data Warehouse). Extrait le 29 Mai 2019, <https://grim.developpez.com/cours/businessintelligence/concepts/conception-datawarehouse/>
- [38] http://eeweb.poly.edu/iselesni/sass/SASS_toolbox/html/Example2.html
- [39] H.EZZIKOURI et M. Fakir, Algorithmes de classification : ID3 C4.5. Université Sultan Moulay Slimane, Faculté des Sciences et Techniques Béni Mellal
- [40] Abd El Hamid Djefal. "Fouille de données Avancé", cour 2018/2019.
- [41] Extrait le 02 juin 2019, <https://www.futura-sciences.com/tech/definitions/internet-java-485/>
- [42] Extrait le 02 juin 2019, <https://www.techno-science.net/definition/517.html>
- [43] Extrait le 02 juin 2019, <https://www.jmdoudoux.fr/java/dej/chap-jdbc.htmjdb>
- [44] Extrait le 02 juin 2019, https://fr.wikipedia.org/wiki/Java_Database_Connectivity
- [45] S. Déjean. Présentation du logiciel R, Institut de Mathématiques de Toulouse UMR 5219 Université Paul-Sabatier (Toulouse III) .Extrait 03 juin 2019, <https://www.math.univ-toulouse.fr/sdejean/PDF/semin-R.pdf>

- [46] Extrait le 02 juin 2019, <https://quanti.hypotheses.org/488/>
- [47] Extrait le 02 juin 2019, <https://www.rdocumentation.org/packages/randomForest/versions/4.6-14/topics/randomForest>
- [48] <http://scikit-learn.org/stable/>
- [49] (5 avril 2018) Machine Learning avec le package « caret ». Extrait 03 juin 2019, http://eric.univ-lyon2.fr/ricco/tanagra/fichiers/fr_Tanagra_package_caret.pdf
- [50] Extrait 03 juin 2019, <https://desgeeksetdeslettres.com/web/xampp-plateforme-pour-heberger-son-propre-site-web>
- [51] cooking-hacks, Extrait 03 juin 2019, <https://www.cooking-hacks.com/documentation/tutorials/ehealth-biometric-sensor-platform-arduino-raspberry-pi-medical>
- [52] Drid.A(2017/2018), pattern recognition for healthcare analytics
- [53] Extrait 03 juin 2019, <https://physionet.org/physiobank/database/ecg>
- [54] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 20(3) :45-50 (May-June 2001). (PMID : 11446209). Extrait 13 Janvier 2019, <https://physionet.org/physiobank/database/mitdb/>
- [55] Greenwald SD. Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information. Ph.D. thesis, Harvard-MIT Division of Health Sciences and Technology, 1990. Extrait 13 Janvier 2019, <https://physionet.org/physiobank/database/svdb/>
- [56] Petrutiu S, Sahakian AV, Swiryn S. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. Europace 9 :466-470 (2007). Extrait 13 Janvier 2019, <https://physionet.org/physiobank/database/Itafdb/>
- [57] Extrait Janvier 2019, <https://www.physionet.org/faq.shtml>
- [58] Extrait Janvier 2019, <https://www.physionet.org/physiotools/wfdb.shtml>
- [59] Extrait Mars 2019, <https://www.codeproject.com/Articles/20995/ECG-Annotation-C-Library>