



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 11/M2/2019

Mémoire

Présenté pour obtenir le diplôme de master académique en

Informatique

Parcours : **Systeme d'information optimisation et de décision**
(SIOD)

Surveillance médicale à distance à base de data mining

Par :

BEKIRI ROUMAISSA

Soutenu le 07 juillet 2019, devant le jury composé de :

Bali Mahboub Abdelmadjid	M.A.A	Président
Abdelhamid Djefal	M.C.A	Rapporteur
Guemeida Abdelbasset	M.C.B	Examineur

Dédicace

Toutes les lettres ne sauraient trouver les mots qu'il faut... Tous les mots ne sauraient exprimer la gratitude, l'amour, le respect, la reconnaissance... Aussi, c'est tout simplement que, Je dédie ce modeste travail à:

À ma chère mère, tu représentes pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement qui n'a pas cessé de m'encourager et de prier pour moi. Ta prière et ta bénédiction m'ont été d'un grand secours pour mener à bien mes études.

À mon cher père, aucune dédicace ne saurait exprimer l'amour, l'estime, le dévouement et le respect que j'ai toujours eu pour vous. Ce travail est le fruit de tes sacrifices que tu as consentis pour mon éducation et ma formation.

À ma chère sœur Manel, qui a toujours été de mon côté, je te souhaite un avenir plein de joie, de bonheur, de réussite et de sérénité..

À mes chères sœurs: Naouel, Rofia, Soulef, Yasmina, Moufida. Votre affection et votre soutien m'ont été d'un grand secours au long de ma vie professionnelle et personnelle.

À mon beau frère Sabre, qui a toujours été de mon côté, et m'avez aidé, je vous souhaite un avenir plein de réussite dans ton vie professionnelle.

À mes chères amies Asma, Saouel, Ibtissem, Amina, Sara. En souvenir de notre sincère et profonde amitié et des moments agréables que nous avons passés ensemble.

Une spéciale dédicace à tous ceux qui m'ont aimé le bien et le succès et qui m'ont soutenu tout au long des années passées et qui était la raison de mon succès.

À tous mes collègues de l'université de Biskra. Et à tous ce qui ont enseigné moi au long de ma vie scolaire.

Remerciement

Tout d'abord, je remercie le Dieu, notre créateur de m'avoir donné la force, la volonté et le courage afin d'accomplir ce travail modeste.

Je veux adresser les grands remerciements les plus sincères à mon encadreur "Dr. Abdelhamid Djeflal" qui a proposé le thème de ce mémoire, pour ses conseils son encouragement, sa patience, son aide précieuse et pour le temps qu'il m'a consacré. Et ses dirigés du début à la fin de ce travail pour l'attention et la disponibilité dont il a su faire preuve au cours de la réalisation de ce mémoire.

Je remercie mon professeur, "Meadi Mohammed Nadjib" qui m'a aidé dans mon mémoire, et qui m'a motivé à donner le meilleur pour mon projet.

Je tiens également à remercier messieurs les membres de jury pour l'honneur qu'ils m'ont fait en acceptant de siéger à notre soutenance.

Finalement, je tiens à exprimer ma profonde gratitude à ma famille qui ma soutenue et encouragée tout au long de mes études. Ainsi que l'ensemble des enseignants qui ont contribué à ma formation.

Résumé

Le domaine de la e-santé connaît des progrès rapides et croissant avec l'évolution des moyens de communication et d'acquisition des données sanitaires à travers les montres intelligentes telles que le rythme cardiaque, le signal cardiaque, la température, le niveau de glucose, ...etc.

Ce travail propose un système de suivi automatique de la santé des individus à distance basé sur les montres intelligentes portées par les patients et les médecins ainsi qu'un serveur centralisé de l'administration. Ce dernier utilise les données accumulées des patients pour construire des modèles de décision en utilisant les méthodes de machine Learning. Ces modèles sont utilisés pour suivre automatiquement la santé des patients.

Le système a été conçu, implémenté et validé sur des données concernant les maladies cardiovasculaires.

Mots clés : Surveillance médicale, mesure physiologique, Data mining, Apprentissage automatique, montres intelligentes, SVM (classification).

ملخص

يعرف مجال الصحة الإلكترونية تقدماً سريعاً وامتزاجاً خاصاً مع تطور وسائل الاتصال والحصول على البيانات الصحية عن طريق الساعات الذكية مثل معدل ضربات القلب، إشارة القلب، درجة الحرارة، مستوى الجلوكوز... الخ

يقترح هذا العمل نظاماً للمراقبة التلقائية لصحة الأفراد عن بعد تعتمد على الساعات الذكية التي يرتديها المرضى والأطباء بالإضافة إلى خادم مركزي للإدارة. يستخدم هذا الأخير بيانات المرضى المتراكمة لإنشاء نماذج قرارات باستخدام أساليب التعلم الآلي. تستخدم هذه النماذج لتتبع صحة المرضى تلقائياً.

تم تصميم النظام وتنفيذه والتحقق من صحته بتطبيقه على البيانات المتعلقة بأمراض القلب والأوعية الدموية.

الكلمات المفتاحية: المراقبة الطبية، القياس الفسيولوجي، استخراج البيانات، التعلم الآلي، الساعات الذكية،

الات الاشعة الحاملة

ABSTRACT

The field of e-health knows rapid and increasing progress with the evolution of the means of communication and acquisition of health data through smartwatches such as heart rate, cardiac signal, temperature, level of glucose, ..etc.

This work proposes a system of automatic remote monitoring of individuals health based on smart watches worn by patients and doctors as well as a centralized server of the administration. The latter uses accumulated patient data to build decision models using machine learning methods. These models are used to automatically track the health of patients.

The system has been designed, implemented and validated on data concerning cardiovascular diseases.

Key words : Medical surveillance, physiological measurement, Data mining, machine learning, smart watches, SVM (classification).

Table des matières

1	SANTÉ NUMÉRIQUE	3
1.1	Introduction	3
1.2	Importance du domaine de la santé numérique	4
1.3	Enjeux des technologies numériques appliquées aux domaines de la santé	4
1.3.1	Des perspectives prometteuses pour la médecine	4
1.3.1.1	Facteur de développement des coopérations professionnelles	5
1.3.1.2	Évolution des modalités de prise en charge des patients	5
1.3.1.3	Un partage d'informations en croissance exponentielle et quasi instantané	5
1.3.2	Enjeux de l'efficacité administrative	6
1.3.3	Enjeux du traitement des informations médicales	6
1.4	Les grands domaines de l'informatique médicale	7
1.4.1	Bio-informatique	7
1.4.2	L'informatique clinique ou bioclinique	8
1.4.3	Informatique de santé publique	8
1.5	L'informatique médicale et l'informatique générale	8
1.5.1	Les bases de données	8
1.5.2	L'intelligence artificielle	9
1.5.3	Les réseaux	9
1.6	Moyens informatiques pour le suivi de la santé	9
1.6.1	Dispositif connectés pour la santé	9
1.6.1.1	Dispositifs à la ceinture	9
1.6.1.2	Téléphones mobiles	10
1.6.1.2.0.1	Application CareZone	10
1.6.1.2.0.2	Application Epocrates	11
1.6.1.3	Médailles	12

1.6.1.4	Bracelet-téléphone	12
1.6.1.5	Montres intelligentes	13
1.6.1.5.0.1	Composants des montres intelligentes	14
1.6.1.5.0.2	Samsung Gear	14
1.6.1.5.0.3	Huawei watch	15
1.6.1.5.0.4	Apple watch	15
1.6.2	Informations de santé mesurables par moyens mobiles connectés . .	16
1.6.2.1	Position	16
1.6.2.2	Mouvement	17
1.6.2.3	Chute	17
1.6.2.4	Tension Artérielle	17
1.6.2.5	Rythme cardiaque	18
1.6.2.6	Signal cardiaque	18
1.6.2.7	Température	19
1.6.2.8	Niveau de glucose	19
1.6.2.9	Sommeil	20
1.6.3	Système d'information médical	20
1.6.3.1	Télé-santé	21
1.6.3.2	Télé-médecine	22
1.6.3.3	M-santé (mobile-santé)	22
1.7	Architectures et systèmes d'information pour la e-santé	22
1.7.1	Caractéristiques d'un système pour la e-santé	22
1.7.2	Systèmes Mobiles de Surveillance de Patients (SMSP)	23
1.7.2.1	Réseau local sans fil (RLSF)	24
1.7.2.2	Serveur local à domicile (SLD)	24
1.7.2.3	Serveur médical distant (SMD)	24
1.7.3	Système utilisant des capteurs	25
1.8	Conclusion	25
2	DATA MINING	27
2.1	Introduction	27
2.2	Définition d'un Data Mining	27
2.3	Les domaines d'application de Data Mining	28
2.3.1	Le data mining dans la banque	29
2.3.2	Le data mining dans l'assurance	29
2.3.3	Le data mining dans le commerce	29
2.3.4	Le data mining dans la détection de fraude	29
2.3.5	Le Data mining dans la médecine et la pharmacie	29
2.4	Processus de Data Mining (ECD)	30

2.4.1	Nettoyage et intégration des données	31
2.4.2	Pré-traitement des données	31
2.4.3	Fouille de données	31
2.4.4	Évaluation et présentation	31
2.4.4.1	La validation statistique	32
2.4.4.2	La validation par expertise	32
2.5	Principales tâches de fouille de données	32
2.5.1	Classification	33
2.5.1.1	Classification supervisée (apprentissage)	34
2.5.1.2	Classification non supervisée	34
2.5.2	L'estimation	34
2.5.3	La prédiction	34
2.5.4	Les règles d'association	34
2.5.5	La segmentation	35
2.6	Apprentissage supervisé Vs Apprentissage non supervisé	35
2.7	Modélisation descriptive Vs prédictive	35
2.8	Les méthodes de data mining	36
2.8.1	Les méthodes classiques	36
2.8.2	Les méthodes sophistiquées	36
2.8.2.1	Arbre de décision	37
2.8.2.2	Les réseaux de neurones	38
2.8.2.3	Deep learning	40
2.8.2.4	Règles d'association	41
2.9	Les machines à support vecteurs (SVM)	42
2.9.1	Principe de la technique SVM	42
2.9.1.1	Hyperplan	43
2.9.1.2	Vecteurs supports	43
2.9.1.3	Marge	44
2.9.1.4	Maximiser la marge	44
2.9.2	Les étapes de construction d'un modèle de décision	45
2.9.3	Avantage et inconvénients des SVM	46
2.10	Conclusion	47
3	CONCEPTION	48
3.1	Introduction	48
3.2	Spécification des besoins	49
3.3	Le diagrammes des cas d'utilisation	50
3.3.1	Application patient	50
3.3.2	Application médecin	54

3.3.3	Administration du système	56
3.4	Architecture du système	57
3.5	Sécurité	58
3.6	Conclusion	60
4	IMPLÉMENTATION ET RÉSULTAT	61
4.1	Introduction	61
4.2	Environnement et outils de programmation	61
4.2.1	XAMPP	62
4.2.2	Apache	62
4.2.3	phpMyAdmin	63
4.2.4	Android Studio	63
4.2.5	Android Wear	64
4.2.6	Weka	65
4.2.7	Bibliothèque LibSVM	66
4.2.7.1	Paramètres SVM pris en charge par la librairie [26]	66
4.2.8	Langages de programmation	67
4.2.8.1	PHP	67
4.2.8.2	Javascript	67
4.2.8.3	JAVA	68
4.2.8.4	SQL	68
4.3	Interface de l'application de Surveillance Médicale à Distance (SMD)	69
4.3.0.1	Application patient	70
4.3.0.2	Application Médecin	73
4.3.0.3	Application d'administration	76
4.4	Validation sur le cas des maladies cardiovasculaire	80
4.4.1	Base de données utilisé	80
4.4.2	Pré-traitements	82
4.4.3	Expérimentations et résultats	82
4.4.4	Discussion	86
4.5	Conclusion	86
5	CONCLUSION GÉNÉRALE	87
	Bibliographie	89

Table des figures

1	L'avancement technologique dans le monde médicale [13]	5
2	Accès par carte au service médical [13]	7
3	Domaines de l'informatique médicale[1]	7
4	Balise Aloïze[36]	10
5	Application CareZone[38]	11
6	Application Epocrates [38]	11
7	Le bracelet-téléphone Columba [47]	12
8	Le bracelet BlueGard [7]	13
9	Sumsung Gear S3[22]	14
10	Huawei watch [39].	15
11	Montre de santé Apple watch Series 4. [4]	16
12	Capteur GPS [32]	17
13	Montre de santé intelligente pour mesurer la tension artérielle.	17
14	Biocapteurs optiques : a) capteur de surveillance cardiaque à électrocardiogramme (ECG) ;(b) capteur de fréquence cardiaque à pince doigt	18
15	Signal électrocardiogramme (ECG) [41]	19
16	Capteur de température du corps humain [9].	19
17	Montre intellegentes(K'Watch) pour mesurer le niveaux de glucose [32].	20
18	Principaux domaines de la santé connecté [13].	21
19	Caractéristiques des montres intelligentes pour la santé [9]	23
20	Système mobile de surveillance de Patient(SMSP)[49]	24
21	Système d'information utilisant des capteurs[24]	25
1	Les différents domaines de Data mining [34].	28
2	Processus d'extraction de connaissances à partir des données [21].	30
3	Tâches de fouille de données.	33
4	Technique de data mining [53].	33
5	Arbre de décision[21]	37

6	Nœud d'un réseau de neurone[58].	39
7	Exemple d'utilisation du Deep learning [5].	41
8	L'hyperplan H qui sépare les deux ensembles de points.	43
9	L'hyperplan H qui sépare les deux ensembles de points.	43
10	L'hyperplan H optimal, vecteurs supports et marge maximale.	44
11	Meilleur hyperplan séparateur.	44
12	Schéma globale de fonctionnement de l'SVM	45
1	Diagramme de contexte d'un Système médical de surveillance.	50
2	"Diagramme de cas d'utilisation générale"	50
3	Diagramme de séquence "Authentification"	51
4	Diagramme de séquence "Profil du patient"	52
5	Diagramme de séquence "Envoie des mesures" par les méthodes "inform et request"	53
6	Diagramme de séquence "Réception des bilan"	53
7	Diagramme de séquence "Inscription de médecin"	54
8	Diagramme de séquence "Ajout de patients"	54
9	Diagramme de séquence "Ajouter de mesures"	55
10	Diagramme de séquence "Suivi de l'état de santé par le médecin"	55
11	Schéma relationnel de la base de données de notre système.	56
12	Fonctionnement globale d'un système d'information.	57
13	Architecture du Système de Surveillance Médicale à Distance(SSMD)	58
1	Informations sur envirement du Xampp [18].	62
2	Android studio [2]	64
3	Android studio wear[2]	64
4	Interface Weka[2]	65
5	Interface d'accueil du système(sur web, sur montre)	70
6	Interface d'inscription de patients(sur web, sur montre)	71
7	Interface de connection (sur web, sur montre	71
8	Application mesuré "rythme cardiaque" (sur montre)	72
9	Interface d'acquisition de données(sur web,sur montre)	72
10	Interface de saisie manuel des paramètres physiologique (sur montre)	72
11	Interface affichage du bilan (sur web)	73
12	Le menu de l'application patient(android wear)	73
13	Le menu de l'application médecin (sur montre)	74
14	Interface d'inscription du médecin(sur web, sur montre)	74
15	Interface d'ajout du patients(sur web, sur montre)	75
16	Interface d'ajout de mesures (sur montre)	75
17	Interface d'ajout de mesures (sur web)	76

18	Interface d'authentification (sur web)	77
19	Page de construction des modèles (sur web)	78
20	Page d'utilisation de modèles (sur web)	79
21	Page d'ajouter des notifications (sur web)	79
22	Description des attributs de la base utilisé	81
23	Fichier Weka "Heart.csv"	82
24	Précision et Recall en fonction du degré d du noyau polynomial pour la méthode SMO	83
25	Précision et Recall en fonction du degré d du noyau polynomial pour la méthode LibSVM	84
26	Précision et Recall en fonction de gamma du noyau RBF pour la méthode SMO	84
27	Précision et Recall en fonction de gamma du noyau RBF pour la méthode LibSVM	85
28	Précision et Recall en fonction de gamma du noyau RBF pour la méthode SMO	85
29	Précision et Recall en fonction de gamma du noyau RBF pour la méthode LibSVM	86

Introduction générale

Selon des rapports de statistiques récemment publiés par les nations unies, l'âge moyen de la population devrait croître rapidement dans les pays développés dans la prochaine décennie. Ce qui entraîne une augmentation très importante du coût des soins de santé. Cependant, les dernières innovations technologiques dans les différents domaines viennent à point nommé pour la mise en place de nouvelles approches plus efficaces et à coût abordable de prise en charge sanitaire des personnes.

Les soins de santé constituent un élément essentiel de la vie humaine. Chacun de nous a besoin d'un suivi périodique des paramètres vitaux et traitements adéquats sur la base de ses données médicales et son état de santé. Ces processus deviennent important quand les gens deviennent âgées et avec des maladies chronique telles que (les maladies cardiovasculaire, diabète, fréquence cardiaque non équilibrés, parkinson, cancer...) et ne sont pas en mesure de suivre leur état de santé correctement sans un personnel médical spécialisé ou des équipements sophistiqués pour la surveillance surtout dans les villes éloignés. Dans les situation d'urgences, le malade doit être transporté à l'hôpital, observé par un personnel médical et traité à temps si certains paramètres sont anormaux. Dans de nombreux cas, même un court délai du trajet pourrait conduire à des conséquences dangereuses y compris le décès du patient.

De nos jours, différents dispositifs et systèmes de surveillance de patients ont été proposés et mis en place pour des différents objectifs telles que : l'obtention de l'information sur la santé des patients de façon rapide et efficace, économies des ressources matérielles et financières relatives aux déplacements des patients aux différents centres médicaux, la surveillance périodique et alerte en cas d'urgences meilleures pour les personnes âgées ou handicapées, capable de fournir une information fiable et de façon continue sur le patient.

La développement des technologies surtout dans le domaine de la santé numérique permettent un contrôle rapide des individus à travers des moyens tels que les montres intelligentes, les téléphone, les tablette,...etc. Dans notre travail nous intéressons aux montres connectés (Smartwatch health).

Basé sur les éléments précédents, nous pouvons formuler l'objectif principal de notre travail, qui consiste en la combinaison des derniers développements dans le marché des montres intelligentes et des dispositifs de détection et de capture de signaux physiologiques de patients avec les techniques d'apprentissage automatique. Le but est la conception et la réalisation d'un système de surveillance continue et efficace à base d'analyse en utilisant des méthodes de data mining.

Notre approche est basée sur l'utilisation des smartwatches connectées pour la collecte des données sanitaires des personnes et leur stockage dans un système d'information centralisé pour les analyser et suivre l'état de santé de ces personnes d'une façon temps réel continue.

Effectivement, nous avons proposé un système basé sur des montres intelligentes portées par les patients et les médecins et connecté à travers les réseaux de communication à un serveur distant d'administration. Ce derniers utilise les données collectés des montres telles que le rythme cardiaque, la température, les mouvements, le niveau du glucose, ...etc. pour apprendre des modèles de décision capable de la détection précoce de l'apparition des maladies chez les personnes.

Nous avons développé, dans ce contexte, pour les patients et les médecins des interfaces web et des applications sur montres leur permettant d'accéder et profiter du système.

Ce mémoire comporte quatre chapitres organisés comme suit : Dans le premier chapitre, nous présentons quelques notions fondamentales liées aux domaine de la santé numérique, l'importance du domaines, les différents dispositifs utilisés dans le domaine de e-santé et les architectures existantes et plus particulièrement pour la surveillance médicale des patients.

Le second chapitre présente le domaine de data mining, les principaux concepts de fouille de données, ses tâches, ses méthodes, et plus particulièrement la méthode de SVM (classification) que nous utilisons dans notre projet ou nous détaillons son principe, ses avantages et ses inconvénients.

Le troisième chapitre est réservé à la conception de notre système, son architecture globale et détaillée. Le dernier chapitre est consacré à l'implémentation et la réalisation de la proposition de notre système, et aussi les expérimentations et la discussion des résultats obtenus. Le mémoire est terminé par une conclusion générale contenant les perspectives envisagées.

SANTÉ NUMÉRIQUE

1.1 Introduction

Les nouvelles technologies ont touché le domaine de la santé de manière dramatique ce qui a donné naissance au terme "E-santé". Le terme e-santé (e-health en anglais) ou encore télé-santé, santé numérique, sybersanté, santé connectée désigne tous les domaines où les technologies de l'information et de la communication (TIC) sont mises au service de la santé. Cela concerne des domaines tels que la télé-médecine, la prévention, le maintien à domicile, le suivi d'une maladie chronique à distance (diabète, hypertension, insuffisance cardiaque ...), les dossiers médicaux électroniques ainsi que les applications et la domotique.

La e-santé est donc un ensemble d'outils numériques permettant la transmission d'une grande variété de données importantes, codées et décodées par des moyens technologiques. Les systèmes d'information (SI) stockent ces données et utilisent des plateformes logicielles et de bases de données conçues pour une meilleure coordination des professionnels de santé.

L'e-santé apparaît comme l'une des solutions majeure aux grands défis de notre époque.

Le vieillissement de la population notamment et l'augmentation du diagnostic des maladies chroniques (cancer, hypertension, diabète, insuffisance cardiaque, parkinson, etc.)[50] Grâce aux nouvelles technologies de l'information et de la communication, particulièrement les systèmes existant de la santé numérique et facilitant la mobilité, la télé-médecine prend de plus en plus d'importance. A l'aide de capteurs, des systèmes mobiles, ou des applications sur les montres disséminés dans la maison, dans les villes isolées ou sur la personne suivie, des ensembles de données peuvent être obtenus, traités et analysés localement par un système d'information.

Cette analyse permet d'obtenir des systèmes experts permettant le suivi automatique et à distance des patients sans besoins d'être présent dans les centres médicaux.

Dans ce chapitre nous nous intéressons en premier lieu à l'importance du domaines de la santé numérique en générale, les différents enjeux technologiques relatifs à ce do-

maine permettant le suivi de la santé des patients à distance en utilisant des moyens informatiques.

1.2 Importance du domaine de la santé numérique

Aujourd'hui le domaine de santé numérique fait une grande évolution permettant de travailler en collaboration en utilisant plusieurs moyennes tels que les réseaux informatiques, apprentissage automatique, appareils connecté de mesures...etc. Cela permet de :

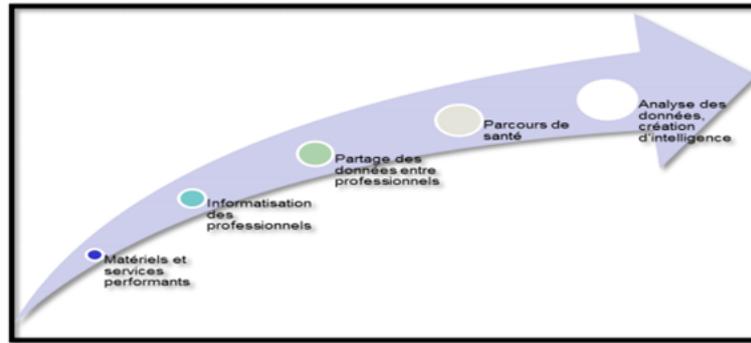
- L'obtention de l'information sur la santé des patients dans les bons délais.
- Rassembler des quantités de données permettant une meilleurs analyse des différentes maladies.
- Un meilleurs suivi des patients par les médecins.
- Utilisation des systèmes experts et apprentissage automatique pour le suivi enligne de la santé des patients.
- Économie des ressources matérielles et financières relatives aux déplacements des patients aux différents centres médicaux.
- Amélioration des délais d'intervention dans les cas d'urgence.
- Améliorer la cordination entre les différents services de santé et de sécurité.
- ...

1.3 Enjeux des technologies numériques appliquées aux domaines de la santé

1.3.1 Des perspectives prometteuses pour la médecine

Dans les dernières décennies, les progrès réalisés dans plusieurs domaines entre autre les technologies numérique, ont contribué à une évolution considérable dans le domaine de la médecine.

La figure suivant décrit l'avancement dans le monde médical en spécifiant la place des technologies numériques.



Source : ARS Pays de la Loire

FIGURE 1 – L'avancement technologique dans le monde médicale [13]

1.3.1.1 Facteur de développement des coopérations professionnelles

La coopération entre le monde de la recherche et celui de la production d'objets, d'outils de communication, de robots, est mise au service de la haute précision chirurgicale, de la cardiologie, de la dermatologie, du suivi et du traitement de maladies chroniques... Un développement rapide et continu des technologies se fait, en lien avec l'habitat intelligent et la domotique, les outils de compensation ou de stimulation, au service des personnes âgées ou en situation d'handicap.

La séquence de développement technologique fait un grand espoirs pour le progrès dans le mode d'exercice des médecins, la relation médecin-patient, et même le modèle de fonctionnement et de gestion du système de santé. L'amélioration de la collaboration entre les professionnels de la santé et du suivi de l'état de santé du patient et de ses traitements (carnet de santé électronique) permet l'obtention d'une meilleure connaissance et une meilleure interprétation des symptômes pathologiques [13].

1.3.1.2 Évolution des modalités de prise en charge des patients

La transmission d'informations à distance facilite la gestion de façon plus rapide et efficace des patients et leurs génère de moindres contraintes de déplacement. Ces développements qu'on appelle "médecine ambulatoire", répondent aux objectifs du système de santé et d'autre part aux attentes des patients qui préfèrent bénéficier de soins à leur domicile plutôt qu'en milieu hospitalier, qu'il s'agisse d'une période de convalescence ou du suivi régulier de maladies chroniques [13].

1.3.1.3 Un partage d'informations en croissance exponentielle et quasi instantané

A l'aide des traitements statistiques des données et des algorithmes, un diagnostic précoce des maladies de façon plus rapide et efficace est rendu possible. L'enregistrement

informatique des pathologies rencontrées et des soins rend potentiellement disponible une véritable (bibliothèque) de données ("big data"). Les informations médicales sont distribuées sur le web, ce qui permet leur disponibilité à une plus grande communauté de chercheurs dans les différents domaines liés à la santé et par conséquent un développement plus rapide des techniques utilisées et des services fournis [13].

1.3.2 Enjeux de l'efficacité administrative

L'enjeu principal pour les administrations des institutions de santé c'est la rapidité des services et l'efficacité pour donner des procédures administratives tout en réduisant le nombre d'étapes intermédiaires.

L'objectif est de faciliter l'accès aux soins, les inscriptions et même le remboursement des patients d'une façon électronique au lieu des transactions papiers. La majorité des programmes gouvernementaux dans le domaine de santé vise à favoriser l'échange d'informations et à aider les professionnels à axer leurs efforts sur les soins et les traitements, plutôt que sur la gestion des dossiers administratifs, notamment à travers les points suivants :

- La rapidité d'inscriptions dans les centres de traitements .
- Collecte plus facile des données administratives.
- Vérification aisée des droits des patients.

1.3.3 Enjeux du traitement des informations médicales

Le stade suivant dans la restructuration de la relation entre les patients, les professionnels de santé et les représentants du secteur public est à l'évidence l'introduction de technologies de transaction numérique ainsi que la création de cartes de santé électroniques. L'avantage majeur de ces cartes c'est l'accès direct à les informations médicale et administratives d'un patient. Éliminant les risques d'urgence tout en augmentant la performance des qualités des services. L'accès aux données personnelles d'un patient lié au processus d'authentification, ce fait de manière secrète et garantit la confidentialité des identités des patients et aussi des équipes médicale.



FIGURE 2 – Accès par carte au service médical [13]

1.4 Les grands domaines de l'informatique médicale

L'utilisation de l'informatique et des TIC dans la médecine et la santé a permis l'émergence d'un certain nombre de domaine de recherche dans les différents secteur relatifs.

La figure suivantes décrit les grands domaines de l'informatique médicale et les différents niveau micro et macro scopiques auxquels ils s'adressent.

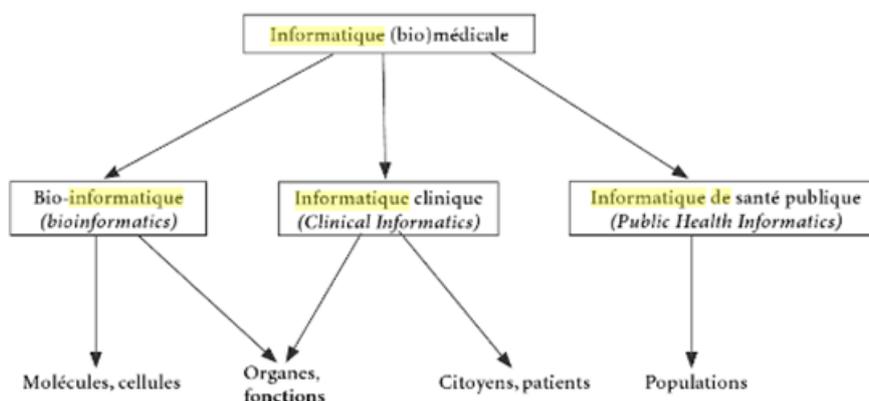


FIGURE 3 – Domaines de l'informatique médicale[1]

1.4.1 Bio-informatique

C'est une discipline hybride récente qui datte de quelques dizaines d'années et qui est fondée sur des concepts et des formalismes issus de la biologie, de l'informatique, des

mathématiques et de la physique. Elle utilise toutes les potentialités de traitement de l'informatique : modèles théoriques, algorithmes et programmes, ordinateurs, réseau Internet, bases de données ... On trouve deux principaux types de bio-information analysés par la bioinformatique : les séquences des nucléotides et les séquences des acides aminés [20].

1.4.2 L'informatique clinique ou bioclinique

Elle permet d'offrir des solutions méthodologiques et techniques pour la représentation des données, et aussi de traiter les données des patients et des connaissances médicales associées à la prise en charge individuelle des patients, leurs organisations, leur communications, la saisie et le stockage de leurs informations, et leur interprétation. La prise en charge aidée par l'informatique peut se faire en présence du malade ou à distance grâce aux outils de la télémédecine [1].

1.4.3 Informatique de santé publique

L'informatique de la santé publique regroupe les outils, techniques et applications informatiques permettant de raisonner non au niveau des individus mais des populations. La santé publique vise à développer des actions éducatives, préventives, curatives et sociales pour améliorer la santé globale des populations et faciliter la gestion de façon plus efficace. Ces actions doivent s'appuyer sur des systèmes d'information performants dont la conception relève de l'informatique médicale [1].

1.5 L'informatique médicale et l'informatique générale

Les développements associés à l'informatique médicale reposent sur les méthodes et les outils développés par les informaticiens, même si dans certains cas les problématiques de la santé sont originales et font progresser même les méthodes informatiques.[31] Dans l'informatique médicale on utilise des méthodes de l'informatique on citera :

1.5.1 Les bases de données

Pour informatiser les dossiers médicaux, l'utilisation des systèmes de gestion de bases de données(SGBD) est indispensable. Une des difficultés est liée à la nécessité d'élaborer et d'utiliser des systèmes terminologiques particuliers pour coder les symptômes, signes cliniques, pathologies et traitements des patients.

1.5.2 L'intelligence artificielle

La prise de décision est le rôle principal des médecins. Elle peut être effectuée après la numérisation en utilisant les méthodes intelligentes issues de l'intelligence artificielle. Les méthodes de l'ingénierie des connaissances servent pour la modélisation des connaissances médicales, les méthodes d'apprentissage automatique sont utilisées pour la fouille de données médicales, le traitement automatique du langage rend possible la structuration automatique d'informations (pathologies, traitements) contenues dans les comptes-rendus d'hospitalisation et les comptes-rendus opératoires.

1.5.3 Les réseaux

Les réseaux informatiques supportent les échanges de données et permettent de faciliter l'accès à un grand volume de données et contribuent à la disponibilité des données pour la prise en charge des patients. Bien que l'utilisation des réseaux informatiques pose des problèmes particuliers en médecine tels que la standardisation du contenu tenant compte de la sémantique des messages, sécurisation et protection de la confidentialité des données respectant les droits des patients.

1.6 Moyens informatiques pour le suivi de la santé

Pour pouvoir suivre la santé des individus en temps réel, plusieurs moyens doivent être utilisés. Premièrement, des dispositifs de mesures doivent être à la fois portés par les individus et connectés aux centres de suivi. Deuxièmement, des réseaux de communications doivent être utilisés pour établir la connexion entre les dispositifs connectés et les centres de suivi. Enfin, des systèmes d'informations doivent être développés pour le stockage des données collectées et permettre aux applications de gestion de fonctionner.

1.6.1 Dispositif connectés pour la santé

Plusieurs dispositifs ont été développés en collaboration entre médecins, biologistes, physiciens et électroniciens pour mesurer des indices de santé chez les individus à domicile ou en mobilité sans avoir besoin de se déplacer vers les centres médicaux. On trouve plusieurs types de ces dispositifs :

1.6.1.1 Dispositifs à la ceinture

La balise Aloïze est un boîtier qui se porte à la ceinture (Figure 4). Il est équipé d'un système GPS et d'une puce GSM, donnant la possibilité à la famille d'accéder via Internet à une carte qui indique la position du malade ou à un serveur téléphonique restituant par synthèse vocale sa localisation [36].

Avec une autonomie de quatre jours, la balise Aloïze à deux modes de fonctionnement : Aloïze, en mode veille, ne transmet rien : elle attend un ordre éventuel pour passer en mode actif (SMS). L'aidant ou le tuteur peut, à son initiative, demander la localisation du malade (remontée des coordonnées GPS par GPRS) qui est transmise soit par téléphone soit par Internet sous forme de cartes graphiques. La localisation donnée par le GPS est transmise en temps réel par SMS à la plateforme d'appel. Aloïze, en mode actif, transmet 10 fois par heure une adresse précise (GPS) et une restitution cartographique via Internet. Une option de sortie de zone est également disponible : 1 à 5 zones de vie avec alarme peuvent être programmées permettant d'être alerté en cas de sortie du patient de chaque périmètre défini.



FIGURE 4 – Balise Aloïze[36]

1.6.1.2 Téléphones mobiles

Dans le domaine de la santé numérique, les applications installés sur les smartphones connaissent des progrès très rapide ces derniers années pour devenir un facteur important pour le déploiement de la santé mobile. Les applications de la santé sont très nombreuses permettant de faciliter l'accès aux données, la consultation et la surveillance du patient, l'imagerie diagnostique etc. Parmi ces applications on cite :

1.6.1.2.0.1 Application CareZone CareZone est parmi les premières applications médicales sur smatphone. C'est une application de style tout-en-un qui facilite la gestion des médicaments de la famille et les doses concernant chaque médicament avec les instructions de suivi par le médecin puis des informations supplémentaire sur le médecin si nécessaire. Elle comprend un calendrier pour créer des horaires, un registre pour consigner les incidents qui peuvent arriver, et permet d'organiser les contacts des médecins et des pharmacies.

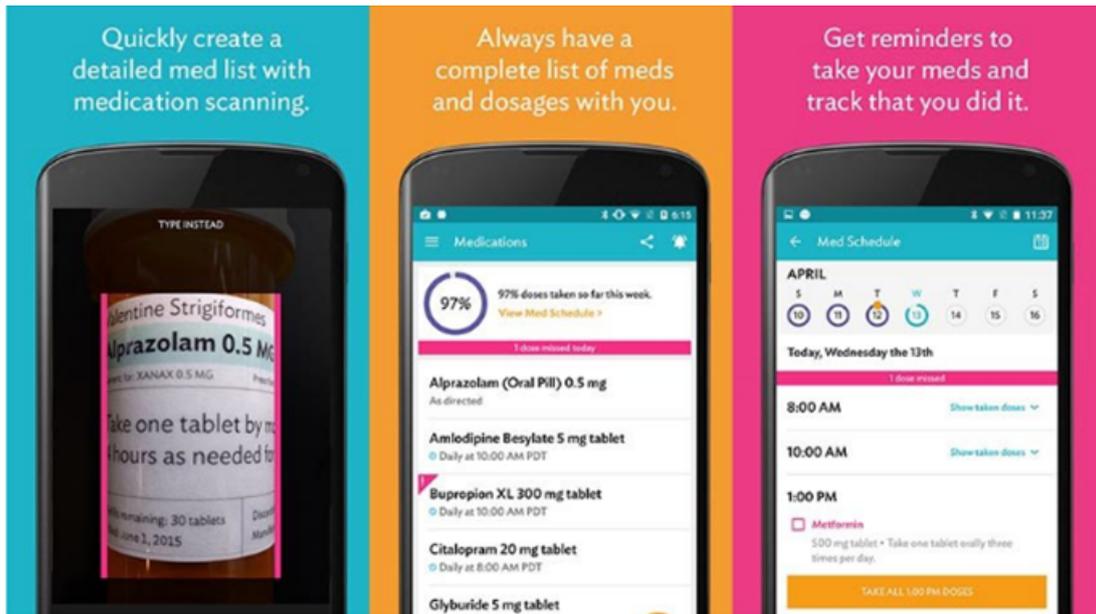


FIGURE 5 – Application CareZone[38]

1.6.1.2.0.2 Application Epocrates Epocrates est parmi les applications les plus utilisées pour les professionnels de la santé que par les personnes normales. L'application propose des informations cliniques sur des milliers de médicaments sur prescriptions, génériques et en libre service, incluant les interactions médicamenteuses et effets secondaires, pouvant aller jusqu'à 30 médicaments à la fois [38].

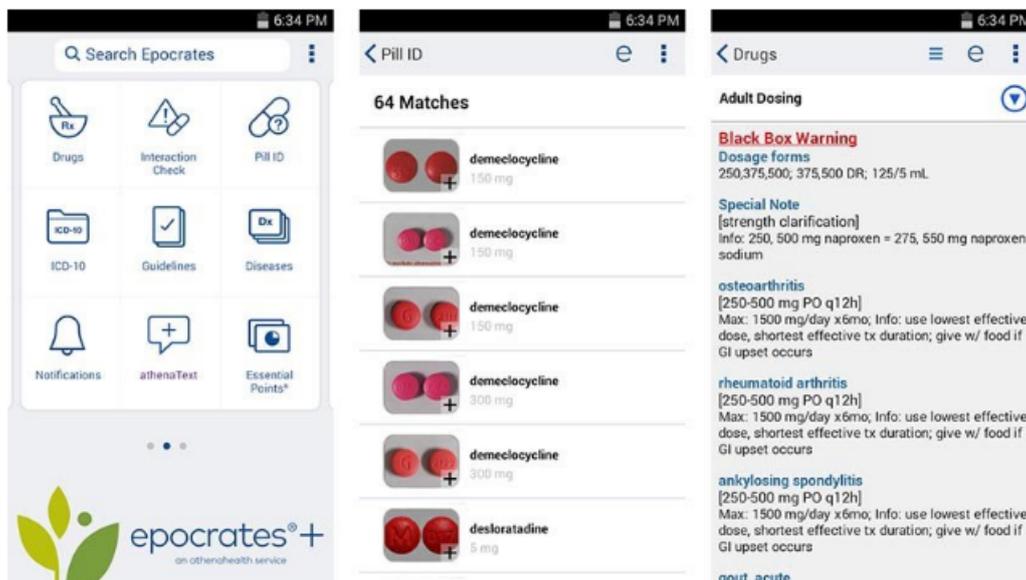


FIGURE 6 – Application Epocrates [38]

1.6.1.3 Médaillons

Les médaillons [8] sont connus depuis de nombreuses années. Simple d'utilisation, ils se présentent généralement sous la forme de bouton d'alarme. En cas de besoin, la personne surveillée appuie sur le bouton qui communique via des ondes-courtes avec un boîtier branché sur la ligne téléphonique standard de la maison. Le boîtier appelle le centre. Cependant, les systèmes de médaillon actuels comportent quelques limitations : Une zone de surveillance restreinte dans l'habitat (rayon 200m) : ils ne fonctionnent pas lorsque la personne va au jardin, rend visite à ses voisins ou fait ses courses. Ce genre de système n'est pas compatible avec les malades atteints de la maladie d'Alzheimer étant donné que l'alerte ne peut être déclenchée que manuellement.

1.6.1.4 Bracelet-téléphone

Le bracelet-téléphone de type Columba est un des premiers dispositifs. Il a été développé par Orange en collaboration avec la société canadienne (Medical Intelligence Technologies). Ce bracelet (Figure 7-1) est équipé d'un système de géolocalisation GPS-Assisté et d'un émetteur/récepteur GSM/GPRS, ainsi que d'un logiciel "intelligent" d'émission d'alertes, ce qui lui permet d'assurer trois fonctions principales :

- Déclenchement d'alertes en cas de sortie de zone (cercle de 500 mètres autour du domicile).
- La géolocalisation du porteur du bracelet,
- L'établissement d'un appel vers la famille ou le personnel médical.



FIGURE 7 – Le bracelet-téléphone Columba [47]

Ce dispositif n'est plus commercialisé à ce jour. Le bracelet BlueGard [7], remplaçant du Columba vise à garantir aux personnes vulnérables une autonomie et une sécurité de déplacement dans leur lieu de vie. Equipé de GPS et du GSM/GPRS, ce bracelet (Figure 8) peut alerter le personnel médical ou la famille si son porteur est localisé hors de la zone préalablement prédéfini (la distance peut varier de 500 mètres jusqu'à plusieurs kilomètres). A l'intérieur du domicile, le bracelet désactive le GPS afin de réduire la

consommation énergétique, pour ne laisser que la connexion RF (Radio Fréquence) avec une station d'accueil qui permet le rechargement de la batterie à distance. Cette méthode de rechargement évite à la famille de retirer le bracelet du poignet.



FIGURE 8 – Le bracelet BlueGard [7]

1.6.1.5 Montres intelligentes

Dans la littérature, on trouve plusieurs définitions d'une montre intelligente.

D'après McIntyre en 2014[32] : "Un dispositif permettant la surveillance médicale et le suivi des paramètres biométriques".

D'après Chuah en 2016[32] : "Un mini-appareil qui est porté et permet l'installation et l'utilisation d'applications".

D'après al Hsiao en 2017[32] : "SmartWatch sont des appareils qui peuvent se connecter avec les smartphones et recevoir beaucoup d'informations, tels que les messages textuels, les horaires et les données GPS. Alors qu'il peuvent effectuer des tâches de base relatives aux données et aux communications, il est également capable d'exécuter des applications mobiles".

Les montres connectées sont des montres intelligentes qui possèdent de multiples fonctions. En effet, ces outils sont les résultats de l'introduction des nouvelles technologies dans le secteur des montres. Avec ces outils de dernière génération, il n'est plus seulement question de mettre la montre pour se renseigner sur l'heure qu'il fait.

Les montres connectées ou Smart Watch sont pour la plupart équipés de Bluetooth, de connexion Wifi et de GPS afin de fournir des données réelles à leur utilisateurs quand ils sont activés[57].

L'évolution de ces montres intelligentes a touché également le domaine de la santé en proposant la prise de plusieurs mesures telles que le rythme cardiaque, la tension artérielle, la température du corps humain, la vitesse de mouvement, les pontes et les distances parcourues et dernièrement le signal cardiaque (ECG).

Ces mesures permettent de suivre en temps réel l'état de la santé des utilisateurs et particulièrement de ceux atteint de plusieurs maladies telles que les maladies cardiovasculaires,

le diabète, le parkinson,...etc. Actuellement, il existe plusieurs montres intelligentes qui proposent différentes mesures de santé.

1.6.1.5.0.1 Composants des montres intelligentes Les montres connectées consistent en un ensemble de capteurs qui les rendent intelligentes, parmi ces composants nous pouvons de citer :

- Capteurs de rythme cardiaque pour mesurer le nombre de battement du coeur,
- Capteur de contact pour mesurer la température de la peau,
- Capteur de la température ambiante ou extérieure,
- Altimètre pour mesurer l'altitude,
- Altitude gain ou perte pour mesurer les paliers d'escalier montés.
- GPS pour mesurer les coordonnées géographiques de l'utilisateur,
- Vélo /course à pied et distance.
- Accéléromètre pour mesurer la vitesse de déplacement.
- Capteur du nombre de pas(pédomètres).
- Capteur de l'état de la montre : portée ou non portée.
- Niveau global d'activité physique.
- Gyroscope pour détecter la posture de la montre afin d'optimiser son utilisation.

1.6.1.5.0.2 Sumsung Gear Sumsung Gear est parmi les montres intelligentes de haut niveau qui porte les meilleurs technologies selon ses caractéristiques techniques particulièrement dans le domaine de e-santé. Elle représente le premier concurrent de Applewatch sur le marché. Sumsung Gear utilise principalement le système d'exploitation android et compatible avec le système OS. Elle assure une connectivité avec d'autre appareil via le connecteur USB, le Wifi et les réseaux cellulaires. Elle contient des capteurs pour offrir différentes mesures telles que : le rythme cardiaque, le mouvement, le nombre de calories, la distance parcourue, le nombre de pas tout au long du jour, le niveau d'activité. La figure suivant représente un modèle d'une montre Sumsung Gear :



FIGURE 9 – Sumsung Gear S3[22]

1.6.1.5.0.3 Huawei watch La Huawei Watch est la première montre connectée de Huawei. Elle se distingue par son boîtier en acier et son cadran rond en cristal saphir réputé pour sa très haute résistance aux chocs.

Huawei a apporté de nombreuses améliorations à sa montre connectée, et le résultat est plus que réussi. La Huawei Watch possède un GPS, une connexion Wi-Fi, Bluetooth et 4G (en option). Elle s'utilise donc seule ou avec le smartphone. Grâce à ses nombreux capteurs, elle permet d'obtenir des données détaillées sur les performances physiques telles que la surveillance du rythme cardiaque en temps réel qui fournit une mesure personnelle en temps réel plus efficace et précise de la fréquence du rythme cardiaque en utilisant un algorithme d'auto-apprentissage et des capteurs innovants) et permet de le consulter directement sur la montre ou via l'application Huawei Health.

La nouvelle montre connectée avec son système d'exploitation android, attire clairement une clientèle plutôt sportive : les fonctions dédiées au sport sont mises en avant. La figure suivant présente un modèle d'un Huawei watch :



FIGURE 10 – Huawei watch [39].

1.6.1.5.0.4 Apple watch L'une des premières montres intelligentes apparues, elle contient plusieurs fonctionnalité : des nouveaux capteurs cardiaques électriques, couronne numérique repensée avec retour haptique. Apple Watch Series 4 redéfinit la norme pour ce que peut être une montre.

Elle inspire de saines habitudes et aide à gérer tous les aspects de la santé de l'être humain, du stress aux dépenses caloriques. Elle prend soin du cœur en alertant la personne en cas d'anomalie et elle veille sur la sécurité comme jamais avec la détection de chute.

Apple Watch permet le suivi de l'état de la personne du matin au soir. Elle aide de prendre la fréquence cardiaque à tout moment et faire le suivi d'un ensemble de données. Elle reçoit aussi une alerte dès qu'une fréquence anormalement basse ou élevée est détectée même si la personne n'éprouve aucun symptôme, les valeurs obtenues sont parfois signe d'une maladie grave. Mais ces symptômes passent souvent inaperçus, ce qui empêche d'en diagnostiquer les causes.

Apple Watch Series 4 veille sur le cœur et alerte au besoin pour prendre des mesures immédiates et consulter un médecin pour la confirmer. Parmi ses fonctionnalités majeures est la détection de chute qui permet de signaler la position du patient et alerte un des appel d'urgence parmi les contacts de la montre. En plus elle apprend son porteur de bonnes habitudes telles que boire régulièrement, penser à faire des pauses et bouger, encourager à relaxer par des exercices de respiration, elle ajout aussi la fonction talkie-walkie permettant la communication avec d'autres utilisateurs de l'Apple Watch via le Wi-Fi ou le réseau cellulaire [4, 57].



FIGURE 11 – Montre de santé Apple watch Series 4. [4]

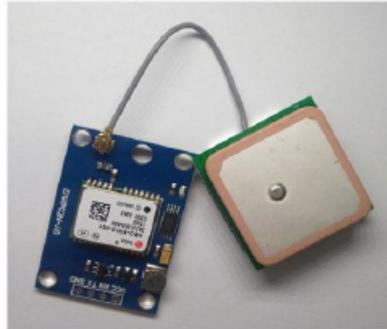
1.6.2 Informations de santé mesurables par moyens mobiles connectés

1.6.2.1 Position

La technologie des montres connectées permet de suivre en temps réel les utilisateurs en vue map ou satellite en calculant leurs coordonnées GPS.

Un capteur GPS (figure 3a) est un capteur à trois axes utilisé en navigation spatiale qui

peut déterminer l'emplacement, l'altitude et la vitesse à tout moment et dans la plupart des conditions météorologiques. D'après son état médical et en cas d'urgence, il est très important de sélectionner la position du patient et déclencher une appel d'urgence via les contacts existant dans la montre.



(a)

FIGURE 12 – Capteur GPS [32]

1.6.2.2 Mouvement

Parmi les fonctions de la montre intelligente est la mesure les mouvements des athlètes à chaque période et le nombre de pas de marche en temps réel. Cette montre connectée sport a une puce GPS intégrée qui affiche une carte de votre emplacement. Il peut détecter la distance, la trajectoire du mouvement.

1.6.2.3 Chute

Grâce à son accéléromètre et à son gyroscope, les smart watch permettent de détecter les chutes. La montre affiche alors une alerte et permet de joindre facilement les services d'urgence. Si la personne ne réagit pas dans les 60 secondes ou moins, elle envoie un message indiquant sa position géographique à un contact d'urgence [57].

1.6.2.4 Tension Artérielle

la montre combine des capteurs EKG (électrocardiogramme) et PPG (photopléthysmographie), ce qui lui permet de mesurer la tension artérielle de son utilisateur en temps réel.



FIGURE 13 – Montre de santé intelligente pour mesurer la tension artérielle.

1.6.2.5 Rythme cardiaque

La plupart des montres intelligentes sont intéressées de mesurer les maladies cardiovasculaires à travers le nombre de battements du cœur, si elle est basse ou élevée sont parfois signe d'une maladie grave. Mais ces symptômes passent souvent inaperçus, ce qui empêche d'en diagnostiquer les causes. Smart Watch alerte de façon automatique l'utilisateur sur l'état médical au besoin pour qu'il puisse prendre des mesures immédiates et consulter un médecin.

Le capteur de fréquence cardiaque à pince digitale illustré à la figure 2b est un capteur optique haute performance qui mesure le changement dans le mouvement du sang dans le corps. Les capteurs sont courants en médecine des composants électroniques destinés à être utilisés à l'intérieur pour surveiller la santé du patient [9].

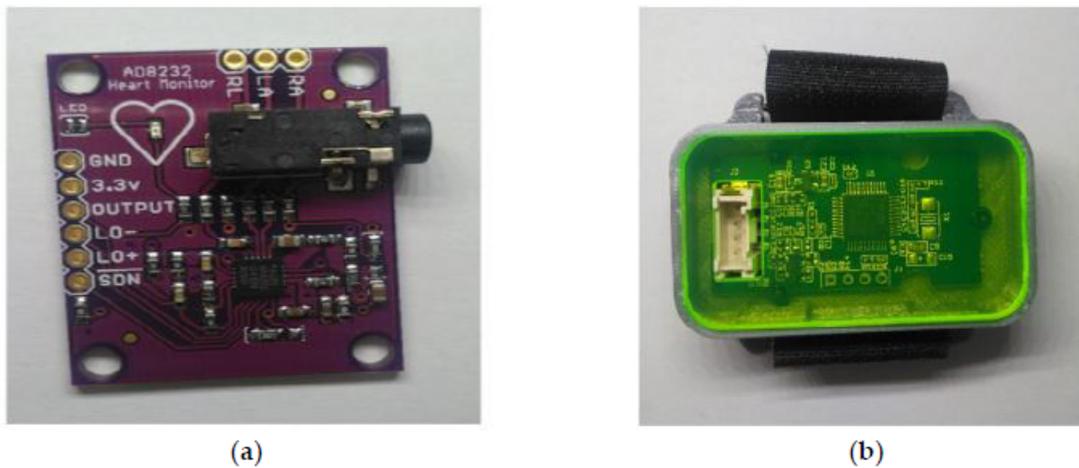


FIGURE 14 – Biocapteurs optiques : a) capteur de surveillance cardiaque à électrocardiogramme (ECG);(b) capteur de fréquence cardiaque à pince doigt

1.6.2.6 Signal cardiaque

L'électrocardiogramme (ECG) est un signal qui représente l'activité électrique du cœur. L'ECG est un élément essentiel que ce soit dans la surveillance des patients ou dans le diagnostic des maladies cardiovasculaires. La figure suivante présente un signal électrocardiogramme (ECG) :

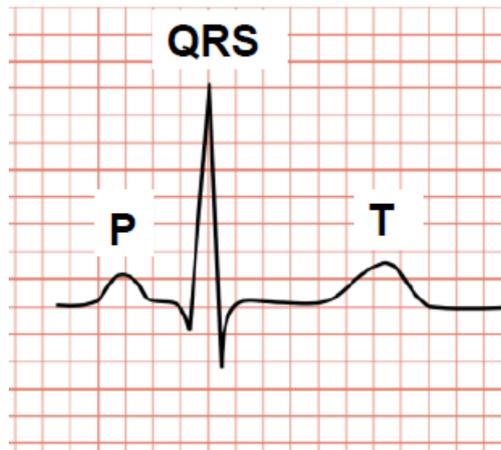


FIGURE 15 – Signal électrocardiogramme (ECG) [41]

Un smartwatch récent peut afficher avec précision l'activité électrique du cœur grâce aux capteurs spécifique et notifier aux personnes atteintes de ces maladies.

1.6.2.7 Température

Les montres connectées prennent aussi via des capteurs la valeur de la température corporelle.

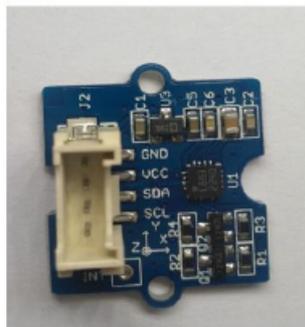


FIGURE 16 – Capteur de température du corps humain [9].

1.6.2.8 Niveau de glucose

A travers des capteurs spécifiques de mesure du glucose, les montres intelligentes permettent la surveillance de la glycémie et suivre l'état du diabète chez un patient à tout moment.

Il est important que la glycémie reste dans une fourchette saine. Si la glycémie devient trop basse, nous pouvons perdre la capacité de penser et de fonctionner normalement. S'ils deviennent trop hauts et restent élevés, cela peut causer des dommages ou des complications au corps au cours de nombreuses années [45].



FIGURE 17 – Montre intelligente(K'Watch) pour mesurer le niveaux de glucose [32].

1.6.2.9 Sommeil

Le suivi du sommeil permet de mesurer la qualité de sommeil d'un individu pour connaître les différents états qui se passent durant le sommeil. Pendant la nuit, le corps passe par plusieurs phases de sommeil, grâce à des capteurs sur les montres connectées (accéléromètre ou sonar) on peut créer un graphique de ces phases. La mesure des cycles de sommeil permet de [55] :

- Visualiser le processus de sommeil dans un graphique de sommeil et des hypnogrammes.
- Permet de calculer le pourcentage de sommeil profond et le nombre de cycles, ce qui peut être un indicateur d'un sommeil sain ou en mauvaise santé.
- Réveil intelligent, elle permet de trouver le meilleur moment pour que l'alarme réveille l'individu.
- Collecter des statistiques pour analyser le sommeil à long terme.
- Fixez les objectifs et améliorer les paramètres de sommeil.

1.6.3 Système d'information médical

Les TIC (Technologies de l'information et de la Communication) regroupent un large ensemble de médias de communication et d'appareils qui relient entre eux les systèmes d'information (SI) et les individus (messageries vocales, internet, intranet, équipements pour visioconférence, smartphones, tablettes, . . .). Le schéma ci-après décrit les principaux domaines couverts.

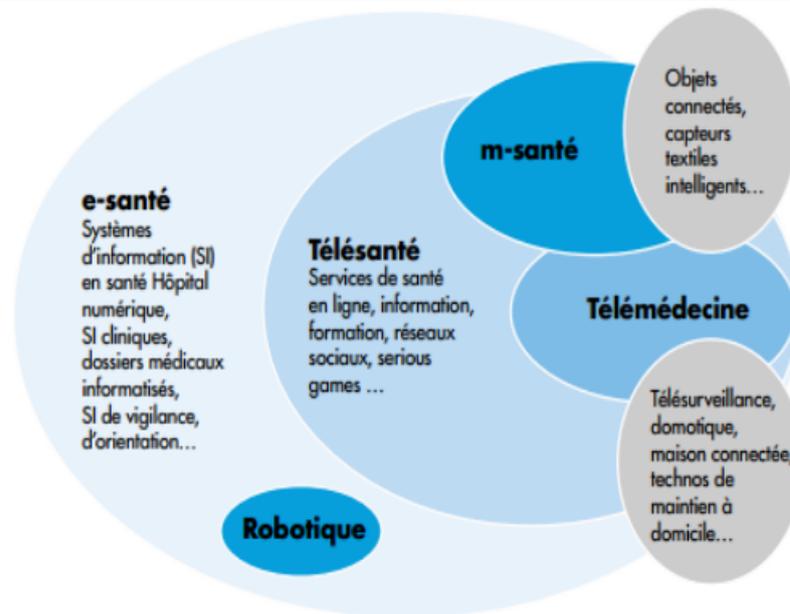


FIGURE 18 – Principaux domaines de la santé connecté [13].

Les SI soulèvent, dans le domaine de la santé, des problèmes de sécurité et d'interopérabilité. Leurs usages et leurs fonctionnalités concernent concrètement trois grands domaines :

- La télé-santé,
- Les applications en télé-médecine,
- Les données médicales : le partage de l'information pour une meilleure coordination des acteurs de santé, et l'exploitation de ces données dans le domaine de la recherche médicale.

1.6.3.1 Télé-santé

La télésanté est l'utilisation des outils de production, de transmission, de gestion et de partage d'informations numérisées au bénéfice des pratiques tant médicales que médico-sociales. Pour ce qui concerne le domaine médico-social,[37], la télé-santé comprend aussi bien les aspects de promotion de la santé et de santé publique via le web que la domotique, les services santé, bien-être , (télé-conseil sur le web, réseaux sociaux, forums," jeux sérieux" utilisés souvent dans le monde de la rééducation, ou à but pédagogique)...

Les personnes atteintes de maladies chroniques utilisent de plus en plus ces applications pour être en lien avec leur médecins, surveiller leur santé, être encouragées dans leurs efforts, se mettre en contact avec les services d'urgence.

1.6.3.2 Télé-médecine

La télé-médecine est une pratique médicale qui utilise différentes technologies de télécommunication afin d'offrir des diagnostics médicaux ou encore des soins de santé à distance. Son objectif est de faciliter l'accès aux services médicaux dans les collectivités éloignées ou en milieu rural.

1.6.3.3 M-santé (mobile-santé)

La m-santé représente des applications disponibles sur les smartphones, tablettes permettant d'offrir des services de façon simple pour les médecins pour suivre leurs patients à distance [13]. Ces usages ont tout leur intérêt, notamment en termes de prévention et permettent des avantages majeurs dont nous citons :

- Offrir un suivi en temps réel des patients.
- Obtenir un diagnostic rapide et précis pour améliorer le résultat des traitements[19].

1.7 Architectures et systèmes d'information pour la e-santé

Dans la littérature, on trouve plusieurs modèles de système d'information mettant en œuvre la santé numérique et chacun à des caractéristiques pour garantir un fonctionnement efficaces de ces derniers :

1.7.1 Caractéristiques d'un système pour la e-santé

Dans les systèmes actuels appliqués dans le domaine de la santé médicale sur les montres intelligentes, qui rassemblent un ensemble des critères pour assurer un suivi efficace de la situation individuelle, nous pouvons citer les éléments suivants :

- Une consommation faible de la batterie.
- Le facteur de la forme, Il est confortable pour l'utilisation grâce à sa petite forme.
- La fiabilité du système.
- Qualité de service.
- Expérience utilisateur enrichie.
- Une plus grande efficacité.
- Capacité d'interopérabilité sur différentes plateformes.
- Facilité de déploiement.
- Popularité du système de soins de santé intelligent pour offrir des soutiens.
- l'évolutivité du système pour passer aux nouvelles versions et les technologies.
- Connectivité suffisante.

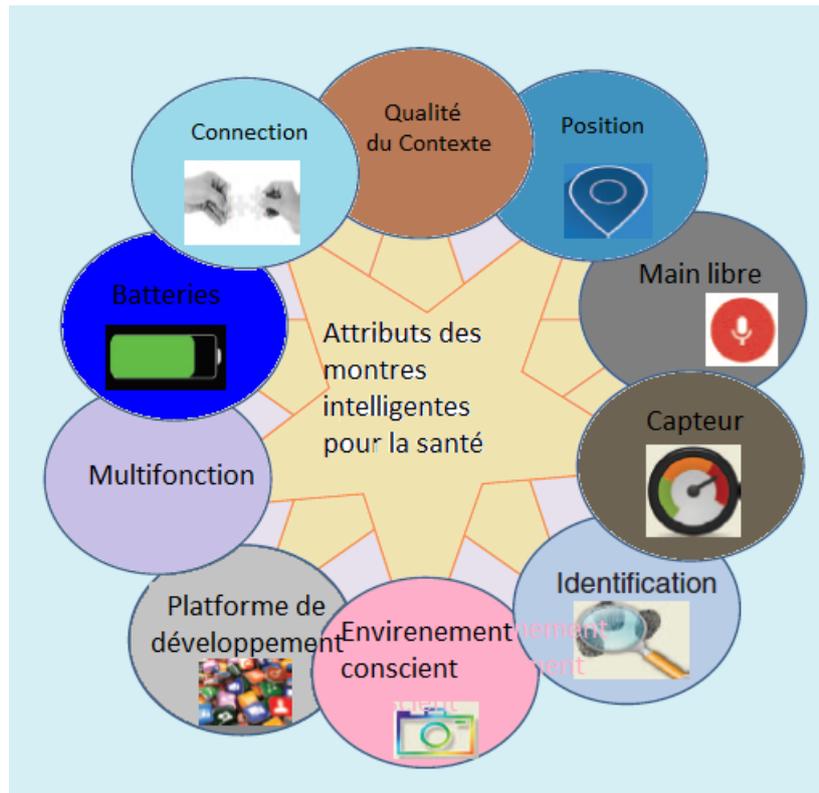


FIGURE 19 – Caractéristiques des montres intelligentes pour la santé [9]

1.7.2 Systèmes Mobiles de Surveillance de Patients (SMSP)

Les auteurs de [49] proposent un système de santé numérique à base de PC et PDA qui représente un système mobile de surveillance de patients et qui rassemble les données physiologiques des patients par les capteurs/biocapteurs. Les données sont acheminées au PC ou à la cellule phone/PDA d'un patient. Après une analyse de situations critiques et détermination d'alerte possible, ces dispositifs transmettent ces données au serveur pour l'analyse complète. Après que les données soient analysées, le serveur médical fournit les résultats aux PCs ou aux cellules phone/PDA des patients. Les patients peuvent prendre des mesures nécessaires selon ces résultats. Sur le plan structurel, le SMSP proposé est constitué de trois composants essentiels (Figure 20) :

- Réseau local sans fil (RLSF).
- Serveur local à domicile (SLD) accessible par le patient, son assistant ou un membre de sa famille.
- Serveur médical distant (SMD) accessible par le médecin ou tout autre autorisé.

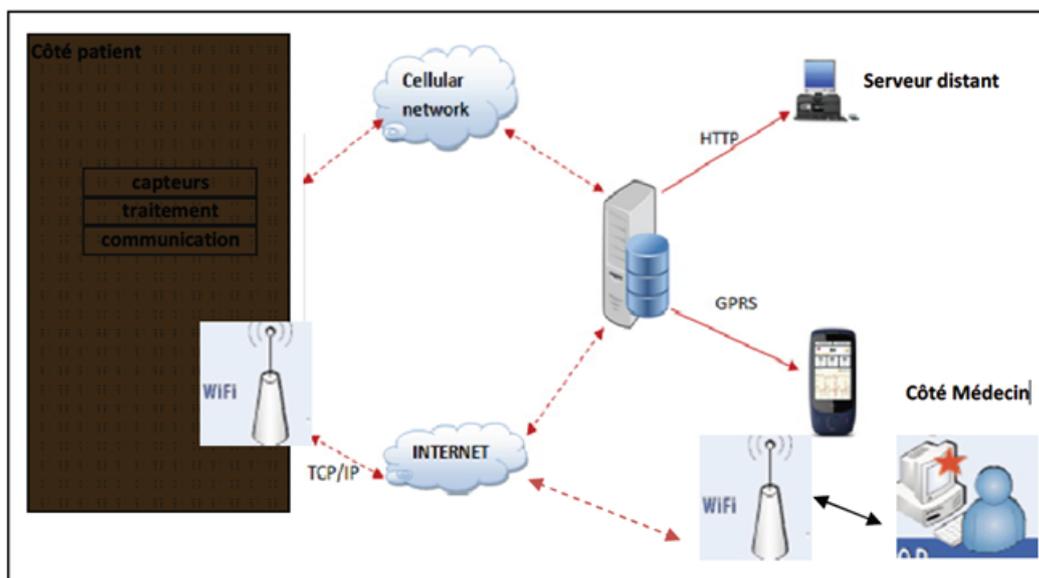


FIGURE 20 – Système mobile de surveillance de Patient(SMSP)[49]

1.7.2.1 Réseau local sans fil (RLSF)

Le réseau local sans fil permet la communication entre les appareils de mesure des paramètres physiologiques du patient et la cellule locale (Smartphone, PDA, tablette, ou PC). La technologie réseaux Bluetooth est utilisée pour des raisons de disponibilité, de faible coût, d'efficacité et de simplicité.

1.7.2.2 Serveur local à domicile (SLD)

Le serveur local à domicile (SLD) du patient représente les différents périphériques mobiles ou PC, les plus utilisés et appropriés que les utilisateurs préfèrent et avec lesquels ils se sont familiarisés. Les capteurs sont des outils dont la fonction principale est l'acheminement d'informations aux smart-phones et la collecte de façon continue des données physiologiques des patients. Ils sont dotés de technologies de communication sans fil (Bluetooth). A travers une analyse, on extrait des informations pertinentes permettant aux patients le suivi d'un processus de soins de santé par leurs périphériques mobiles. Dans les cas d'anomalie, le patient et le médecin sont alertés en même temps à partir du serveur distant(SMD).

1.7.2.3 Serveur médical distant (SMD)

Le serveur médical distant est un serveur dédié au médecin qui peut l'installer au niveau de son cabinet ou clinique. Il importe ses données des périphériques mobiles (tels que tablette, smartphone, et même PC) du patient connectés par moyen de communication sans fil. De plus, l'information fournie par les spécialistes dans les scénarios critiques peut

alimenter le serveur. Le serveur médical distant(SMD) est l'organe principal de cette architecture, il reçoit les données enregistrées des patients en provenance des SLDs. Il est capable de déterminer les seuils spécifiques des paramètres physiologiques des patients. Tous les renseignements du patient sont stockés dans une base de données du SMD et peuvent être envoyés aux SLDs dans le cas de consultation par le patient.

1.7.3 Système utilisant des capteurs

Les auteurs de [24] proposent une architecture permettant de quantifier les rythmes cardiaques et respiratoires selon certaines conditions de mouvement. Les éléments utilisés sont : ordinateurs(PC), webCam, Capteurs PPG au doigt. Leur fonctionnement est à l'intérieur et relatif aux sources d'éclairage.

Un capteur est placé sur le doigt d'un patient et mesure simultanément le rythme cardiaque et les amplitudes du signal PPG. Toutes ces données sont enregistrées au niveau de PC. La figure suivante décrit l'architecture proposé :

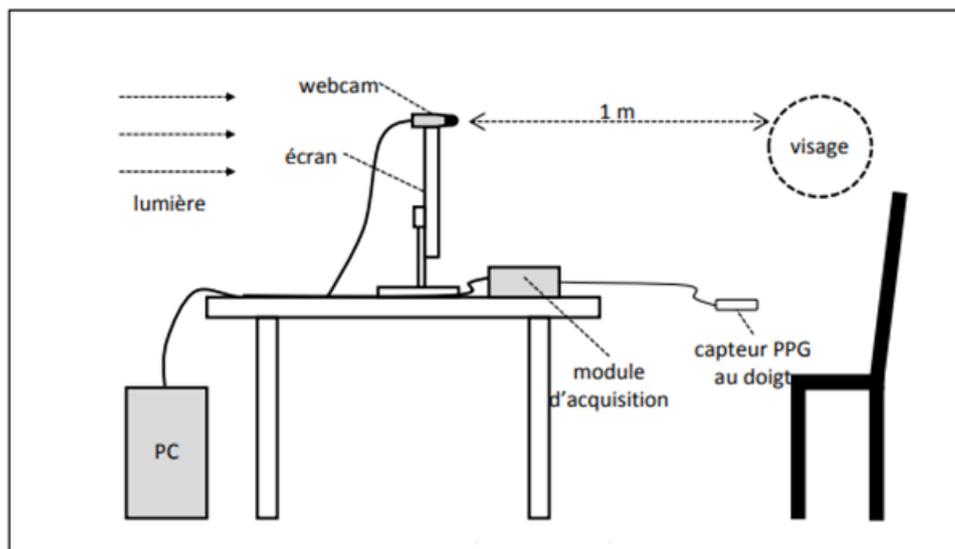


FIGURE 21 – Système d'information utilisant des capteurs[24]

1.8 Conclusion

Dans ce chapitre, nous avons présenté l'importance du domaine de la santé numérique ainsi que les différents moyens informatique nécessaire à sa mise en place. En effet, les dispositifs mobiles connectés telles que les montres intelligentes ainsi que les réseaux de communication et les systèmes d'informations sont des éléments nécessaire pour construire des systèmes informatique pour la e-santé.

Nous avons également présenté les montres intelligentes et les possibilités qu'elles

offrent pour la prise de plusieurs mesures de santé telles que le rythme cardiaque, les mouvements, le niveau du glucose,...etc.

Grâce aux réseaux de communication auxquels ces montres sont connectés, les données relatives à la santé peuvent être collectées dans des centres d'analyse et utilisées pour une meilleure prise de décision médicale.

Dans le chapitre suivant, nous allons présenter le domaine de data mining rassemblant les méthodes d'analyse intelligentes de données et que nous prévoyons utiliser pour concevoir notre système de suivi médical à distance.

DATA MINING

2.1 Introduction

Avec la croissance rapide des nombres d'habitants dans le monde pour atteindre selon les statistiques de l'ONU un nombre 9,4 milliards en 2050, le suivi médical de la santé des citoyens devient de plus en plus difficile.

Le développement remarquable dans les moyens technologiques surtout dans les domaines des montres intelligentes connectés et les réseaux de communications permettent de suivre en temps réel plusieurs indices de santé des individus. Les données obtenues à travers ces moyens peuvent être stockées et analysées dans des centres d'analyse par des méthodes intelligentes (de data mining) pour une prise de décision médicale plus rapide et efficace et répondant aux nouvelles conditions relatives à la croissances de la population.

Dans ce chapitre nous allons présenter les concepts de base ainsi que la classification de ces méthodes.

2.2 Définition d'un Data Mining

Le Data mining, est un domaine qui est apparu avec l'explosion des quantités de données stockées, Il consiste à l'analyse de grands ensembles de données observationnelles pour découvrir des modèles et des relations intéressantes et utiles. Le domaine combine des outils de statistiques et d'intelligence artificielle (tels que les réseux de neurones, les SVM, les arbres de décisions,..) avec la gestion de bases de données pour analyser de grandes collections numériques, appelées ensembles de données. L'exploration de données (fouille de données) est largement utilisée dans les domaines des affaires (assurances, banque, commerce de détail), de la recherche scientifique (astronomie, médecine) et de la sécurité publique (détection des criminels et des terroristes) [14].

L'objectif global du processus d'exploration de données est d'extraire des connaissances à partir d'un ensemble de données et de les transformer en une structure compréhensible pour une utilisation ultérieure.

La tâche d'exploration de données consiste en l'analyse automatique ou semi-automatique de grandes quantités de données afin d'extraire des modèles auparavant inconnus et intéressants, tels que des groupes d'enregistrements de données (analyse de grappes), des enregistrements inhabituels (détection d'anomalies) et des dépendances (exploration de règles d'association) [16].

Grâce aux techniques d'extraction de connaissances, les bases de données volumineuses sont devenues des sources riches et fiables pour la génération et la validation de connaissances. La fouille de données n'est qu'une phase du processus d'ECD, et consiste à appliquer des algorithmes d'apprentissage sur les données afin d'en extraire des modèles (motifs). L'extraction de connaissances à partir des données se situe à l'intersection de nombreuses disciplines [34], comme l'apprentissage automatique, la reconnaissance de formes, les bases de données, les statistiques, la représentation des connaissances, l'intelligence artificielle, les systèmes experts, etc.(figure 3-1)

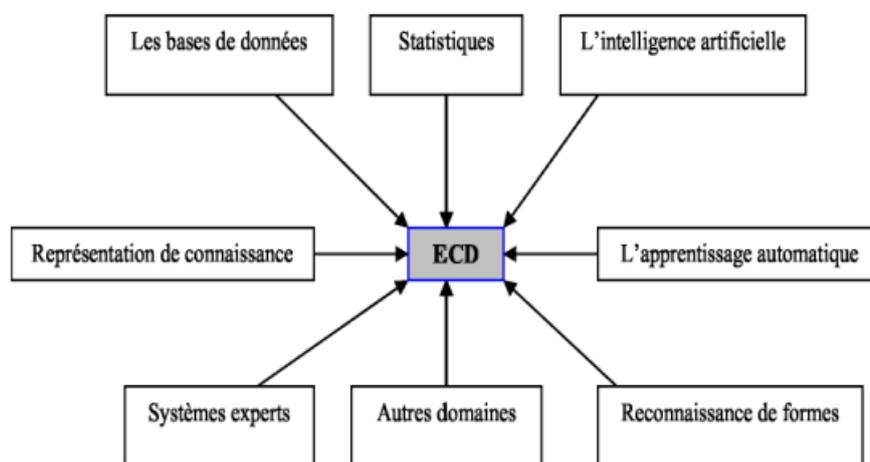


FIGURE 1 – Les différents domaines de Data mining [34].

2.3 Les domaines d'application de Data Mining

La fouille de données a aujourd'hui une grande importance économique du fait qu'elle permet d'optimiser la gestion des ressources (humaines et matérielles). Elle est utilisée par exemple dans différents domaines on cite :

2.3.1 Le data mining dans la banque

Dans les banques, le data mining est utilisé notamment pour :

- Prédire la réaction des clients aux changements des taux d'intérêt.
- Identifier les clients qui seront les plus réceptifs aux nouvelles offres de produits.
- Déterminer les clients qui posent le risque le plus élevé de manquer à leurs engagements aux prêts.
- Détecter les activités frauduleuses dans les transactions par cartes de crédit.
- Prédire les clients qui sont susceptibles de changer leurs cartes d'affiliation au cours du prochain trimestre.

2.3.2 Le data mining dans l'assurance

Dans l'assurance, On peut traduire le data mining sur différents manière :

- Des produits obligatoires (automobile, habitation) :
- Analyse des sinistres.
- Recherche des critères explicatifs du risque ou de la fraude,
- Extraction de caractéristiques à partir de ces déclarations (type d'accident, de blessures, etc.)

2.3.3 Le data mining dans le commerce

Dans le domaine du commerce, l'extraction de données est appliquée pour :

- L'évaluation des risques pour chaque transaction.
- Connaitre la distribution et l'emplacement géographique des clients.
- L'analyse de la fidélité des clients dans les opérations à base de crédit.
- Évaluation de la menace concurrentielle dans une région.

2.3.4 Le data mining dans la détection de fraude

Pour la detection de fraude, on utilise le data mining pour :

- Détection de fraude de cartes de crédits.
- Détection de fraude dans les listes des électeurs en utilisant les réseaux de neurones en combinaison avec le data mining.
- La détection des fraudes dans les demandes de passeport par la conception d'un système de diagnostic par apprentissage en ligne.
- Détection de fausses demandes de remboursement médicale.

2.3.5 Le Data mining dans la médecine et la pharmacie

Dans le domaine médical, le data minig est utilisée pour :

- Déterminer des segments de patients susceptibles d'être soumis à des protocoles thérapeutiques déterminés, chaque segment regroupant tous les patients réagissant identiquement.
- Mettre en évidence des facteurs de risque ou de rémission dans certaines maladies. Choisir le traitement le plus approprié.
- Optimisation des plans d'action des visiteurs médicaux pour le lancement de nouvelles molécules.
- Prédire les effets sur la peau humaine de nouveaux cosmétiques, en limitant le nombre de tests sur les animaux.

2.4 Processus de Data Mining (ECD)

"L'Extraction de Connaissances à partir des Données (ECD) est un processus itératif et interactif d'analyse d'un grand ensemble de données brutes afin d'en extraire des connaissances exploitables par un utilisateur analyste qui y joue un rôle central" [61].

D'après[21], un processus d'ECD est constitué de quatre phases qui sont : le nettoyage et intégration des données, le pré-traitement des données, la fouille de données et enfin l'évaluation et la présentation des connaissances. La figure 1.1 présente ces différentes étapes ainsi que les enchaînements possibles entre ces étapes. Cette séparation est théorique car en pratique, ce n'est pas toujours le cas. En effet, dans de nombreux systèmes, certaines de ces étapes sont fusionnées [33] .

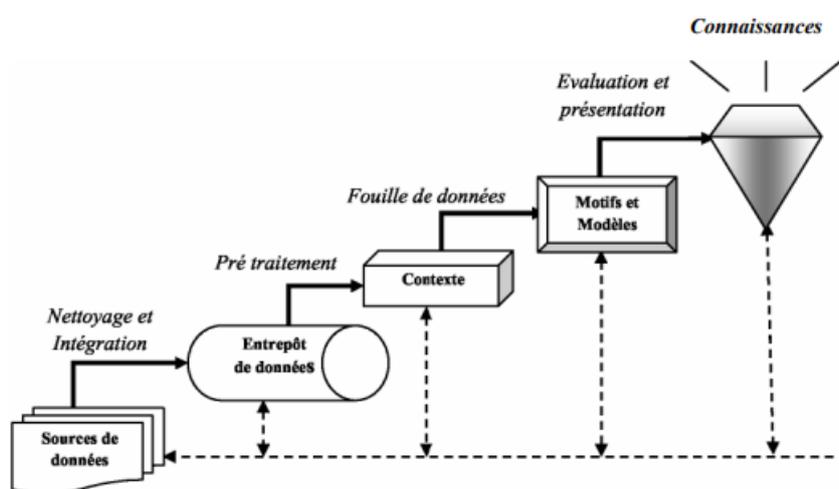


FIGURE 2 – Processus d'extraction de connaissances à partir des données [21].

2.4.1 Nettoyage et intégration des données

Le nettoyage des données consiste à retravailler ces données bruitées, soit en les supprimant, soit en les modifiant de manière à tirer le meilleur profit. L'intégration est la combinaison des données provenant de plusieurs sources (base de données, sources externes, etc). Le but de ces deux opérations est de générer des entrepôts de données et/ou des magasins de données spécialisés contenant les données retravaillées pour faciliter leurs exploitations futures.

2.4.2 Pré-traitement des données

Il peut arriver parfois que les bases de données contiennent à ce niveau un certain nombre de données incomplètes et/ou bruitées. Ces données erronées, manquantes ou connaissances inconsistantes doivent être retravaillées si cela n'a pas été fait précédemment. Dans le cas contraire, durant l'étape précédente, les données sont stockées dans un entrepôt. Cette étape permet de sélectionner et transformer des données de manière à les rendre exploitables par un outil de fouille de données. Cette seconde étape du processus d'ECD permet d'affiner les données. Si l'entrepôt de données est bien construit, le pré-traitement de données peut permettre d'améliorer les résultats lors de l'interrogation dans la phase de fouille de données.

2.4.3 Fouille de données

La fouille de données (data mining en anglais), est le cœur du processus d'ECD. Il s'agit à ce niveau de trouver des pépites de connaissances à partir des données. Tout le travail consiste à appliquer des méthodes intelligentes dans le but d'extraire cette connaissance. Les méthodes d'apprentissage automatique (des champs d'étude de l'intelligence artificielle) sont les plus utilisées. Il est possible de définir la qualité d'un modèle en fonction de critères comme les performances obtenus, la fiabilité, la compréhensibilité, la rapidité de construction et d'utilisation et enfin l'évolutivité. Tout le problème de la fouille de données réside dans le choix de la méthode adéquate à un problème donné. Il est possible de combiner plusieurs méthodes pour essayer d'obtenir une solution optimale globale [16].

2.4.4 Évaluation et présentation

Cette phase est constituée de l'évaluation, qui mesure l'intérêt des motifs extraits, et de la présentation des résultats à l'utilisateur grâce à différentes techniques de visualisation. Cette étape est dépendante de la tâche de fouille de données employée. En effet, bien que l'interaction avec l'expert soit importante quelle que soit cette tâche, les techniques ne sont pas les mêmes. Ce n'est qu'à partir de la phase de présentation que l'on peut employer le

terme de connaissance à condition que ces motifs soient validés par les experts du domaine. Il y a principalement deux techniques de validation qui sont la technique de validation statistique et la technique de validation par expertise.

2.4.4.1 La validation statistique

Consiste à utiliser des méthodes de base de statistiques descriptives. L'objectif est d'obtenir des informations qui permettront de juger le résultat obtenu, ou d'estimer la qualité ou les biais des données d'apprentissage. Cette validation peut être obtenue par :

- Le calcul des moyennes et variances des attributs,
- si possible, le calcul de la corrélation entre certains champs,
- ou la détermination de la classe majoritaire dans le cas de la classification.

2.4.4.2 La validation par expertise

Elle est réalisée par un expert du domaine qui jugera la pertinence des résultats produits. Par exemple pour la recherche des règles d'association, c'est l'expert du domaine qui jugera la pertinence des règles. Pour certains domaines d'application (le diagnostic médical, par exemple), le modèle présenté doit être compréhensible. Une première validation doit être effectuée par un expert qui juge la compréhensibilité du modèle. Cette validation peut être, éventuellement, accompagnée par une technique statistique [11].

2.5 Principales tâches de fouille de données

On dispose de données structurées. Les objets sont représentés par des enregistrements (ou descriptions) qui sont constitués d'un ensemble de champs (ou attributs) prenant leurs valeurs dans un domaine. De nombreuses tâches peuvent être associées au Data Mining, nous pouvons citer par exemple :

- 1- La classification.
- 2- L'estimation.
- 3- La prédiction.
- 4- Les règles d'association
- 5- La segmentation.

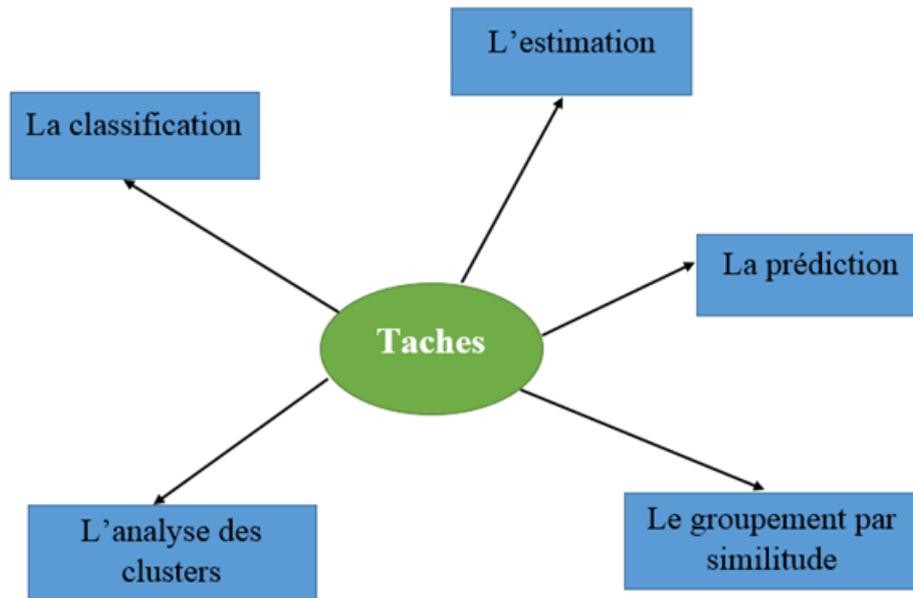


FIGURE 3 – Tâches de fouille de données.

La figure suivante présente les techniques de data mining et ses différentes méthodes :

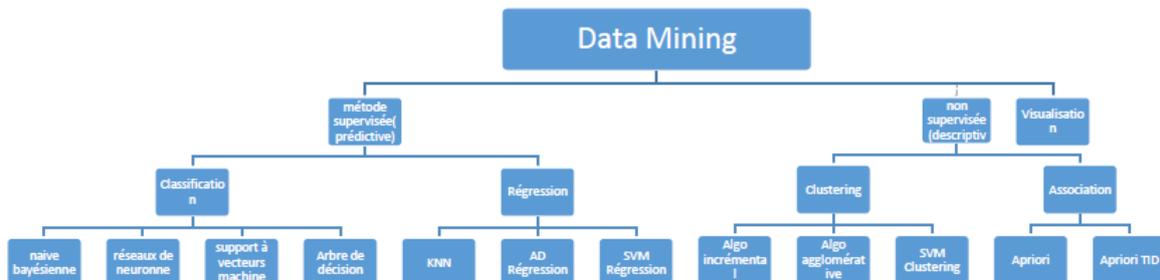


FIGURE 4 – Technique de data mining [53].

2.5.1 Classification

La classification est la tâche la plus commune de la fouille de données qui semble être une tâche humaine primordiale. Elle consiste à examiner les caractéristiques d'un objet et lui attribuer une classe, la classe est un champ particulier à valeurs discrètes. Il existe deux types :

2.5.1.1 Classification supervisée (apprentissage)

Elle sert à obtenir un critère de séparation destiné à prédire l'appartenance à une classe tel que le nombre de classe est défini a priori. On présente quelque méthode de classification supervisée :

- Analyse discriminante.
- k-Plus proches voisins.
- Arbre de décision.
- Réseau de neurones.

2.5.1.2 Classification non supervisée

Elle sert à établir des représentations des données dans des espaces à faible dimensions pour y lire des typologies d'individus tel que le nombre de classes n'est pas connu. Parmi les méthodes de classification non supervisée en trouve :

- Analyse en composante principales (ACP)[40].
- Analyse des associations.
- Analyse en clusters.

2.5.2 L'estimation

L'estimation est similaire à la classification à part que la variable de sortie est numérique plutôt que catégorique. Elle consiste à estimer la valeur d'un champ à partir des caractéristiques d'un objet. Il suffit d'attribuer une classe particulière pour un intervalle de valeurs du champ estimé [16].⁴

2.5.3 La prédiction

Cela consiste à estimer une valeur future. En général, les valeurs connues sont historisées. On cherche à prédire la valeur future d'un champ. Cette tâche est proche des précédentes. Les méthodes de classification et d'estimation peuvent être utilisées en prédiction. Des exemples de tâches de prédiction sont :

- prédire les valeurs futures d'actions,
- prédire, au vu de leurs actions passées, les départs de clients.

2.5.4 Les règles d'association

Cette tâche, plus connue comme l'analyse du panier de la ménagère, consiste à déterminer les variables qui sont associées. Elle consiste à rechercher les relations ou les dépendances existantes entre plusieurs caractéristiques d'un individu. L'exemple type est la détermination

des articles (le pain et le lait, la tomate, les carottes et les oignons) qui se retrouvent ensemble sur un même ticket de supermarché. Cette tâche peut être effectuée pour identifier des opportunités de vente croisée et concevoir des groupements attractifs de produit. [6]

2.5.5 La segmentation

Le clustering (ou la segmentation) est le regroupement d'enregistrements ou des observations en classes d'objets similaires. Un cluster est une collection d'enregistrements similaires l'un à l'autre, et différents de ceux existants dans les autres clusters. La différence entre le clustering et la classification est que dans le clustering il n'y a pas de variables sortantes. La tâche de clustering ne classe pas, n'estime pas, ne prévoit pas la valeur d'une variable sortantes. Au lieu de cela, les algorithmes de clustering visent à segmenter la totalité de données en des sous groupes relativement homogènes. Ils maximisent l'homogénéité à l'intérieur de chaque groupe et la minimisent entre les différents groupes.

2.6 Apprentissage supervisé Vs Apprentissage non supervisé

Les techniques supervisées : appartiennent à la modélisation prédictive. Dans l'apprentissage supervisé (classification), les données d'apprentissage (observations, mesures) sont accompagnées d'étiquettes indiquant la classe de l'observation. Les nouvelles données sont classées en fonction de l'ensemble de formation.

Les techniques non supervisées :appartiennent à la modélisation descriptive. Dans l'apprentissage non supervisé (clustering), la classe de données d'apprentissage est inconnue. Nous cherchons à trouver des classes ou des groupes d'observations ou de mesures.

2.7 Modélisation descriptive Vs prédictive

La modélisation prédictive : est utilisée lorsque l'objectif est d'estimer la valeur d'un attribut cible particulier et qu'il existe des exemples de données de formation pour lesquels les valeurs de cet attribut sont connues. Un exemple est la classification, qui prend un ensemble de données déjà divisées en groupes prédéfinis et recherche des modèles dans les données qui différencient ces groupes. Ces modèles découverts peuvent ensuite être utilisés pour classifier d'autres données pour lesquelles la désignation de groupe appropriée pour l'attribut cible est inconnue (bien que d'autres attributs puissent être connus) [14].

La modélisation descriptive : ou clustering, divise également les données en groupes. Avec le regroupement, cependant, les groupes appropriés ne sont pas connus à l'avance, les modèles découverts en analysant les données sont utilisés pour déterminer les groupes. Par exemple, un annonceur pourrait analyser une population générale afin de classer les clients

potentiels dans différents clusters, puis développer des campagnes publicitaires distinctes ciblées sur chaque groupe. La détection des fraudes utilise également le clustering pour identifier des groupes d'individus présentant des schémas d'achat similaires [14].

2.8 Les méthodes de data mining

Dans le domaines de data mining il existe des méthodes spécifiques pour un problème particulier. On peut les classer en fonction de :

- La tâche à réaliser.
- Le type et la nature des données.
- L'ensemble des connaissances et des compétences disponibles,
- La finalité du modèle construit,
- L'environnement social, technique, philosophique de l'entreprise,
- ... etc.

On distingue deux grandes catégories de méthodes d'analyse consacrées à la fouille de données[23] :

2.8.1 Les méthodes classiques

On y retrouve des outils généralistes de l'informatique ou des mathématiques :

- Les requêtes dans les bases de données, simples ou multi-critères, dont la représentation est une vue,
- Les requêtes d'analyse croisée, représentées par des tableaux croisés.
- Les différents graphes, graphiques et représentations.
- Les statistiques descriptives.
- L'analyse de données : analyse en composantes principales(ACP).

2.8.2 Les méthodes sophistiquées

Elles ont été élaborées pour résoudre des tâches bien définies. Ce sont :

- Les algorithmes de segmentation,
- les règles d'association,
- Les algorithmes de recherche du plus proche voisin,
- Les arbres de décision,
- Les réseaux de neurones,
- Les algorithmes génétiques,
- etc.

La section suivante n'est pas une présentation exhaustive de l'ensemble des techniques de la fouille de données, mais une présentation de quelques méthodes sophistiquées pour fournir un aperçu du domaine.

2.8.2.1 Arbre de décision

Développées autour de 1960 (AID de Morgan et Sonquist) et très utilisées en marketing, ces méthodes délaissées par les statisticiens ont connu un regain d'intérêt avec les travaux de Breiman et al. (1984) qui en ont renouvelé la problématique : elles sont devenues un des outils les plus populaires du data mining en raison de la lisibilité des résultats [52].

Fournit des règles explicites de classement et supporte bien les données hétérogènes, manquantes et les effets non linéaires, actuellement la seule grande concurrente de l'arbre de décision est la régression logistique, cette méthode étant préférée dans la prédiction du risque en raison de sa plus grande robustesse. Remarquons que les arbres de décision sont à la frontière entre les méthodes prédictives et descriptives, puisque leur classement s'opère en segmentant la population à laquelle ils s'appliquent : ils ressortissent donc à la catégorie des classifications hiérarchiques descendantes supervisées.

- Description de l'algorithme

La technique de l'arbre de décision est employée en classement pour détecter des critères permettant de répartir les individus d'une population en n classes (souvent $n=2$) prédéfinies. On commence par choisir la variable qui, par ses modalités, sépare le mieux les individus de chaque classe, de façon à avoir des sous-populations, que l'on appelle nœuds, contenant chacune le plus possible d'individus d'une seule classe, puis on réitère la même opération sur chaque nouveau nœud obtenu jusqu'à ce que la séparation des individus ne soit plus possible ou plus souhaitable. Par construction, les nœuds terminaux (les feuilles) sont tous majoritairement constitués d'individus d'une seule classe avec une assez forte probabilité, quand il satisfait l'ensemble des règles permettant d'arriver à cette feuille. L'ensemble des règles de toutes les feuilles constitue le modèle de classement (Figure 5).

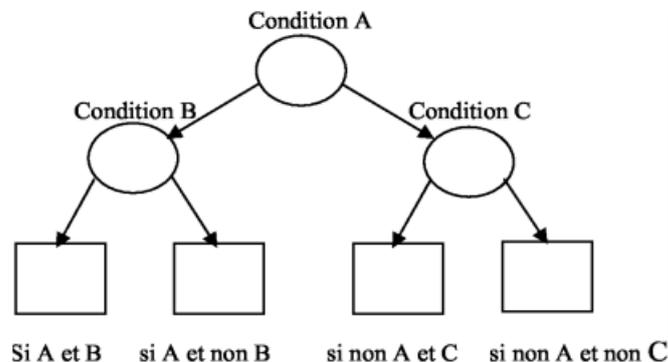


FIGURE 5 – Arbre de décision[21]

- Critiques de la méthode

1. Avantage

- Donne une classification efficace : l'attribution d'une classe à un exemple à l'aide d'un arbre de décision est un processus très efficace (parcours d'un chemin dans un arbre).
- Disponibilité des outils : les algorithmes de génération d'arbres de décision sont disponibles dans tous les environnements de fouille de données.
- Adaptabilité aux attributs de valeurs manquantes : les algorithmes peuvent traiter les valeurs manquantes (descriptions contenant des champs non renseignés) pour l'apprentissage, mais aussi pour la classification.
- Traitement de tout type de données : l'algorithme peut prendre en compte tous les types d'attributs et les valeurs manquantes. Il est robuste au bruit.

2. Inconvénients

- Méthode sensible au nombre de classes : les performances tendent à se dégrader lorsque le nombre de classes devient trop important.
- Manque d'évolutivité dans le temps : l'algorithme n'est pas incrémental, c'est-à-dire, que si les données évoluent avec le temps, il est nécessaire de relancer une phase d'apprentissage sur l'échantillon complet (anciens exemples et nouveaux exemples).

2.8.2.2 Les réseaux de neurones

Les réseaux de neurones sont apparus dans les années cinquante avec les premiers perceptrons, et sont utilisés industriellement depuis les années quatre-vingt. Un réseau de neurone "ou réseau neuronal" à une architecture calquée sur celle du cerveau, organisée en neurones et synapses, et se présente comme un ensemble de nœuds "ou neurones formels, ou unités" connectés entre eux, chaque variable prédictive continue correspondant à un nœud d'un premier niveau, appelé couche d'entrée, et chaque variable prédictive catégorique (ou chaque modalité d'une variable catégorique) correspondant également à un nœud de la couche d'entrée.

Le cas échéant, lorsque le réseau est utilisé dans une technique prédictive, il y a une ou plusieurs variables à expliquer, elle correspondant alors chacune à un nœud (ou plusieurs dans le cas des variables catégorielles) d'un dernier niveau : la couche sortie. Les réseaux prédictifs sont dits "à apprentissage supervisé" et les réseaux descriptifs sont dits "à apprentissage non supervisé". Entre la couche d'entrée et la couche sortie sont parfois connectés à des nœuds appartenant à un niveau intermédiaire : la couche cachée. Il peut exister plusieurs couches cachées [58].

— Définition

"Les réseaux de neurones sont des outils très utilisés pour la classification, l'es-

timation, la prédiction et la segmentation. Ils sont issus des modèles biologiques, sont constitués d'unités élémentaires (les neurones) organisées selon une architecture" [58].

Un noeud est initialisées par des valeurs entre 0 et 1 (ou parfois entre -1 et 1) applées des poids et obtenu en sorties entre 0 et n valeurs ,selon la fonction de transfert .

Une fonction de combinaison calcule une première valeur à partir des noeuds connectés en entrée et poids des connexions. Dans les réseaux les plus courants, les perceptrons, il s'agit de la somme pondérée "nipi" des valeurs des noeuds en entrée. Afin de déterminer une valeur en sortie, une seconde fonction, appelée fonction de transfert (ou d'activation), est appliquée à cette valeur. Les noeuds de la couche d'entrée sont triviaux, dans la mesure où ils ne combinent rien, et ne font que transmettre la valeur de la variable qui leur correspond.

parmi les applications de les réseaux de neurones on cite :

- Statistiques : analyse de données, prévision, classification
- Robotique : contrôle et guidage de robots ou de véhicules autonomes.
- Imagerie, reconnaissance de formes.
- Traitement du signal.
- Simulation de l'apprentissage [15].

Un noeud de perceptron se présente donc comme suit :

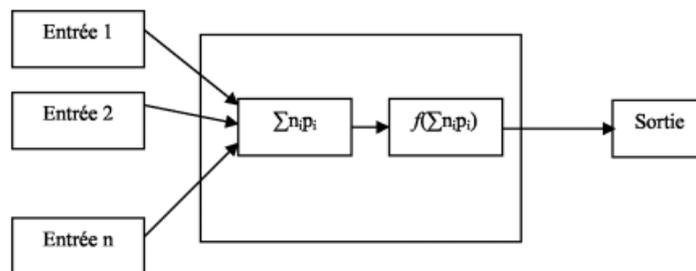


FIGURE 6 – Noeud d'un réseau de neurone[58].

Dans cette figure, on utilise les notions suivantes :

- n_i est la valeur du noeud i du niveau précédent (la sommation sur i correspond à l'ensemble des noeuds du niveau précédent connectés au noeud observé).
- p_i est le poids associé à la connexion entre i et le noeud observé.
- f est la fonction de transfert associée au noeud observé.

- Critiques de la méthode

1. Avantages

- Lisibilité du résultat : le résultat de l'apprentissage est un réseau constitué de cellules organisées selon une architecture, définies par une fonction

d'activation et un très grand nombre de poids à valeurs réelles.

- Les données réelles : les réseaux traitent facilement les données réelles "préalablement normalisées" et les algorithmes sont robustes au bruit. Ce sont, par conséquent, des outils bien adaptés pour le traitement de données complexes éventuellement bruitées comme la reconnaissance de formes (son, images sur une rétine, etc.).
- Classification efficace : le réseau étant construit, le calcul d'une sortie à partir d'un vecteur d'entrée est un calcul très rapide.
- En combinaison avec d'autres méthodes : pour des problèmes contenant un grand nombre d'attributs pour les entrées, il peut être très difficile de construire un réseau de neurones. On peut, dans ce cas, utiliser les arbres de décision pour sélectionner les variables pertinentes, puis générer un réseau de neurones en se restreignant à ces entrées.

2. Inconvénients

- Un réseau de neurones artificiel représente une boîte noire, et il est très difficile voire impossible d'analyser et comprendre son fonctionnement en face d'un problème donné, ce qui empêche de choisir la structure (type, nombre de nœuds, organisation, connexions,...etc) la mieux adaptée à ce problème.
- L'ordre de présentation des exemples d'entraînement au réseau influe directement sur les résultats obtenus. Pour surmonter ce problème, il est nécessaire de répéter au moins la phase d'entraînement avec des ordres différents des exemples ce qui augmente considérablement le temps d'apprentissage.
- Dans le cas des bases de données, les réseaux de neurones artificiels ne permettent pas de traiter des exemples avec des attributs symboliques (catégoriels) qu'après un encodage adapté, à l'inverse de plusieurs autres techniques d'apprentissages tel que les SVMs et les arbres de décision.
- Les RNAs représentent un inconvénient majeur qui est leur sensibilité aux minima locaux et la possibilité de leur divergence. L'ambiguïté de leur fonctionnement empêche d'éviter de tels cas[16].

2.8.2.3 Deep learning

L'apprentissage en profondeur est une classe d'algorithmes d'apprentissage automatique qui :

- Utilisez une cascade de plusieurs couches d'unités de traitement non linéaires pour l'extraction et la transformation de caractéristiques. Chaque couche successive utilise la sortie de la couche précédente en entrée.

- Apprendre de manière supervisée (classification, par exemple) et, ou non supervisée (analyse de modèle, par exemple).
- Apprendre plusieurs niveaux de représentations correspondant à différents niveaux d'abstraction, les niveaux forment une hiérarchie de concepts.

Le deep Learning est utilisé dans de nombreux domaines :

- Reconnaissance d'image,
- Traduction automatique,
- Voiture autonome,
- Diagnostic médical,
- Recommandations personnalisées,
- Modération automatique des réseaux sociaux,
- Prédiction financière et trading automatisé,
- Identification de pièces défectueuses.
- etc.

Dans la figure suivant un exemple de deep learning qui permet d'identifier un ensemble d'entrée si le résultat c'est une car ou nn : .

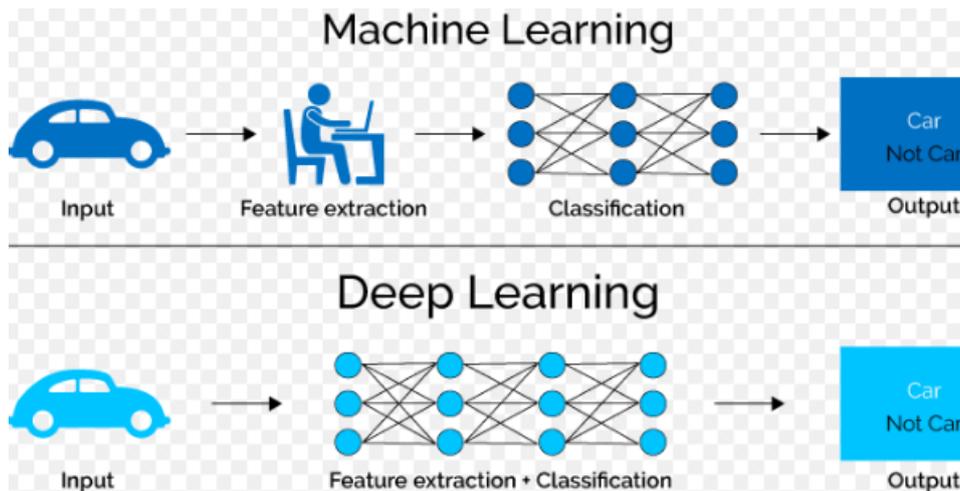


FIGURE 7 – Exemple d'utilisation du Deep learning [5].

2.8.2.4 Règles d'association

est une méthode d'apprentissage automatique basée sur des règles permettant de découvrir des relations intéressantes entre des variables dans de grandes bases de données. Il est destiné à identifier les règles fortes découvertes dans les bases de données en utilisant certaines mesures d'intérêt tel que la confiance, lift, levrage, conviction qui permet de présente la coexistence des deux paries (gauche et droite) de la règle et d'après ces mesure on obtient un ensembles des règles solide qui dépassent un minimum de support et un minimum de qualité . Sur la base du concept de règles strictes, l'apprentissage par l'asso-

ciation a introduit des règles d'association permettant de détecter les régularités entre les observations dans des données de rangées à grande échelle . Algorithmes utilisés, Apriori, AprioriTID, etc. la recherche des règles d'association parmi les méthodes non supervisée car on dispose en entrée que la description des achats [11].

les règles d'association sont principalement facile à l'utilisation et défini comment les produits ou des services se situent les uns par rapport aux autres. Concernant l'utilité de ses règles dans des domaines différents spécifiant dans le système de l'entreprise et aussi en marketing. quand le produire des règles intéressantes, peut aussi produire des règles triviales (déjà bien connues des intervenants du domaine) ou inutiles (provenant de particularités de l'ensemble d'apprentissage). Des exemples de règles d'association sont :

- Si un client achète des plantes alors il achète du terreau,
- Si un client achète une télévision, il achètera un magnétoscope dans un an.

2.9 Les machines à support vecteurs (SVM)

Les machines à vecteur support(en anglais support vector machine SVM) se situent sur l'axe de développement de la recherche humaine des techniques d'apprentissage. Les SVMs sont une classe de techniques d'apprentissage introduite par Vladimir Vapnik au début des années 90, elles reposent sur une théorie mathématique solide. Les SVMs sont dans leur origine utilisées pour la classification binaire et la régression. Aujourd'hui, elles sont utilisées dans différents domaines de recherche et d'ingénierie tel que le diagnostic médical, le marketing, la biologie, la reconnaissance de caractères manuscrits et de visages humains. Notre travail est intéressé par l'utilisation des SVMs pour la classification [17].

2.9.1 Principe de la technique SVM

L'idée principale des SVM consiste a projeter les données dans un espace de plus grande dimension appelé "espace de caractéristiques", afin que les données non linéairement séparables dans l'espace d'entrée deviennent linéairement séparables dans l'espace de caractéristiques. En appliquant dans cet espace la technique de construction d'un hyperplan optimal séparant les deux classes, on obtient une fonction de classification qui dépend d'un produit scalaire des images des données de l'espace d'entrée dans l'espace des caractéristiques.

Ce produit scalaire peut être exprimé, sous certaines conditions, par fonctions définies dans l'espace d'entrée, qu'on appelle "les noyaux". Ce multiple choix de noyaux rend les SVMs plus intéressantes et surtout plus riches puisqu'on peut toujours chercher de nouveaux noyaux qui peuvent être mieux adaptés à la tâche qu'on veut accomplir. Les trois noyaux les plus utilisés sont : le noyau linéaire, le noyau polynomial et le noyau gaussien.

pour deux classes d'exemples donnés, le but de SVM est de trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes. Avec SVM, ce classificateur est un classificateur linéaire appelé hyperplan. Dans le schéma qui suit, on détermine un hyperplan qui sépare les deux ensembles de points. Les points le plus

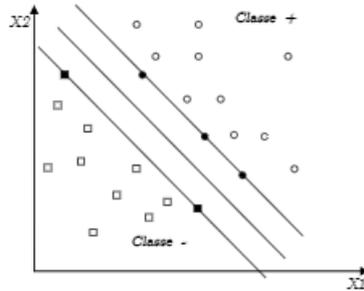


FIGURE 8 – L'hyperplan H qui sépare les deux ensembles de points.

proches, qui seuls sont utilisés pour la détermination de l'hyperplan, sont appelés vecteurs de support.

2.9.1.1 Hyperplan

On appelle hyperplan séparateur un hyperplan qui sépare les deux classes, en particulier il sépare leurs points d'apprentissage. Il est évident qu'il existe une multitude d'hyperplan valide mais la propriété remarquable des SVM est que cet hyperplan doit être optimal. Nous allons donc en plus chercher parmi les hyperplans valides, celui qui passe au milieu des points des deux classes d'exemples. Intuitivement, cela revient à chercher l'hyperplan optimal [46].

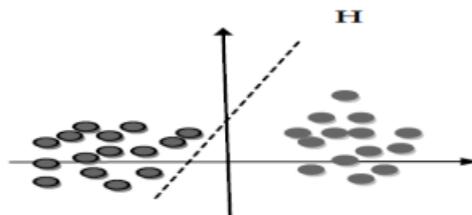


FIGURE 9 – L'hyperplan H qui sépare les deux ensembles de points.

2.9.1.2 Vecteurs supports

Pour une tâche de détermination de l'hyperplan séparable des SVM est d'utiliser seulement les points les plus proches (les points de la frontière entre les deux classes des données) parmi l'ensemble total d'apprentissage, ces points sont appelés "vecteurs supports".

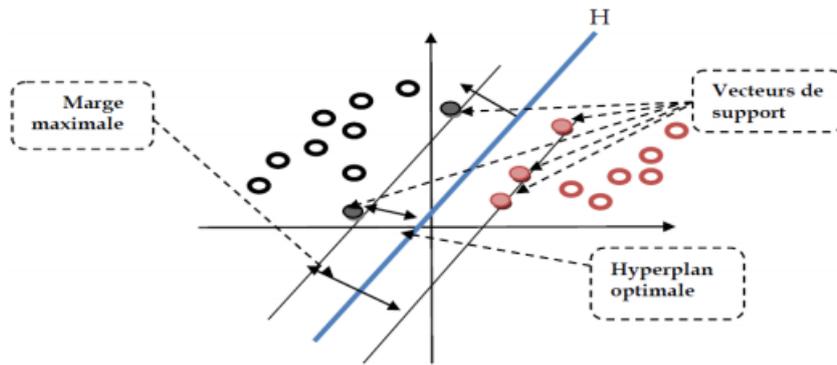


FIGURE 10 – L’hyperplan H optimal, vecteurs supports et marge maximale.

2.9.1.3 Marge

Il existe une infinité d’hyperplans capable de séparer parfaitement les deux classes d’exemples. Le principe des SVM est de choisir celui qui va maximiser la distance minimale entre l’hyperplan et les exemples d’apprentissage (la distance entre l’hyperplan et les vecteurs supports), cette distance est appelée "la marge" [46].

2.9.1.4 Maximiser la marge

Intuitivement, le fait d’avoir une marge plus large procure plus de sécurité lorsqu’on classe un nouvel exemple. De plus, si l’on trouve le classificateur qui se comporte le mieux vis-à-vis des données d’apprentissage, il est clair qu’il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie droite nous montre qu’avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu’il tombe dans la marge. On constate sur la partie gauche qu’avec une plus petite marge, l’exemple se voit mal classé. .

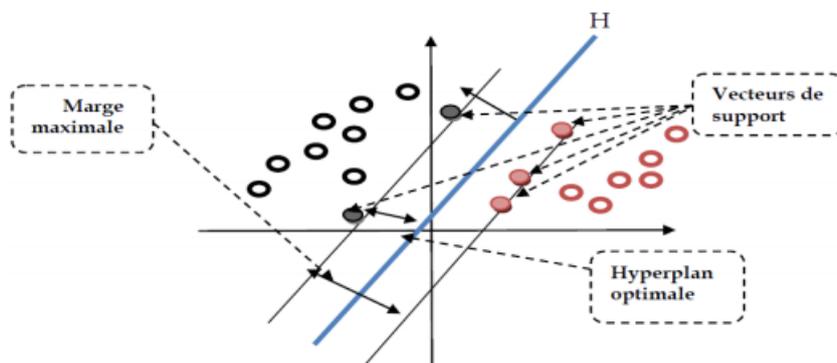


FIGURE 11 – Meilleur hyperplan séparateur.

2.9.2 Les étapes de construction d'un modèle de décision

La construction d'un modèle de décision est basé sur deux étapes la première pour l'apprentissage et la deuxième pour l'utilisation. Le schéma suivant montre le fonctionnement global de la méthode SVM :

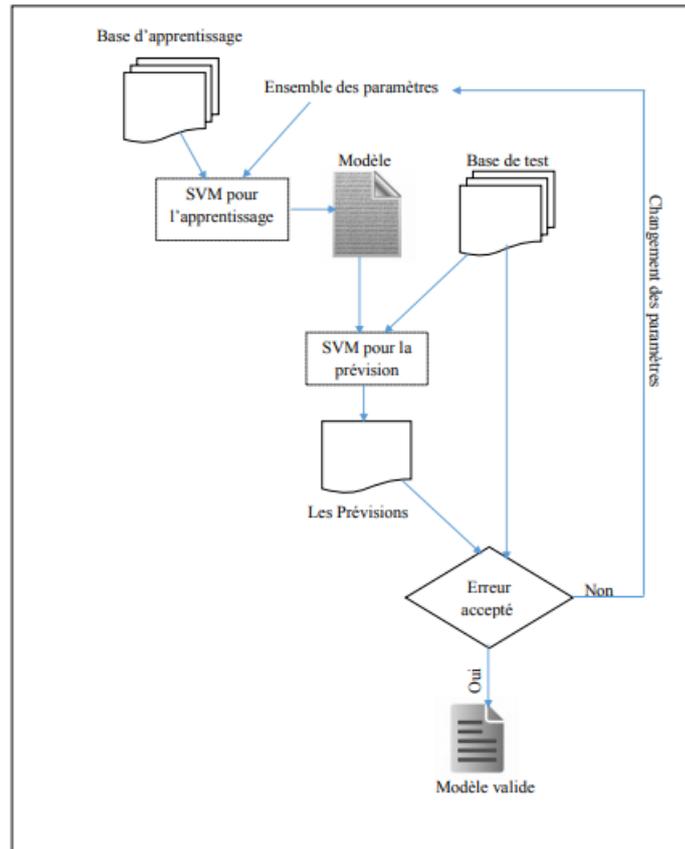


FIGURE 12 – Schéma globale de fonctionnement de l'SVM

2.9.3 Avantage et inconvénients des SVM

1. Avantages

- SVM est une méthode de classification intéressante car le champ de ses applications est large, parmi ses avantages nous avons :
- Un grand taux de classification et de généralisation par rapport aux méthodes classiques.
- L'algorithme est robuste face aux changements d'échelle.
- Résultat en général équivalents et souvent meilleurs.
- Décision rapide. la classification d'un nouvel exemple consiste à voir le signe de la fonction de décision.
- Les exemples sont comparés juste avec le support vecteur et non pas avec tous les exemples d'apprentissage.

2. Inconvénients

- Grand quantité d'exemples en entrées implique un calcul matriciel important.
- Temps de calcul élevé lors d'une régularisation des paramètres de la fonction noyau.
- Classification binaire d'où la nécessité d'utiliser l'approche un-contre-un.

2.10 Conclusion

Le data mining est un domaine très variées permettant d'analyse des grandes quantité de données collectées, utilisé dans divers domaines d'application de l'informatique, de l'intelligence artificielle, des statistiques et des bases de données dans le but d'extraire des connaissances à partir des données.

Dans ce chapitre, nous avons présenté les principaux concepts de fouille de données, son processus, son tâches et son méthodes les plus utilisés en data mining ainsi que les avantages et les inconvénients de chaque méthode.

Dans notre travail, nous nous intéressons plus aux méthodes supervisées puisque l'objectif est la prise de décision à partir des expériences de médecins.

Dans le chapitre suivant, nous allons présenter notre système proposé pour le suivis de la santé aux individus en temps réel en utilisant les montres intelligentes et les méthodes d'apprentissage automatique.

CONCEPTION

3.1 Introduction

Les services de surveillance médicale à distance connaissent une demande croissante provenant généralement des patients souffrant de maladies chroniques [35] et particulièrement des personnes âgées vu leurs contraintes de déplacements fréquents [29].

Les personnes à surveiller à distance peuvent être des personnes âgées, à mobilité réduite, les personnes malades qui ont des maladies chroniques ou toute personne désirant surveiller sa santé. Ces services reposent sur les nouvelles technologies pour surveiller les activités des patients sans imposer à ces derniers de consulter des centres médicaux ou d'analyse.

L'application de surveillance à réaliser devrait tenir compte d'un certain nombre d'objectifs :

- Mettre en place un système de surveillance médicale à distance.
- Surveiller à tout moment l'état de santé de la personne (rythme cardiaque, pression artérielle, température, signal cardiaque, ...).
- Avoir des interfaces simples et claires afin d'améliorer les délais d'intervention.
- Prélever des mesures réparties dans le système à surveiller.
- Respecter la contrainte temps réel lors de l'utilisation de la valeur mesurée ;
- Collecter les données correspondantes dans une base de données dans laquelle peuvent être associés des algorithmes de traitement et de diagnostic.
- Générer éventuellement des alarmes ou initier des interventions.

Ce chapitre est composé de trois parties. La première partie présente la spécification des besoins et les exigences auxquelles notre application devra répondre et les objectifs à atteindre.

La deuxième partie décrit l'architecture technique du système SMD, à travers la définition d'un ensemble de matériels et logiciels s'appuyant sur la technologie des montres intelligentes. L'architecture est basée sur des applications concernant des différentes maladies et mettant en relation le patient et le médecin pour garantir un suivi temps réel de la santé. La troisième partie concerne l'administration supervisant l'inscription et la mise en œuvre des systèmes intelligents de suivi. Nous présentons cette architecture sous les deux aspects : structurel et dynamique.

Comme exemple sur les systèmes intelligents, nous proposons une approche de détection de maladies cardiaques par Data Mining. Rappelons que le principal but de notre système SMD est de fournir des services de surveillance de santé, à toute personne et à tout moment, surmontant les contraintes de l'endroit, du temps et de l'espace.

Par ailleurs, les performances de notre système de surveillance de signes physiologiques de patient doivent se mesurer par :

- La capacité à reconnaître la présence d'une anomalie.
- La capacité à reconnaître l'absence d'une anomalie.
- La rapidité de détection.

3.2 Spécification des besoins

Pour présenter notre système de surveillance médicale à distance, il faut suivre un ensemble d'étapes concernant la construction. Pour cela, la phase de spécification des besoins permet de comprendre d'une façon claire le système. Cette phase est indispensable pour la collection des informations nécessaires. Le moyen le plus précis et réussi pour la spécification des besoins d'un système informatique est le concept de cas d'utilisation. Ce concept sera par la suite raffiné de façon à aboutir à la conception de l'architecture. Un cas d'utilisation est utilisé pour donner une vision globale du comportement fonctionnel d'un système. Ils sont utiles pour des présentations auprès de la direction ou des acteurs d'un projet, mais pour le développement, les cas d'utilisation sont plus appropriés. Un cas d'utilisation représente une unité discrète d'interaction entre un utilisateur (humain ou machine) et un système. Il est une unité significative de travail [30].

— Contexte de la solution

On peut définir le contexte de notre système à l'aide de la modélisation Uml (Unified Modeling Language) qui est le diagramme de contexte. Ce diagramme exprime l'environnement du système dans une situation donnée, permet d'identifier les interactions avec les acteurs qui sont en dehors du système. La figure suivante présente explicitement la limite entre le système et le milieu extérieur [62].

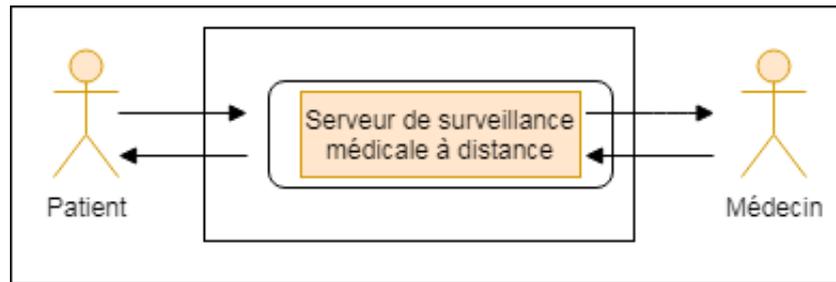


FIGURE 1 – Diagramme de contexte d'un Système médical de surveillance.

3.3 Le diagrammes des cas d'utilisation

Un cas d'utilisation est donc l'abstraction d'une fonction typique du système à modéliser (figure 2)

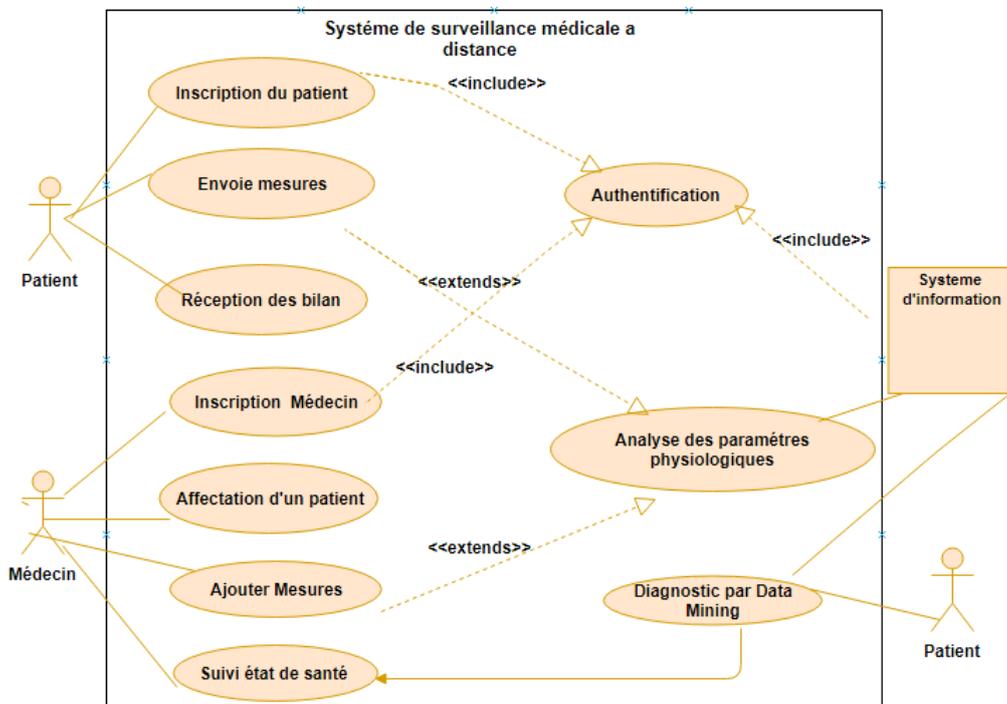


FIGURE 2 – "Diagramme de cas d'utilisation générale"

Comme il vient d'être explicité, notre système est composé de trois applications principales :

3.3.1 Application patient

L'application patient est destinée à être exécuté au niveau de la montre du patient. Elle doit lui permettre de s'inscrire, de se connecter, d'envoyer les mesures aux serveur d'administration, les consulter et d'en recevoir les notifications.

— Cas d'utilisation 1 : Authentification

Dans le scénario d'authentification, les utilisateurs des montres intelligentes sont invités à saisir leurs identifiants suivis de leurs mots de passe pour pouvoir accéder à l'application d'une façon sécurisée. Cette étape ne sera réussie que si et seulement si le couple, mot de passe et identifiant, seront valide. La réalisation de ce scénario est dans le but de renforcer la sécurisation de l'application et d'améliorer sa performance.

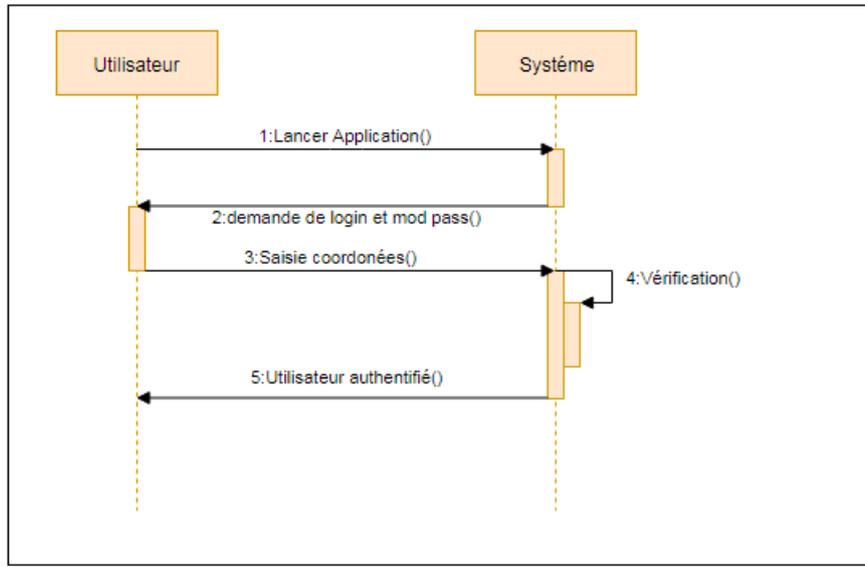


FIGURE 3 – Diagramme de séquence "Authentification"

— Cas d'utilisation 2 : Inscription patient

Le patient dans son premier contact avec le système doit saisir ses informations personnelles sur son profil. Il s'agit de son nom, son prénom, son sexe, son âge, son groupe sanguin, son adresse, son numéro de téléphone et son email. Ces propriétés sont envoyées au système d'information et stockées dans la base de données distante. Après cette opération qui est effectuée une seule fois, le patient pourra se connecter au système et d'en utiliser.



FIGURE 4 – Diagramme de séquence "Profil du patient"

— Cas utilisation 3 : Envoie des mesures au système

Après l'inscription du patient dans le système et l'affectation à son médecin et l'indication des mesures surveillées et leur périodicité par le médecin, on peut avoir deux méthodes pour l'envoi des mesures par l'application sur la montre vers le serveur :

- Méthode inform : La montre arme un temporisateur pour une mesure après une demande du serveur et à chaque déclenchement du temporisateur elle prend la valeur de la mesure et l'envoie avec son code et le temps de prélèvement au serveur.
- Méthode request : A chaque fois que la montre reçoit une demande de mesure du serveur, elle effectue un prélèvement et l'envoie au serveur.

La méthode request nécessite un nombre de message plus élevées (presque le double) que la méthode informe et par conséquent elle consomme plus de bande passante sur le réseaux. En contre partie, elle préserve la batterie de la montre puisque le temporisateur est au niveau du serveur.

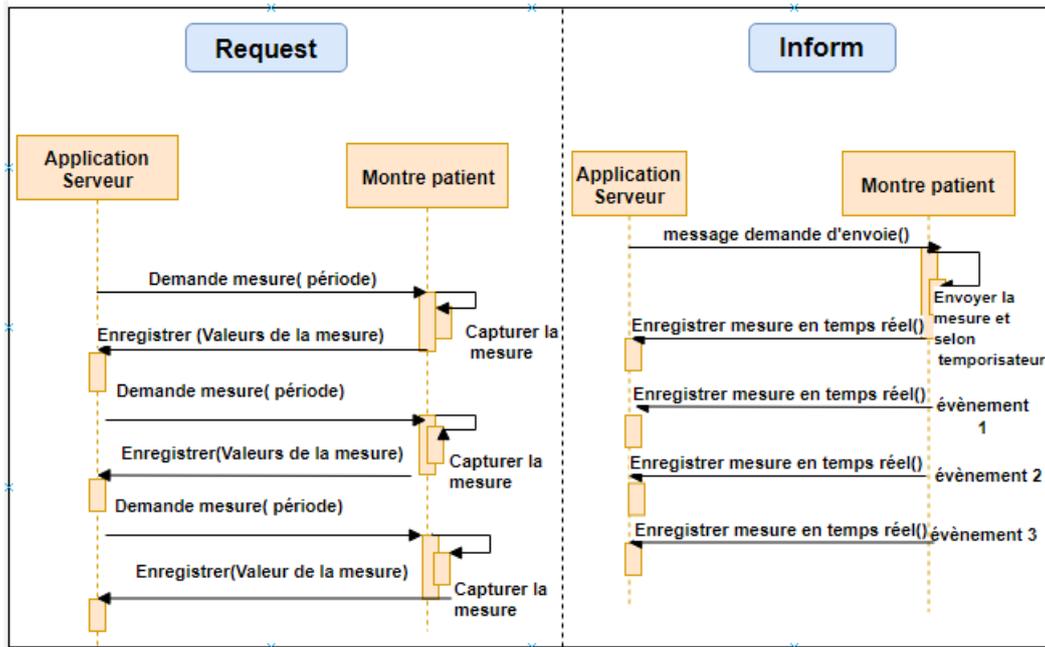


FIGURE 5 – Diagramme de séquence "Envoi des mesures" par les méthodes "inform et request"

- Cas utilisation 4 : Réception de résultat du bilan médical

Cette activité permet au patient de consulter son historique concernant une mesure donnée. Dans ce cas, il doit spécifier le code de la mesure et la période dont il souhaite consulter. Le serveur répond alors en envoyons une liste de valeur de cette mesure avec les dates correspondante.

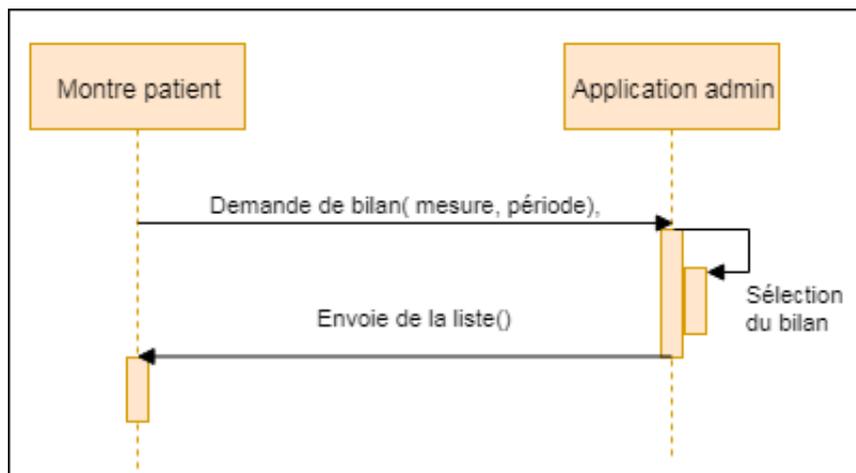


FIGURE 6 – Diagramme de séquence "Réception des bilan"

- Cas utilisation 5 : Réception des notifications

A l'application des modèles de décision entraîné, les patients peuvent recevoir des notifications pour les informer sur le résultat et la décision de diagnostic effectué par le modèle de décision.

3.3.2 Application médecin

L'application médecin est installée au niveau de la montre du médecin. Elle lui permet plusieurs tâches :

- Cas utilisation 1 :Inscription du médecin

Dans cette activité, le médecin doit fournir un ensemble d'informations personnelles et professionnelles et attendre une confirmation de l'administrateur du système pour valider son inscription. Il pourra alors commencer à utiliser le système

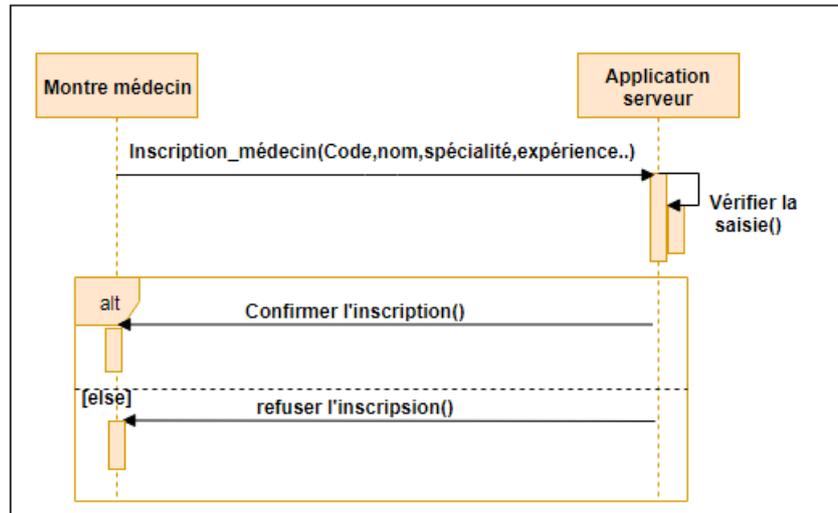


FIGURE 7 – Diagramme de séquence "Inscription de médecin"

- Cas utilisation 2 : Ajout du Patient

Dans cette activité, le médecin choisit parmi les patients inscrits ceux qu'il veut suivre.

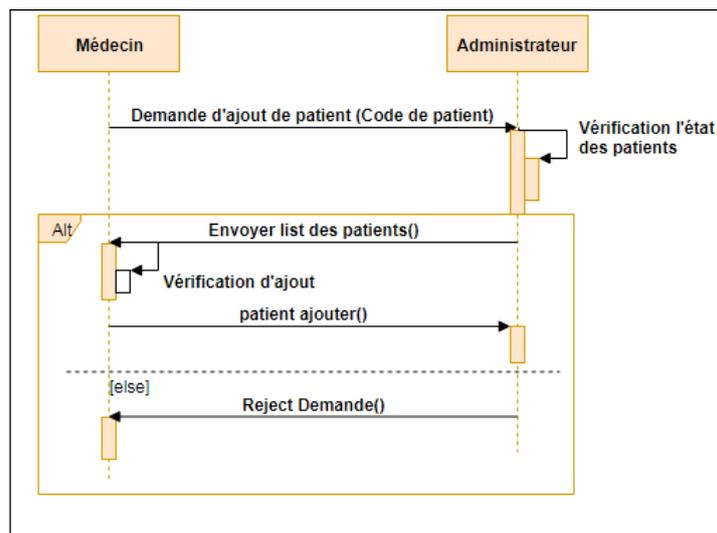


FIGURE 8 – Diagramme de séquence "Ajout de patients"

— Cas utilisation 3 : Ajouter/suppression de mesures

Dans cette activité, le médecin indique à l'administrateur du système pour un patient dont il suit les mesures à contrôler et leur périodicité. Il peut aussi supprimer une mesure déjà enregistrée

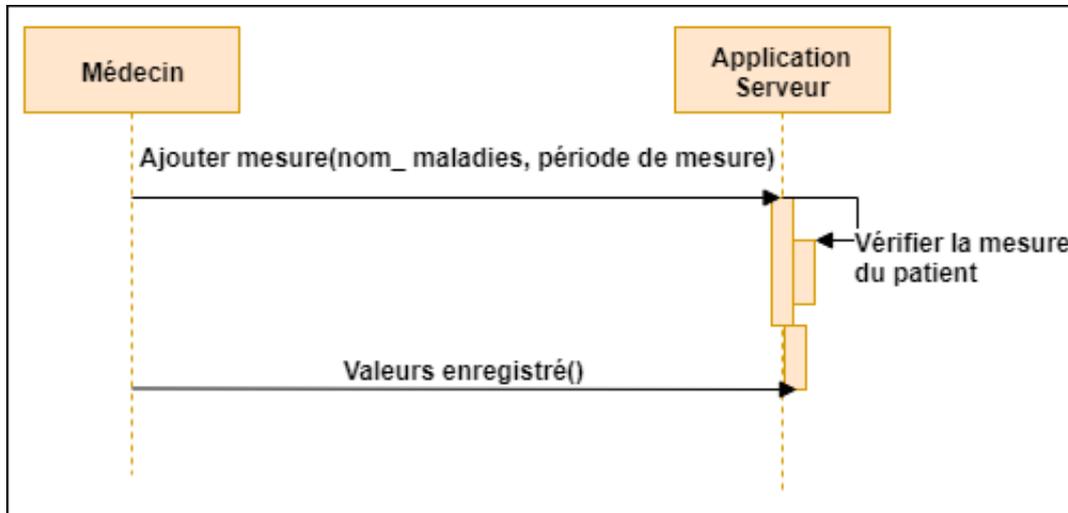


FIGURE 9 – Diagramme de séquence "Ajouter de mesures"

— cas utilisation 4 :Suivi de l'état de santé par le médecin

Le système après l'application d'un modèle de décision sur les données des patients informe les médecins par les résultats observés par le modèle pour attirer leur attention sur les états de leurs patients et vérifier aussi l'efficacité du modèle et améliorer ses performances.

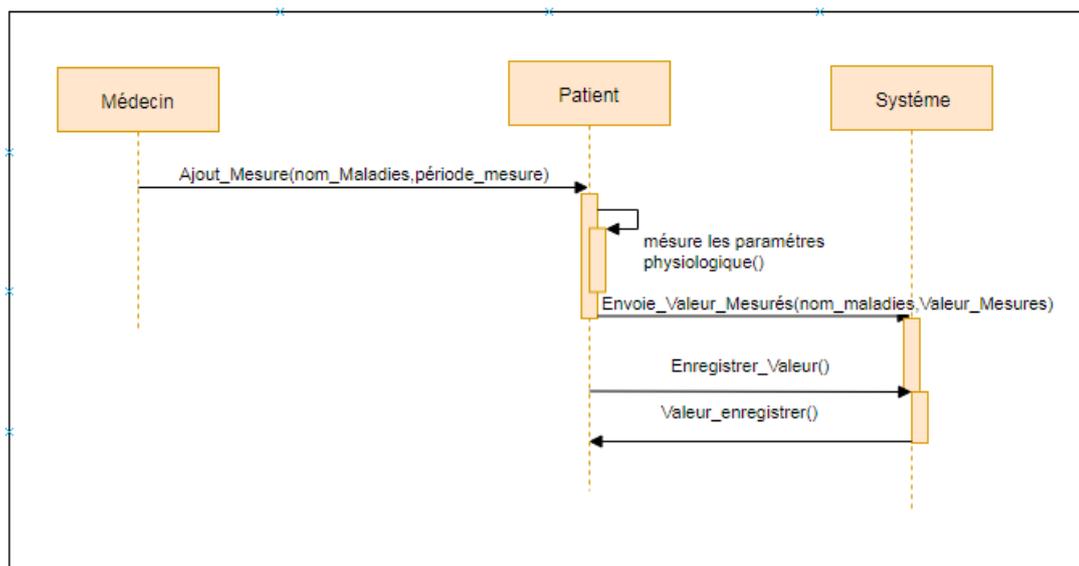


FIGURE 10 – Diagramme de séquence "Suivi de l'état de santé par le médecin"

3.3.3 Administration du système

Le système d'information au niveau de l'administration rassemblent sous forme d'une base de donnée toutes les informations nécessaires à son fonctionnement. Il contient les tables des patients, des médecins, des maladies, des mesures ainsi que celles les reliant. La figure suivante représente le schéma relationnel de la base de données :

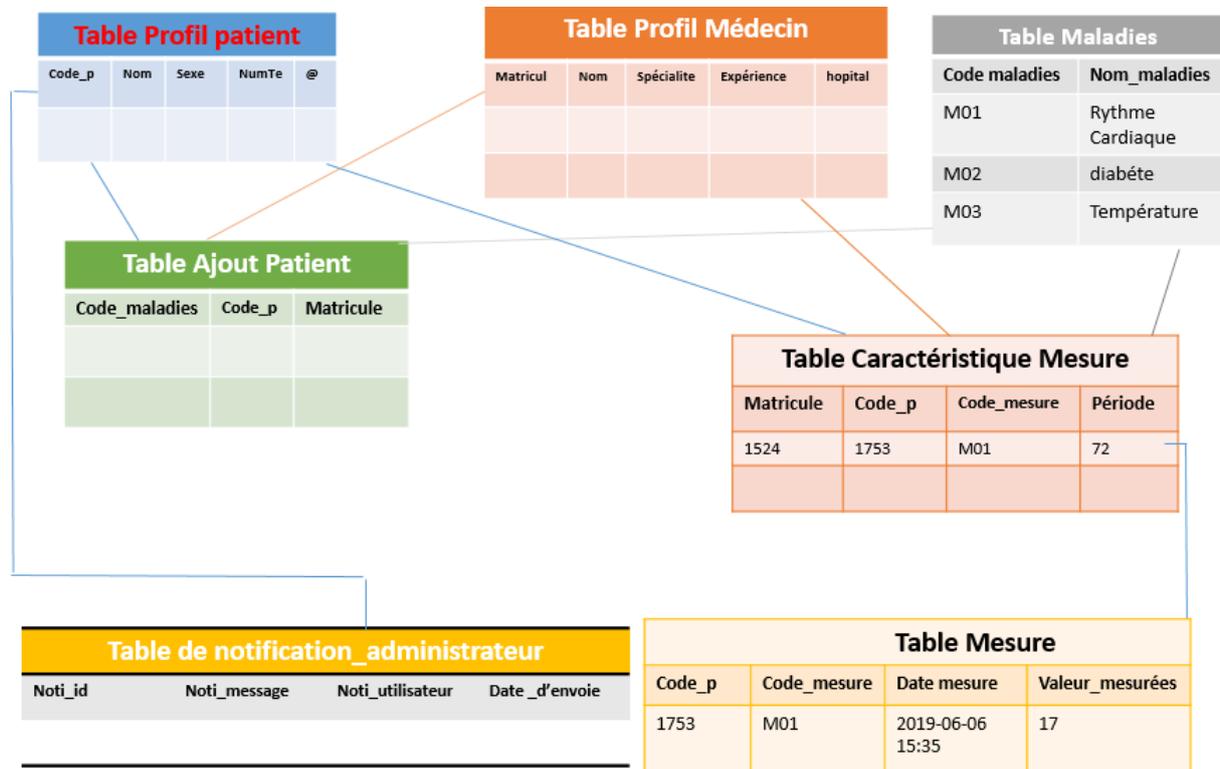


FIGURE 11 – Schéma relationnel de la base de données de notre système.

L'administrateur du système d'information supervise son fonctionnement. Il valide les médecins inscrit après de leur profil, construit et valide et active les modèles de décision et contrôle les paramètres du système telle que les périodes d'application des modèles de décision. Il effectue principalement deux tâches importantes :

- Création de modèle :

La création d'un modèle de décision peut se faire à partir de deux sources :

- Base de données externe : Dans le cas où une base interne n'est pas encore prête lors des premiers temps d'utilisation du système, l'administrateur acquiert une base d'entraînement à partir d'un hôpital, un centre médical, un centre de recherche,...
- Base de données interne : Après le fonctionnement du système pour une période suffisante pour collecter une base d'entraînement. Dans ce cas une phase de pré-traitement sur les données collectées dans les tables de mesures est nécessaire.

Dans tous les cas, l'administrateur utilise une méthode d'apprentissage automatique pour construire un modèle de décision à partir de la base utilisée.

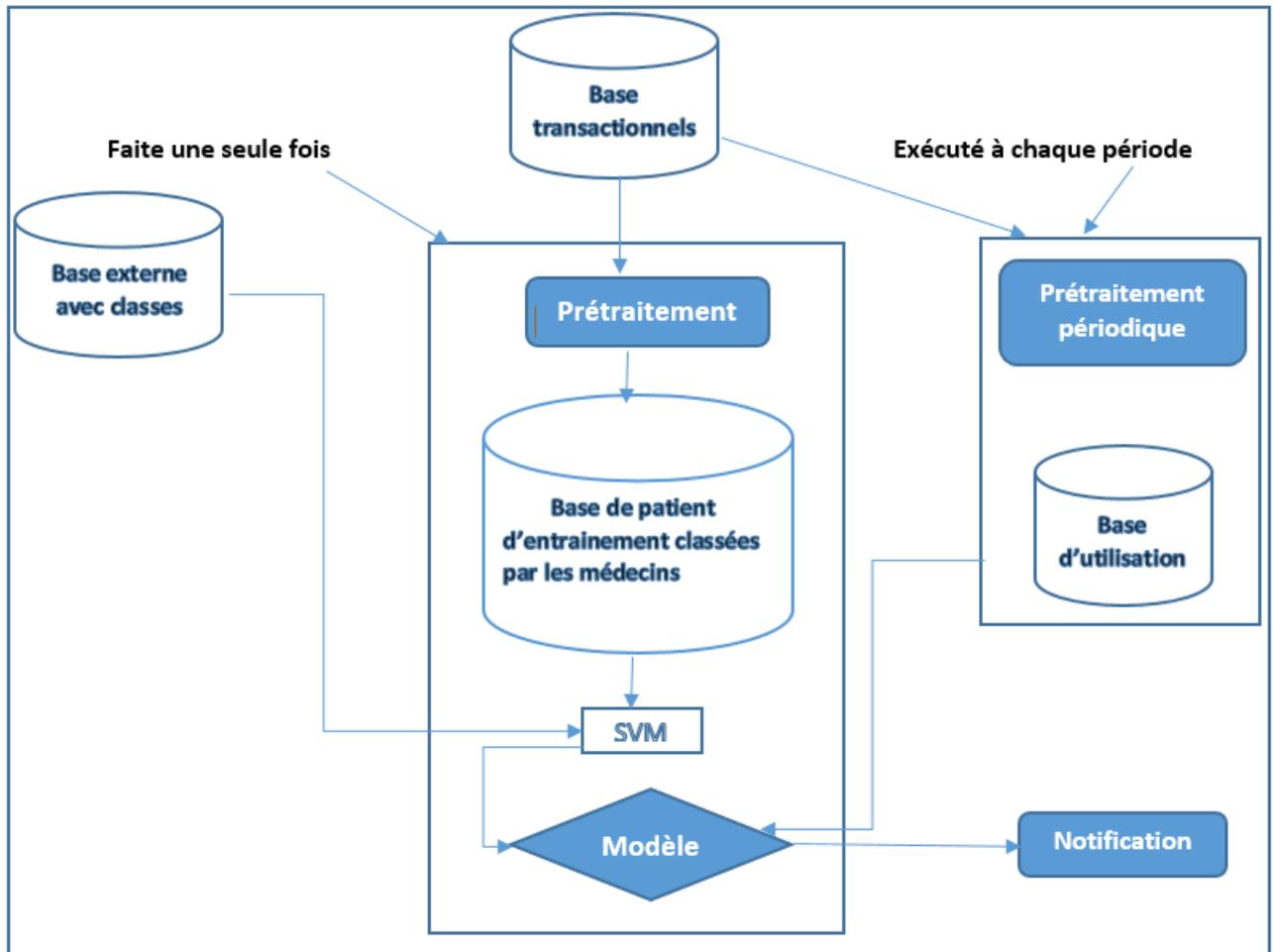


FIGURE 12 – Fonctionnement globale d'un système d'information.

— Application du modèle :

L'administrateur peut lancer un modèle de décision sur les données actuelles des patients pour obtenir des décisions automatiques sur leur situation pour une maladie donnée, comme il peut programmer une application périodique d'un modèle déjà construit.

3.4 Architecture du système

Notre système de surveillance médicale à distance fonctionne en deux interfaces :

- Un site web (application Desktop).
- Une application android (wear) sur une montre intelligente.

Ce système de surveillance de patients à distance (SSPD) rassemble les données physiologiques à travers des sensors existant sur la montre connectée du patient pour mesurer

différentes indices tels que le signal cardiaque, le rythme Cardiaque, la température, la tension artérielle...

Ces mesures seront envoyées au système d'information géré par l'administrateur, et enregistrées dans une base de données pour être utilisées dans des analyses ultérieures.

la figure suivante représente l'architecture globale de notre système d'information.

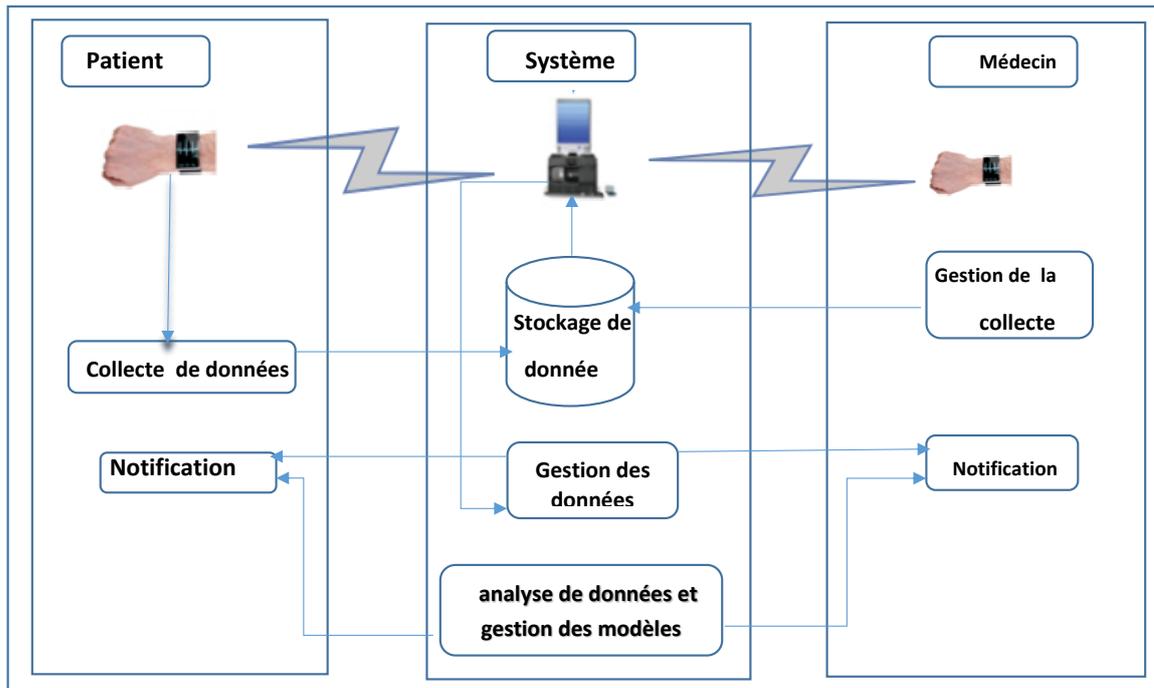


FIGURE 13 – Architecture du Système de Surveillance Médicale à Distance(SSMD)

3.5 Sécurité

Dans le but de protéger les données privées des patients et des médecins, nous avons introduit des mécanismes de sécurité visant à prendre en charge les exigences suivantes :

- Protection des données des patients lors de leur transmission.
- Protection des données des patients lors de leur stockage.
- Protéger l'accès aux système.

Pour cela nous avons opté pour les mécanismes suivants :

- Les données sont transmises sous forme chiffrée.
- Les données d'accès sont stockées sous forme chiffrée, et ne sont pas visibles même par l'administrateur.

- L'authentification est basée session.
- Utilisation des noms d'utilisateurs et mots de passe pour tous les acteurs du système.

3.6 Conclusion

Par rapport aux objectifs assignés à notre travail, nous avons proposé une architecture de système de surveillance médicale à distance (SSMD). Elle est composée essentiellement d'activités de collecte périodiques de données sur les signes physiologiques et le traitement de ces données afin de construire des modèles de décision automatique par les méthodes d'apprentissage automatique et cela dans le but de détecter d'éventuelles déviations par rapport au comportements normaux. Dans le chapitre suivant, nous allons présenter la mise en œuvre de notre conception, les moyens utilisés et les résultats obtenus.

IMPLÉMENTATION ET RÉSULTAT

4.1 Introduction

Après avoir détaillé dans le chapitre précédent l'approche de mise en œuvre pour les parties du système ainsi que la démarche de modélisation conceptuelle de notre système, le présent chapitre a pour objectif d'implémenter les étapes proposées pour réaliser le système conçu.

Étant donné que la conception détaillée tient compte de la plateforme de réalisation, nous commençons dans la première partie, par préciser l'environnement et les technologies utilisés dans le développement. Par la suite, nous allons expliquer toutes les expérimentations que nous avons appliquées sur la méthode proposée et les résultats obtenus.

Enfin, nous nous intéressons à présenter quelques parties de code et des interfaces des principaux modules développés.

4.2 Environnement et outils de programmation

Pour valider notre conception, nous avons utilisé plusieurs outils de codage, de conversion, de communication, et d'administration. Plus précisément, nous avons utilisé :

- PhpMyAdmin pour gérer la base de données au niveau du serveur.
- XAMPP pour gérer le serveur web au niveau de serveur.
- Android studio et android wear pour créer l'application sur montre.
- LibSVM pour gérer les modèles de décision au niveau du serveur.
- Weka pour l'analyse et la validation des modèles de décision.
- PHP et Javascript pour coder les applications.

4.2.1 XAMPP

XAMPP est un ensemble de logiciels permettant de mettre en place facilement un serveur Web local, un serveur FTP et un serveur de messagerie électronique. Il s'agit d'une distribution de logiciels libres (X (cross) Apache MariaDB 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 [44]. Cette « distribution » se chargera donc d'installer l'ensemble des outils qui à besoin lors de la création d'un site Web. Plus d'une dizaines d'utilitaires sont intégrés, comme MySQL, PHP, Perl ou encore phpMyAdmin. Il est distribué avec différentes bibliothèques logicielles qui élargissent la palette des services de façon notable : OpenSSL, Expat(parseur XML), PNG, SQLite, zlib,...

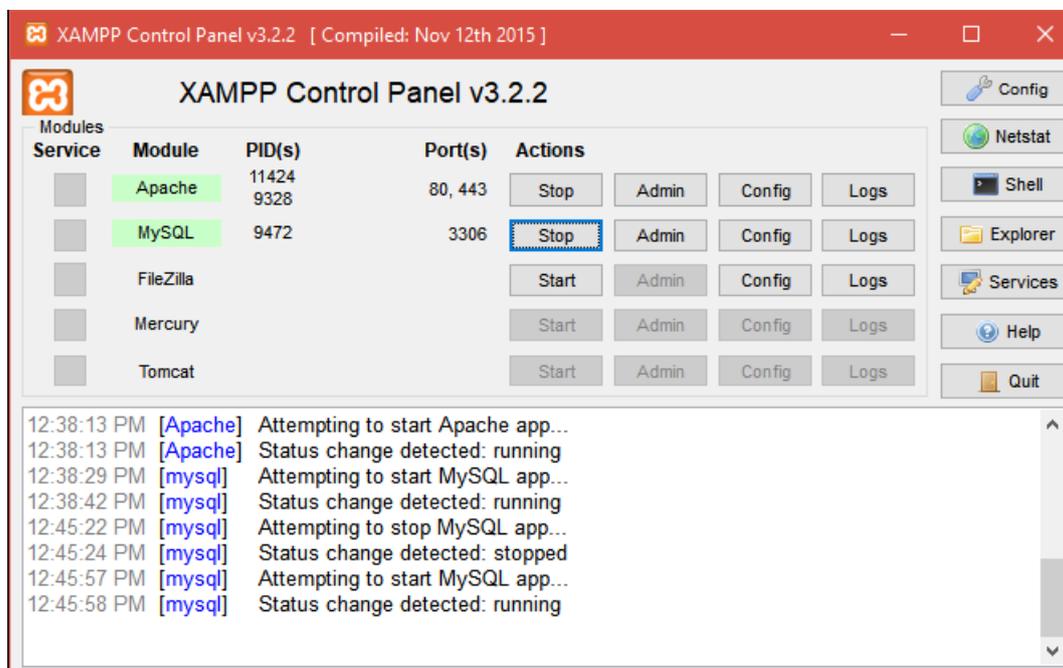


FIGURE 1 – Informations sur environnement du Xampp [18].

4.2.2 Apache

Noyau principal de XAMPP, l'option Apache crée un serveur Apache sur PC. C'est ce qui vous permettra d'exécuter et d'éditer des sites Web à l'aide de navigateurs Web, tout comme on le ferait pour un site. Sauf, bien sûr, que ce site fonctionne sur votre PC, par opposition à un serveur connecté à Internet.

Apache Web Serveur est conçu pour créer des serveurs Web pouvant héberger un ou plusieurs sites Web basés sur HTTP. Les fonctionnalités notables incluent la possibilité de prendre en charge plusieurs langages de programmation, les scripts côté serveur, un

mécanisme d'authentification et la prise en charge de la base de données.

Il est également largement utilisé par les sociétés d'hébergement Web dans le but de fournir un hébergement partagé/virtuel car, par défaut, Apache Web Serveur prend en charge et distingue différents hôtes résidant sur le même ordinateur [3].

4.2.3 phpMyAdmin

PhpMyAdmin (PMA) est une application Web de gestion pour les systèmes de gestion de base de données MySQL réalisée principalement en PHP et distribuée sous licence GNU/GPL [48].

Il s'agit de l'une des plus célèbres interfaces pour gérer une base de données MySQL sur un serveur PHP. De nombreux hébergeurs, le proposent ce qui évite à l'utilisateur d'avoir à l'installer.

Cette interface pratique permet d'exécuter, très facilement et sans grandes connaissances en bases de données, des requêtes comme les créations de table de données, insertions, mises à jour, suppressions et modifications de structure de la base de données, ainsi que l'attribution et la révocation de droits et l'import/export. Ce système permet de sauvegarder commodément une base de données sous forme de fichier .sql et d'y transférer ses données, même sans connaître SQL.

4.2.4 Android Studio

Android Studio est l'IDE recommandé pour développer des applications Android et est disponible gratuitement pour toute personne qui développe des applications Android professionnelles. Android Studio est basé sur le JetBrains Logiciel IntelliJ IDEA, ce qui pourrait expliquer pourquoi même les versions bêta et prévisualisation d'Android studio étaient déjà mieux que Eclipse et pourquoi de nombreux développeurs Android l'utilisaient comme leur IDE depuis le début. La première version stable d'Android Studio est sortie en décembre 2014 et a remplacé Eclipse (avec les outils de développement Android) en tant que principal IDE pour le développement Android. À présent, avec Android Studio, nous n'avons pas seulement un IDE plus stable et plus rapide, il est également livré avec trucs sympas tels que Gradle, de meilleures méthodes de refactoring, et un éditeur de mise en page bien meilleur pour nommer juste quelques uns. Il peut être téléchargé sous les systèmes d'exploitation Windows, macOS et Linux.

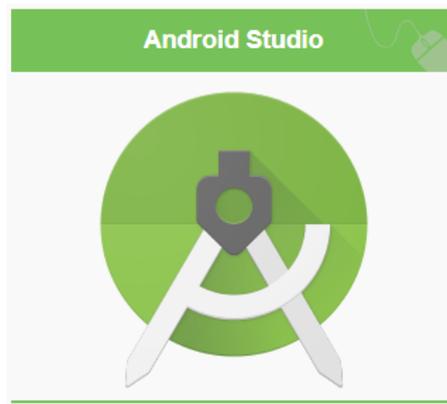


FIGURE 2 – Android studio [2]

Il propose entre autres des outils pour gérer le développement d'applications multilingues et permet de visualiser rapidement la mise en page des écrans sur des écrans de résolutions variées simultanément. Il intègre par ailleurs un émulateur permettant de faire tourner un système android virtuel sur un ordinateur. Dans notre système on besoins d'un émulator spécifiant pour les montres qui appelée "android wear".

4.2.5 Android Wear

Android Wear est une version spéciale du SDK Android et est dédiée aux périphériques portables souvent plus limités en matériel et en capteurs disponibles et ont des écrans plus petits. Andoid wear peuvent apparaître sous forme de montres, de lunettes ou peut-être à l'avenir de lentilles de contact, de tatouages ou de vêtements[59]. Est une nouvelle système d'exploitation spécialement conçue pour les montres, qui a été annoncée par Google en mars 2014. Actuellement, de nombreux appareils fonctionnant sous Android Wear ont été commercialisés, notamment les logiciels bien conçus Moto 360 et la Samsung Gear live et la LG G watch [27].

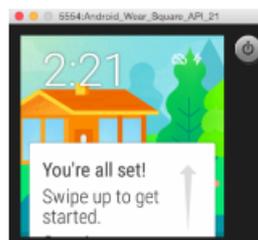


FIGURE 3 – Android studio wear[2]

4.2.6 Weka

Weka(Waikato Environment for Knowledge Analysis) est un environnement de fouille de données développé par le groupe de recherche "machine learning" du département d'informatique de l'université de Waikato en Nouvelle-Zélande. Il est utilisé dans le domaine de la recherche, de l'éducation et de l'industrie. Il est écrit dans le langage Java et testé sur plusieurs plateformes tels que Linux et Windows. Cet environnement est un logiciel "open source" et est disponible sur le site du groupe de recherche "machine learning" du département d'informatique de l'université de Waikato.[60]

Le format des données utilisées par les tâches de fouille de données est "arff". Weka possède plusieurs composants à savoir :

Explorer : ce module regroupe tous les packages importants de Weka à savoir le prétraitement, les algorithmes d'apprentissage, le groupement (clustering), les associations, la sélection des attributs et la visualisation.[54]

Expérimenter : permet d'exécuter plusieurs algorithmes d'apprentissage en mode lot (batch) et de comparer leurs résultats.

Knowledge Flow Environment : fournit les mêmes fonctionnalités que le composant "Explorer".

Ces fonctionnalités sont représentées sous forme graphique et sont utilisées pour construire un schéma de flux de connaissances via une interface drag-and-drop.



FIGURE 4 – Interface Weka[2]

4.2.7 Bibliothèque LibSVM

LIBSVM est une librairie dédiée aux Machines à Support de Vecteurs (Support Vector Machines). Son but est d'aider les utilisateurs à utiliser SVM facilement comme un outil. Elle est disponible en langage C++ et en JAVA, et compatible avec diverses plateformes logicielles (Python, R, MATLAB, Perl, Ruby, Weka, Common LISP, CLISP, Haskell, LabVIEW, interfaces PHP, C .NET, extensions CUDA). Cette bibliothèque est développée pour implémenter dans la résolution des problèmes liée à la classification avec l'utilisation des machines à vecteur support que ce soit la classification mono-class, binaire, multi-class et même pour la régression[12].

Le fonctionnement de LIBSVM se traduit par la définitions des paramètres d'entrée, il consiste à définir des paramètres d'entrée tels que le fichier d'exemples, le noyau à utiliser et ses paramètres, les paramètres de précision et de validation pour fournir un modèle de décision sous forme d'un fichier de type "model". Lors de l'utilisation du modèle fourni, LibSVM demande en plus du fichier modèle un fichier de test pour fournir un fichier contenant les classes prédites.

4.2.7.1 Paramètres SVM pris en charge par la librairie [26]

Afin d'entamer le processus d'apprentissage avec la LIBSVM, certains paramètres sont à renseigner selon qu'on souhaite faire une classification, une régression ou autre, le choix des bons paramètres est déterminant pour obtenir des résultats satisfaisants.

-s svm_type : C'est le type de l'algorithme SVM à utiliser, peut être l'une des fonctions : C_SVC, NU_SVC, ONE_CLASS, EPSILON_SVR, NU_SVR.

-t kernel_type : C'est le type de la fonction noyau à utiliser, peut être défini à : LINEAR, POLY, RBF ou SIGMOID.

- linéaire \bar{u}^*v
- polynomiale $(\gamma \bar{u}^*v + \text{coef0})^{\text{degré}}$
- Radiale $\exp(-\gamma \|u-v\|^2)$
- sigmoïde $\tanh(\gamma \bar{u}^*v + \text{coef0})$

Tels que : \mathbf{u}' :représente la transposé du vecteur contenant les valeurs des attributs de l'ensemble d'apprentissage, et \mathbf{v} le vecteur des labels (étiquettes). **Le gamma, degré et coef0** sont des paramètres (rentrés pas l'utilisateur).

Paramètres des fonctions noyau :

- d degree :** Paramètre degré de la fonction noyau , par défaut 3
- g gamma :** Paramètre gamma de la fonction noyau, par défaut 1
- r coef0 :** Paramètre coef0 de la fonction noyau, par défaut 0.

Paramètres dépendants du type SVM choisi :

-c cost : C'est le paramètre C (coût), qui représente la pénalité de l'erreur, à renseigner lors de l'utilisation du type SVM C-SVC, epsilon-SVR et nu-SVR, par défaut le coût est égal à 1.

-w weight : pour changer le paramètre C à $(weight * C)$, s'il n'est pas renseigné weight est égale à 1 sa valeur par défaut, et par conséquent neutre.

-n nu : Paramètre nu du type nu-SVC, One-class-SVM et nu-SVR, par défaut 0.5

-p epsilon : Paramètre epsilon de la fonction de perte (Loss Function) pour le type epsilon-SVR, par défaut égal à 0.1 .

4.2.8 Langages de programmation

4.2.8.1 PHP

PHP (PHP Hypertext PreProcessor) est un langage de programmation. Sa principale application se situe au niveau de la gestion des sites Web dynamiques [51].

Les capacités de PHP ne s'arrêtent pas à la création de pages Web. Il est aussi possible de manipuler des images, de créer des fichiers PDF, de se connecter à des bases de données ou des serveurs LDAP, et même d'instantanée des objets Java. Un module annexe lui permet également de fournir des interfaces graphiques classiques.

PHP est à l'origine un langage de script qui a été conçu spécifiquement pour agir sur les serveurs Web. En ajoutant quelques lignes de PHP à une page HTML, le serveur exécute les instructions correspondantes pour écrire du code HTML à la place.

PHP est extrêmement performant et fiable, même selon les critères d'application critiques. Avec un seul serveur standard, on peut répondre à des millions de requêtes par jour. Pour des sites à très fort trafic, il existe diverses solutions permettant d'optimiser et d'améliorer les performances globales de PHP.

En 2002, PHP est utilisé par plus de 8 millions de sites Web à travers le monde, en 2007 par plus de 20 millions et en 2013 par plus de 244 millions.

4.2.8.2 Javascript

Javascript est un langage de script incorporé dans un document HTML, Historiquement il s'agit même du premier langage de script pour le Web. Ce langage est un langage de programmation qui permet d'apporter des améliorations au langage HTML en permettant d'exécuter des commandes du côté client, c'est-à-dire au niveau du navigateur et

non du serveur web.

JavaScript est largement utilisé dans la programmation Web et est implémenté dans tous les principaux navigateurs. En tant que langage de programmation, JavaScript prend en charge la programmation fonctionnelle avec des fonctions anonymes, qui sont largement utilisées pour gérer les événements du navigateur tels que les clics de souris [44].

4.2.8.3 JAVA

Java est un langage de programmation orienté objet créé par James Gosling et Patrick Naughton, employés de Sun Microsystems, présenté officiellement le 23 mai 1995 au SunWorld.

Java est conçu pour être orienté objet à partir de la base. La technologie objet a finalement trouvé sa place dans la programmation après une période de gestation de trente ans. Les besoins des systèmes distribués, basés sur le client et le serveur, coïncident avec les paradigmes encapsulés de la transmission de messages du logiciel basé sur les objets. Pour fonctionner dans des environnements de plus en plus complexes basés sur le réseau, les systèmes de programmation doivent adopter des concepts orientés objet. Java fournit une plate-forme de développement propre et efficace basée sur les objets [25].

Les programmeurs utilisant Java peuvent accéder aux bibliothèques existantes d'objets testés offrant des fonctionnalités allant des types de données de base aux interfaces réseau, en passant par les interfaces d'E/S et réseau, aux kits d'outils d'interface utilisateur graphique. Ces bibliothèques peuvent être étendues pour fournir un nouveau comportement.

La particularité et l'objectif central de Java est que les logiciels écrits dans ce langage doivent être très facilement portables sur plusieurs systèmes d'exploitation tels que Unix, Windows, Mac OS ou GNU/Linux, avec peu ou pas de modifications, mais qui ont l'inconvénient d'être plus lourd à l'exécution (en mémoire et en temps processeur) à cause de sa machine virtuelle. Pour cela, divers plateformes et frameworks associés visent à guider, sinon garantir, cette portabilité des applications développées en Java.

4.2.8.4 SQL

Le SQL (Structured Query Language) est un langage permettant de communiquer avec une base de données. Ce langage informatique est notamment très utilisé par les développeurs web pour communiquer avec les données d'un site web [28]

Fonctionnalité de SQL dans un site web

- SQL peut exécuter des requêtes sur une base de données
- SQL peut récupérer des données d'une base de données
- SQL peut insérer des enregistrements dans une base de données

- SQL peut mettre à jour des enregistrements dans une base de données
- SQL peut supprimer des enregistrements d'une base de données
- SQL peut créer de nouvelles bases de données
- SQL peut créer de nouvelles tables dans une base de données
- SQL peut créer des procédures stockées dans une base de données
- SQL peut créer des vues dans une base de données
- SQL peut définir des autorisations sur les tables, les procédures et les vues.

Utilisation de SQL dans un site web Pour créer un site Web affichant les données d'une base de données, vous aurez besoin des éléments suivants : **Utilisation de SQL**

- Un programme de base de données RDBMS (c-à-d. MS Access, SQL Server, MySQL)
- Pour utiliser un langage de script côté serveur, tels que PHP ou ASP,
- Utiliser SQL pour obtenir les données souhaitées,
- Pour utiliser HTML / CSS pour styliser la page.

4.3 Interface de l'application de Surveillance Médicale à Distance (SMD)

L'IHM (Interface Homme Machine) représente un élément clé dans l'utilisation de tout système et conditionne pour une large part son succès. En théorie, une interface homme machine doit être ergonomique aussi bien que efficace.

De plus, ces interfaces doivent être faciles à utiliser et compréhensibles par les utilisateurs pour garantir un bon degré de fiabilité lors des interactions ainsi qu'un temps d'apprentissage réduit.

De plus, elles doivent avoir un certain niveau d'intelligence et de standardisation pour finir par présenter les fonctionnalités de l'application avec une manière plus conviviale. Dans cette partie nous décrivons les maquettes de l'application, c'est une étape très intéressante pour notre projet, la réalisation des maquettes est un loisir qui permet de comprendre le produit avant qu'il n'existe physiquement et de bien faciliter l'utilisation des interfaces de l'application.

Nous avons développé un ensemble d'applications qui mettent en œuvre notre système de surveillance des patients à distance. Les applications développées permettent l'accès au système à travers le web ou à travers d'application sur les montres intelligente. Le système proposé se compose donc des application suivantes :

- Application patient sur montre,
- Application patient sur web,
- Application médecin sur montre,

- Application médecin sur web,
- Application d'administration du système.
- Application d'entraînement pour la construction des modèles.
- Application d'utilisation de modèle (prédiction).

Ces applications peuvent être rassemblées en trois parties représentant chacune un acteur de notre système :

La figure suivante représente l'interface principale de notre système :

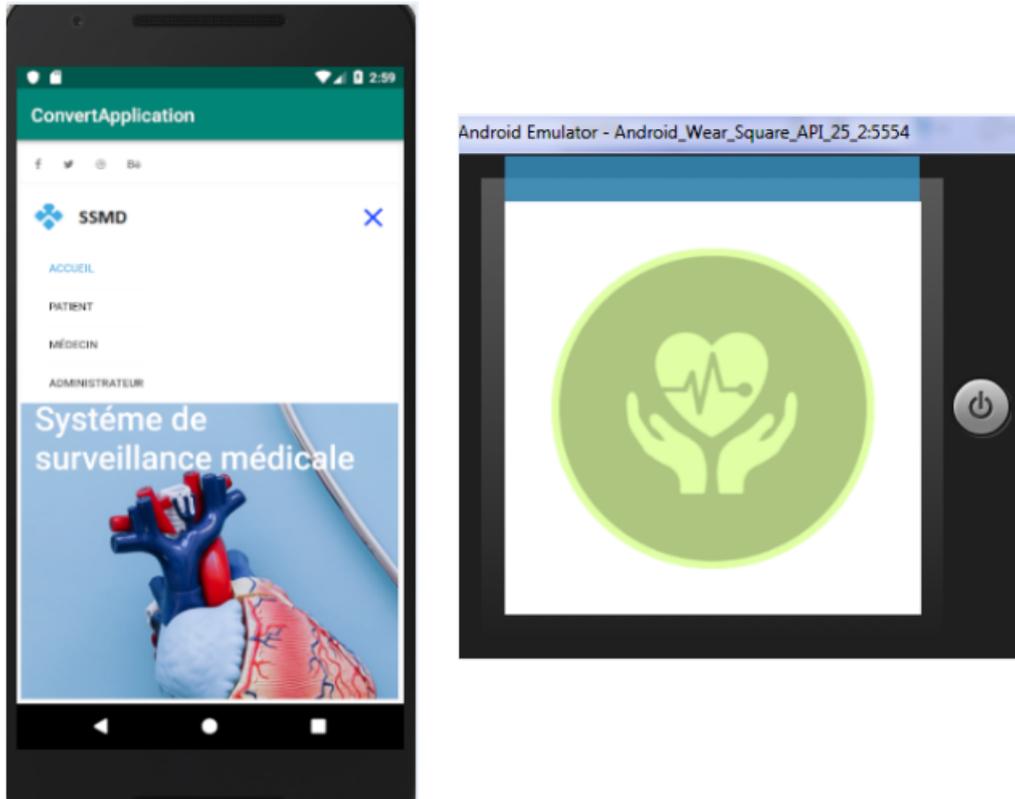


FIGURE 5 – Interface d'accueil du système(sur web, sur montre)

4.3.0.1 Application patient

L'application du patient permet quatre activités principales :

- Inscription du patient :

Le patient doit saisir toutes ses informations personnelles concernant son profil (son nom, son prénom, son sexe, son âge, son adresse,..).

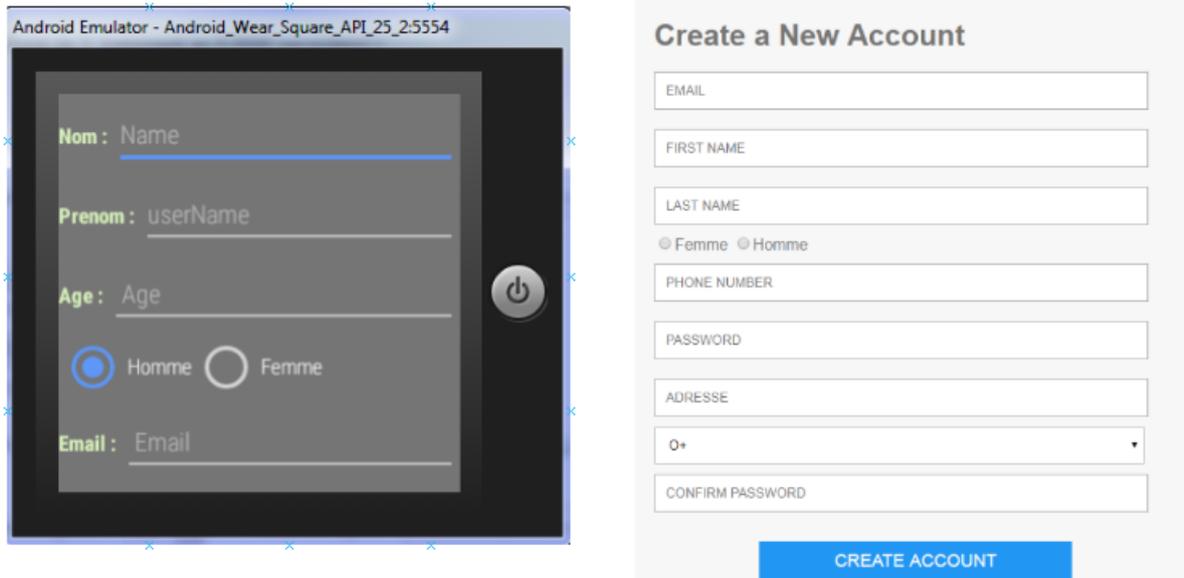


FIGURE 6 – Interface d'inscription de patients(sur web, sur montre)

- Connexion du patient : Une fois le patient enregistré, il pourra se connecter au système d'une façon sécurisé :

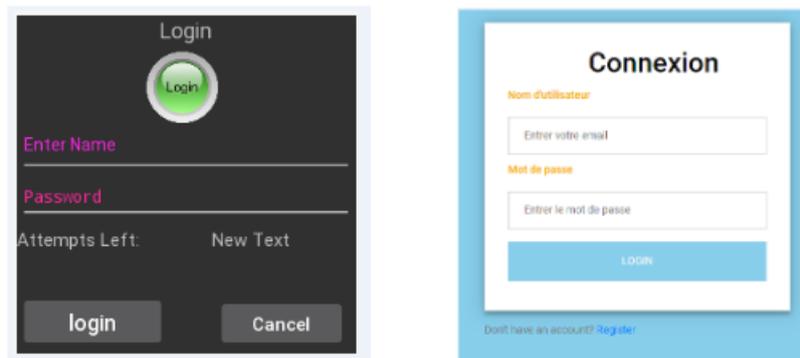


FIGURE 7 – Interface de connexion (sur web, sur montre)

- Acquisition de données :

La prise de mesure est faite périodiquement à la demande de serveur par les capteurs de la montre. La figure suivantes représentes la prise de nombre de battement du coeur par minute :



FIGURE 8 – Application mesuré "rythme cardiaque" (sur montre)

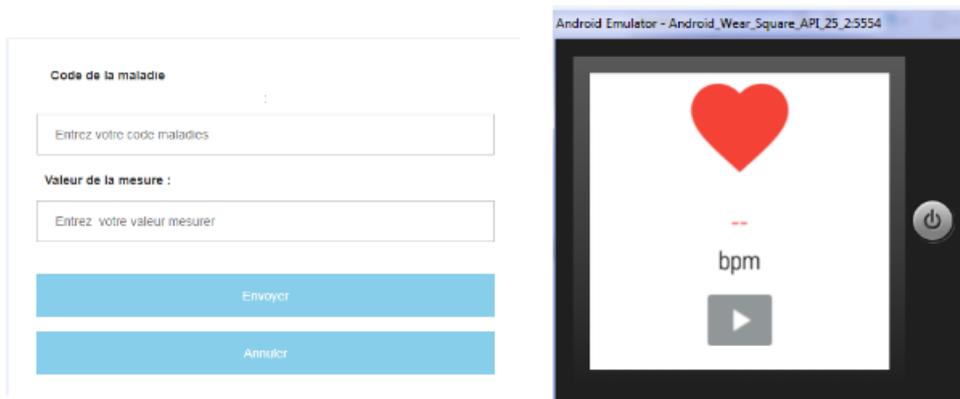


FIGURE 9 – Interface d'acquisition de données(sur web,sur montre)

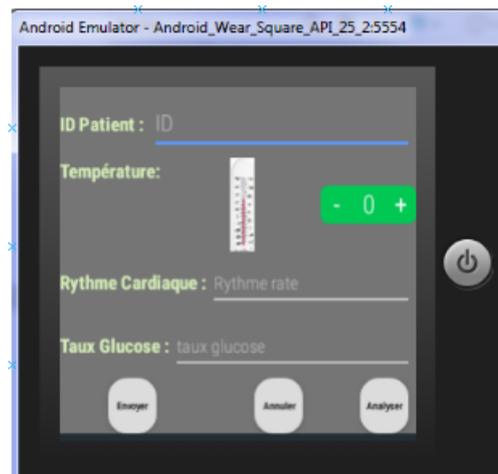


FIGURE 10 – Interface de saisie manuel des paramètres physiologique (sur montre)

— Affichage du bilan :

Le patient peut demander du système l'historique d'une mesure donnée pour une période spécifique.

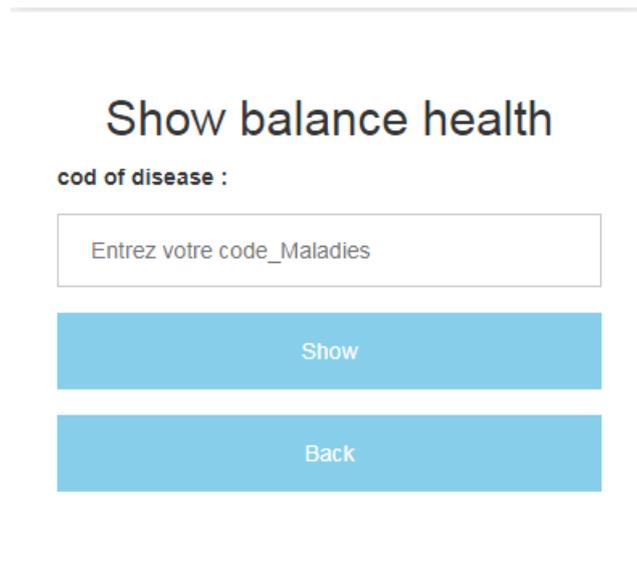


FIGURE 11 – Interface affichage du bilan (sur web)

Dans l'application du patient sur sa montre, l'interface principale permettant d'accéder au activités présentes lui apparait comme suit :

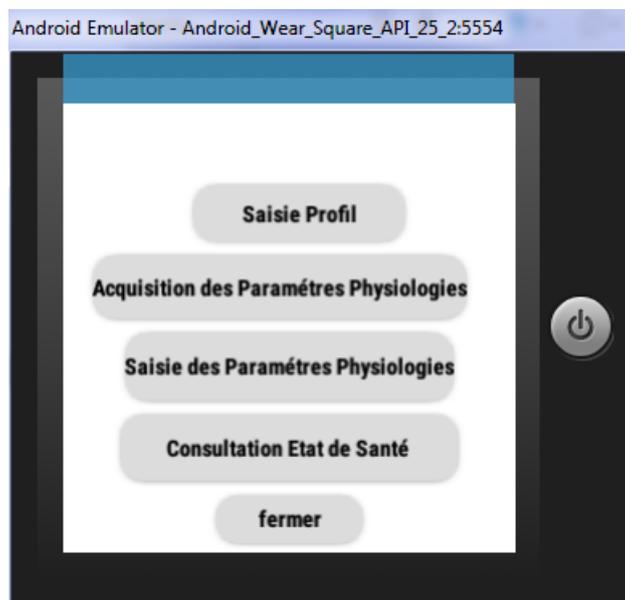


FIGURE 12 – Le menu de l'application patient(android wear)

4.3.0.2 Application Médecin

L'application du médecin lui permet, en plus de son inscription et sa notification, deux activités principales :

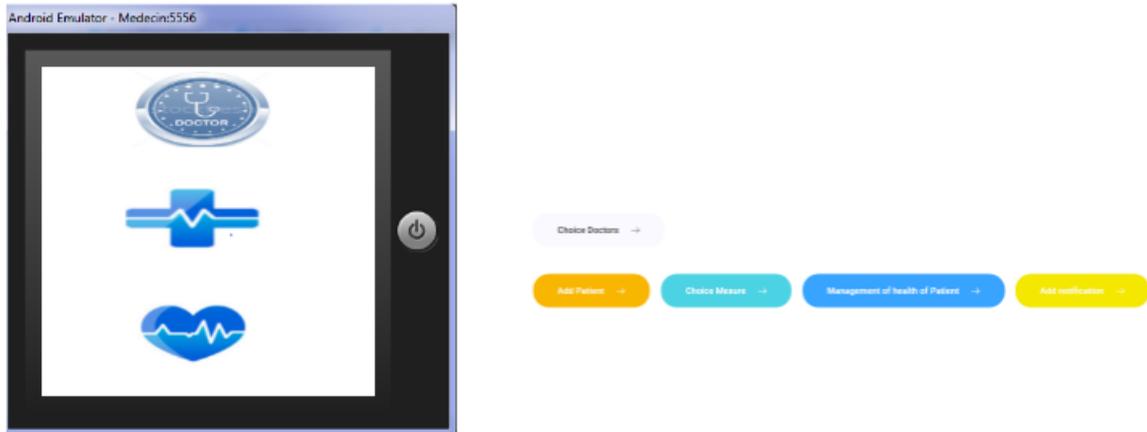


FIGURE 13 – Le menu de l'application médecin (sur montre)

— Inscription médecin :

Le médecin doit saisir toutes ses informations personnelles et professionnelles concernant son profil (son matricule correspondant à le code d'assurance, son prénom, son email, son adresse, son spécialité, son période d'expérience...).

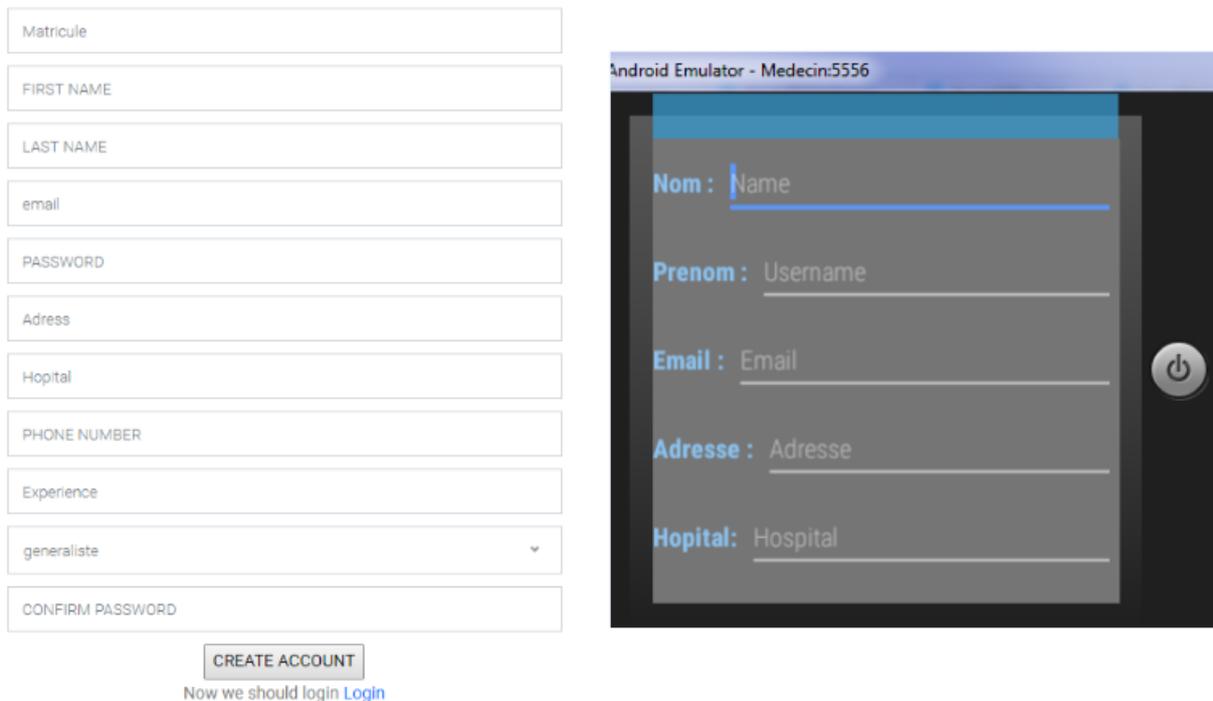


FIGURE 14 – Interface d'inscription du médecin (sur web, sur montre)

— Ajout de patients :

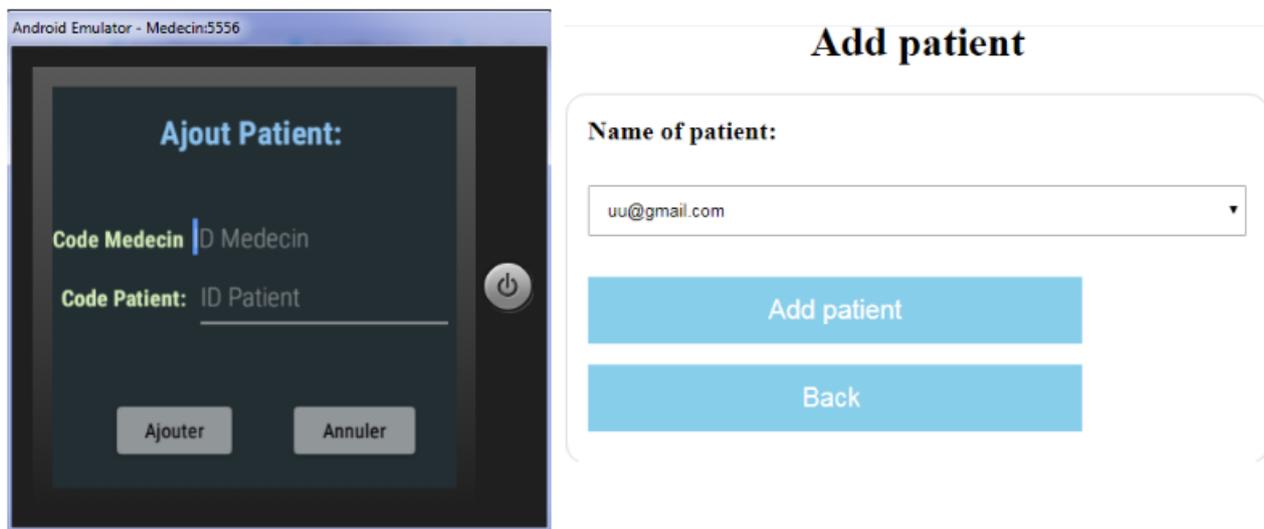


FIGURE 15 – Interface d'ajout du patients(sur web, sur montre)

— Ajout de mesures :

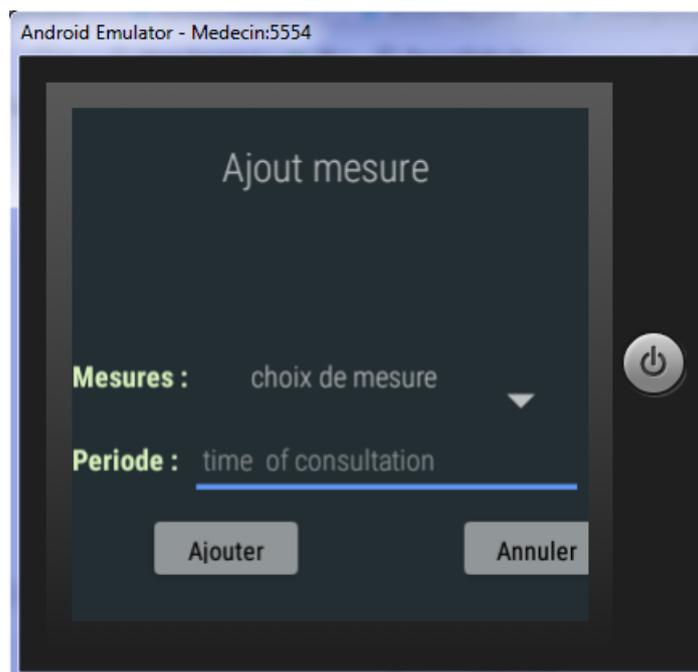


FIGURE 16 – Interface d'ajout de mesures (sur montre)

Choice of measure with user

Name of patient:

uu@gmail.com ▼

Select the appropriate measure for the user

blood pressure

heartbeat

Temperature

Diabetes

Periode of measure :

FIGURE 17 – Interface d'ajout de mesures (sur web)

4.3.0.3 Application d'administration

L'application d'administration permet à l'administrateur de systèmes les activités suivantes :

- Authentification

L'administrateur doit saisir son "username" et son "mot de passe" dans la page web :

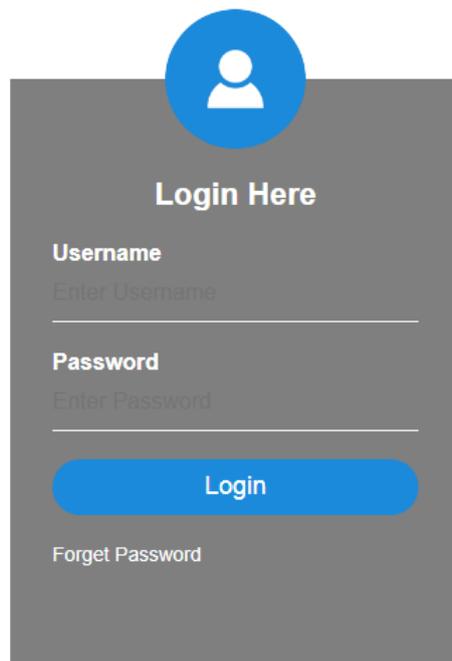
The image shows a web-based authentication interface. At the top center is a blue circular icon containing a white silhouette of a person. Below this icon, the text "Login Here" is displayed in a bold, white font. Underneath, there are two input fields: the first is labeled "Username" and contains the placeholder text "Enter Username"; the second is labeled "Password" and contains the placeholder text "Enter Password". Below these fields is a prominent blue button with rounded corners and the word "Login" in white text. At the bottom of the interface, there is a link labeled "Forget Password" in a smaller, lighter font. The entire interface is set against a dark gray background.

FIGURE 18 – Interface d'authentification (sur web)

— Construction des modèles

Dans cette phase, l'administrateur doit spécifier un fichier d'entraînement contenant les attributs des exemples de patients ainsi que leur classe et indiquer les paramètres de la méthode SVM pour construire un fichier modèle de type "model" pourtant le même nom de fichier d'entraînement.

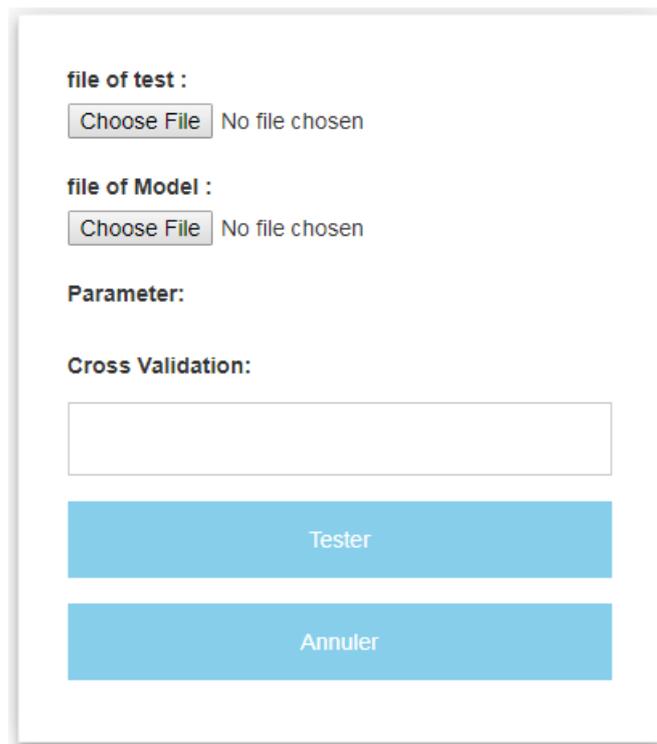
The image shows a web-based interface for model construction. It is divided into several sections:

- file train :** Contains a "Choose File" button and the text "No file chosen".
- Parameter:** This section contains two sliders:
 - Gamma:** A slider with a dark blue knob positioned at the left end, and a dark blue box below it displaying the value "100".
 - C:** A slider with a dark blue knob positioned towards the right end, and a dark blue box below it displaying the value "400".
- Buttons:** At the bottom of the interface, there are two large, light blue buttons with white text: "Analyser" and "Annuler".

FIGURE 19 – Page de construction des modèles (sur web)

— Utilisation de modèles :

Dans la phase d'utilisation, L'administrateur utilise l'outil "svm_predict.exe" pour prédire la décision concernant le vecteur de caractéristique d'un patient donné. Pour cela, il doit spécifier dans la page web le fichier modèle utiliser ainsi que le fichier contenant les valeurs de attributs. Le résultat sera un fichier contenant la listes des classes des exemples présent dans le fichier de test.



file of test :
Choose File No file chosen

file of Model :
Choose File No file chosen

Parameter:

Cross Validation:

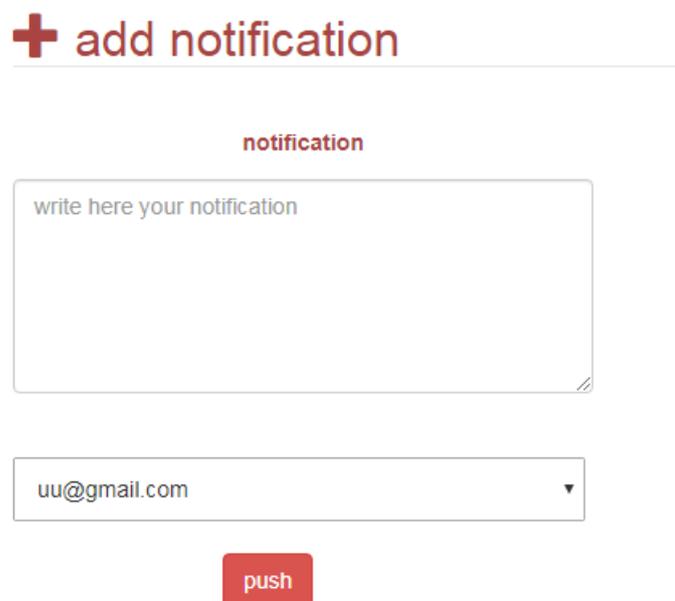
Tester

Annuler

FIGURE 20 – Page d'utilisation de modèles (sur web)

— Envoie notification :

L'administrateur après avoir prédit la classe d'un patient pourra lui envoyer une notification pour l'informer.



+ add notification

notification

write here your notification

uu@gmail.com

push

FIGURE 21 – Page d'ajouter des notifications (sur web)

4.4 Validation sur le cas des maladies cardiovasculaire

Pour valider notre proposition de système, nous avons choisi de l'appliquer pour le suivi des maladies cardiovasculaires. Nous avons utilisé pour cela une base de données externe disponible en accès libre sur un site de data mining.

Les maladies cardiaques ou les maladies cardiovasculaires (MCV) sont une classe de maladies qui impliquent le cœur ou les vaisseaux sanguins.

Les maladies cardiovasculaires comprennent les maladies coronariennes telles que l'angine de poitrine et l'infarctus du myocarde (communément appelé crise cardiaque). Les autres maladies cardiovasculaires comprennent les accidents vasculaires cérébraux, l'insuffisance cardiaque, les cardiopathies hypertensives, les cardiopathies rhumatismales, la cardiomyopathie, les arythmies cardiaques, les cardiopathies congénitales, les cardiopathies valvulaires, les cardites, les anévrismes aortiques, les artères périphériques, les thromboemboliques et les thromboses veineuses [43].

Les mécanismes sous-jacents varient en fonction de la maladie en question. La maladie coronarienne, les accidents vasculaires cérébraux et les maladies de l'artère périphérique impliquent l'athérosclérose.

Cela peut être causé par l'hypertension, le tabagisme, le diabète, le manque d'exercice, l'obésité, l'hyperdolicocéphale, une mauvaise alimentation et une consommation excessive d'alcool, entre autres. L'hypertension artérielle est responsable de 13% des décès par MCV, tandis que le tabac entraîne 9%, le diabète 6%, le manque d'exercice 6% et l'obésité 5%. La maladie cardiaque rhumatismale peut faire suite à une angine streptococcique non traitée [10].

La prévention de l'athérosclérose implique l'amélioration des facteurs de risque par une alimentation saine, l'exercice physique, la prévention du tabagisme et la réduction de la consommation d'alcool. Traiter les facteurs de risque, tels que l'hypertension artérielle, les lipides sanguins et le diabète est également bénéfique. Traiter les antibiotiques avec des antibiotiques peut réduire le risque de cardiopathie rhumatismale [42].

Les maladies cardiovasculaires sont la principale cause de décès dans le monde. Cela est vrai dans toutes les régions du monde, sauf en Afrique.

4.4.1 Base de données utilisé

Le figure présente une description de tous les attributs de la base de données :

Attributs	type	Min/Max	Moyenne
âge	réel	29/77	54.43
sexe	Binaire		
type de douleur thoracique	Nominal	(4 valeurs)	
tension artérielle au repos	réel	94/200	131.344
cholestérol sérique en mg / dl	réel	126/564	249.659
Glycémie à jeun > 120 mg / dl	Binaire		
résultats électrocardiographiques au repos	Nominal	(0, 1,2)	
fréquence cardiaque maximale atteinte	Réel	71/202	149.678
angine de poitrine induite par l'exercice	Binaire		
oldpeak = ST dépression induite par l'exercice par rapport au repos	Réel	0/6.2	1.05
la pente du segment ST du pic d'exercice	Ordered		
nombre de vaisseaux principaux (0-3) colorés par la fluorescence	réel	0/3	0.67
Thal	nominal	{normale; fixe réversible }	

FIGURE 22 – Description des attributs de la base utilisé

A des fins d'expérimentation, nous avons utilisé dans ce mémoire la base Heart deas[56] sous forme d'un fichier (Heart.csv).

Cette base disponible en accès libre sur le site de data mining UCI (référence de site) est une collection de rapports de diagnostic médicaux de 270 patientes âgées d'au moins de 29 ans, dont 120 cas sont en maladie chronique et 150 cas sain. Les 13 premiers attributs de la base représentent des facteurs de risque de la maladie tandis que le 14^{eme} représente la classe du patient (Absence (1) ou présence (2) de maladie cardiaque).

4.4.2 Pré-traitements

Pour pouvoir utiliser la base de données(Heart.csv) dans la plateforme Weka, nous avons appliqué un pré-traitement pour l'adaptation des types des attributs au format utilisé par weka. La figure suivante représente un exemple des enregistrements obtenus.

```
A1,A2,A3,A4,A5,A6,A7,A8,A9,A10,A11,A12,A13,A14
70.0,1.0,4.0,130.0,322.0,0.0,2.0,109.0,0.0,2.4,2.0,3.0,3.0,2
67.0,0.0,3.0,115.0,564.0,0.0,2.0,160.0,0.0,1.6,2.0,0.0,7.0,1
57.0,1.0,2.0,124.0,261.0,0.0,0.0,141.0,0.0,0.3,1.0,0.0,7.0,2
64.0,1.0,4.0,128.0,263.0,0.0,0.0,105.0,1.0,0.2,2.0,1.0,7.0,1
74.0,0.0,2.0,120.0,269.0,0.0,2.0,121.0,1.0,0.2,1.0,1.0,3.0,1
65.0,1.0,4.0,120.0,177.0,0.0,0.0,140.0,0.0,0.4,1.0,0.0,7.0,1
56.0,1.0,3.0,130.0,256.0,1.0,2.0,142.0,1.0,0.6,2.0,1.0,6.0,2
59.0,1.0,4.0,110.0,239.0,0.0,2.0,142.0,1.0,1.2,2.0,1.0,7.0,2
60.0,1.0,4.0,140.0,293.0,0.0,2.0,170.0,0.0,1.2,2.0,2.0,7.0,2
63.0,0.0,4.0,150.0,407.0,0.0,2.0,154.0,0.0,4.0,2.0,3.0,7.0,2
59.0,1.0,4.0,135.0,234.0,0.0,0.0,161.0,0.0,0.5,2.0,0.0,7.0,1
53.0,1.0,4.0,142.0,226.0,0.0,2.0,111.0,1.0,0.0,1.0,0.0,7.0,1
44.0,1.0,3.0,140.0,235.0,0.0,2.0,180.0,0.0,0.0,1.0,0.0,3.0,1
61.0,1.0,1.0,134.0,234.0,0.0,0.0,145.0,0.0,2.6,2.0,2.0,3.0,2
57.0,0.0,4.0,128.0,303.0,0.0,2.0,159.0,0.0,0.0,1.0,1.0,3.0,1
71.0,0.0,4.0,112.0,149.0,0.0,0.0,125.0,0.0,1.6,2.0,0.0,3.0,1
46.0,1.0,4.0,140.0,311.0,0.0,0.0,120.0,1.0,1.8,2.0,2.0,7.0,2
53.0,1.0,4.0,140.0,203.0,1.0,2.0,155.0,1.0,3.1,3.0,0.0,7.0,2
64.0,1.0,1.0,110.0,211.0,0.0,2.0,144.0,1.0,1.8,2.0,0.0,3.0,1
40.0,1.0,1.0,140.0,199.0,0.0,0.0,178.0,1.0,1.4,1.0,0.0,7.0,1
67.0,1.0,4.0,120.0,229.0,0.0,2.0,129.0,1.0,2.6,2.0,2.0,7.0,2
48.0,1.0,2.0,130.0,245.0,0.0,2.0,180.0,0.0,0.2,2.0,0.0,3.0,1
43.0,1.0,4.0,115.0,303.0,0.0,0.0,181.0,0.0,1.2,2.0,0.0,3.0,1
47.0,1.0,4.0,112.0,204.0,0.0,0.0,143.0,0.0,0.1,1.0,0.0,3.0,1
54.0,0.0,2.0,132.0,288.0,1.0,2.0,159.0,1.0,0.0,1.0,1.0,3.0,1
48.0,0.0,3.0,130.0,275.0,0.0,0.0,139.0,0.0,0.2,1.0,0.0,3.0,1
46.0,0.0,4.0,138.0,243.0,0.0,2.0,152.0,1.0,0.0,2.0,0.0,3.0,1
51.0,0.0,3.0,120.0,295.0,0.0,2.0,157.0,0.0,0.6,1.0,0.0,3.0,1
58.0,1.0,3.0,112.0,230.0,0.0,2.0,165.0,0.0,2.5,2.0,1.0,7.0,2
71.0,0.0,3.0,110.0,265.0,1.0,2.0,130.0,0.0,0.0,1.0,1.0,3.0,1
57.0,1.0,3.0,128.0,229.0,0.0,2.0,150.0,0.0,0.4,2.0,1.0,7.0,2
66.0,1.0,4.0,160.0,228.0,0.0,2.0,138.0,0.0,2.3,1.0,0.0,6.0,1
37.0,0.0,3.0,120.0,215.0,0.0,0.0,170.0,0.0,0.0,1.0,0.0,3.0,1
```

FIGURE 23 – Fichier Weka "Heart.csv"

4.4.3 Expérimentations et résultats

Nous avons mené plusieurs tests pour faire un tuning et choisir les bon paramètres permettant de donner les meilleurs résultats en terme de taux de reconnaissance et recall. Le taux de reconnaissance représentent le pourcentage du nombre d'exemple correctement classer dans les exemples de test tandis que le recall représente le pourcentage des exemples positifs correctement classé par rapport au nombre total des exemples positifs. Ce dernier permet de faire le compromis entre les erreurs positives (Un patient sain classé malade) et les erreurs négatives (les patients malades classés sains). Les paramètres utilisé pour le tuning sont les suivant :

- Données de test : le modèle est testé sur les données d'entraînement et sur les données de test externe.

- Le noyau utilisé :RBF et polynomial.
- le paramètre gamma pour le noyau RBF.
- le paramètres d (degrés de polynôme) pour le noyau linéaire.

La figure suivante représente l'évolution de la précision et le Recall en fonction du degré d en testant sur les données d'entraînement et en utilisant le noyau Polynomial de la méthode SMO :

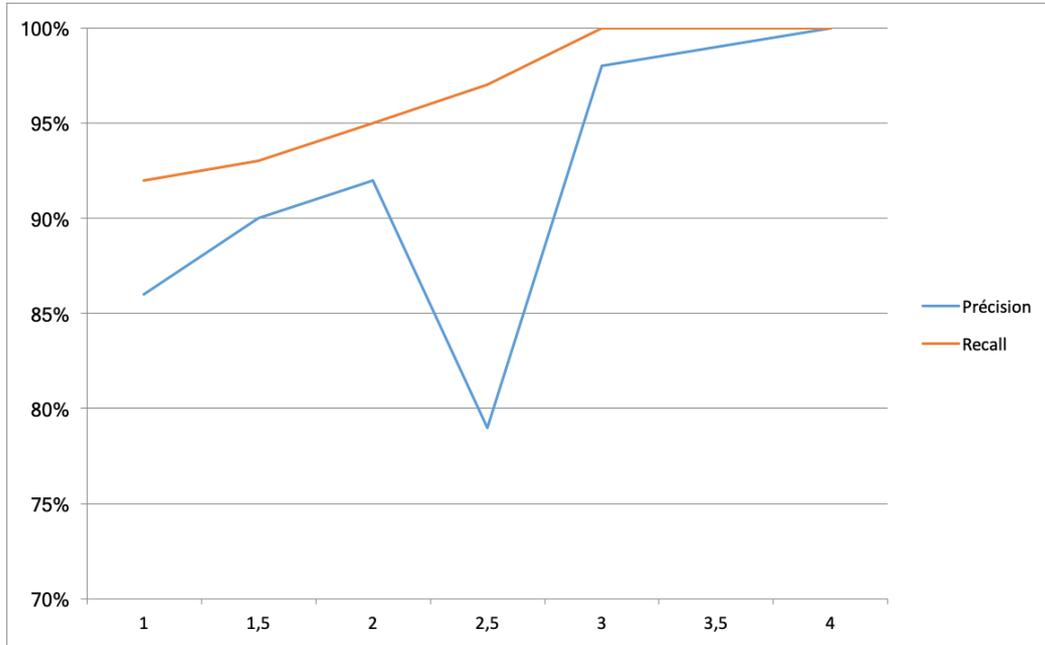


FIGURE 24 – Précision et Recall en fonction du degré d du noyau polynomial pour la méthode SMO

La figure suivante représente l'évolution de la précision et le Recall en fonction du degré d en testant sur les données d'entraînement et en utilisant le noyau Polynomial de la méthode LibSVM :

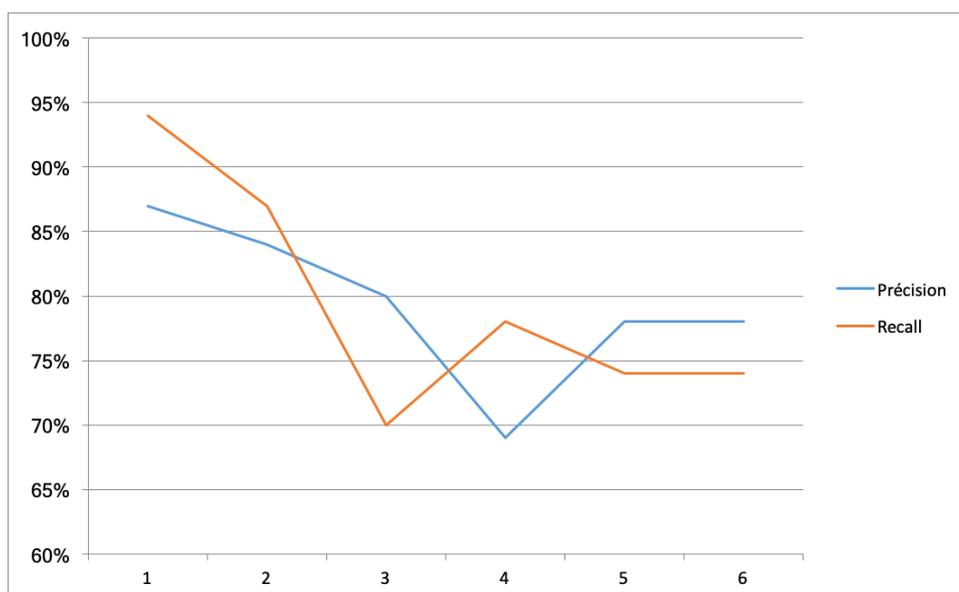


FIGURE 25 – Précision et Recall en fonction du degré d du noyau polynomial pour la méthode LibSVM

La figure suivante représente l'évolution de la précision et le Recall en fonction de gamma en testant sur les données d'entraînement et en utilisant le noyau RBF de la méthode SMO :

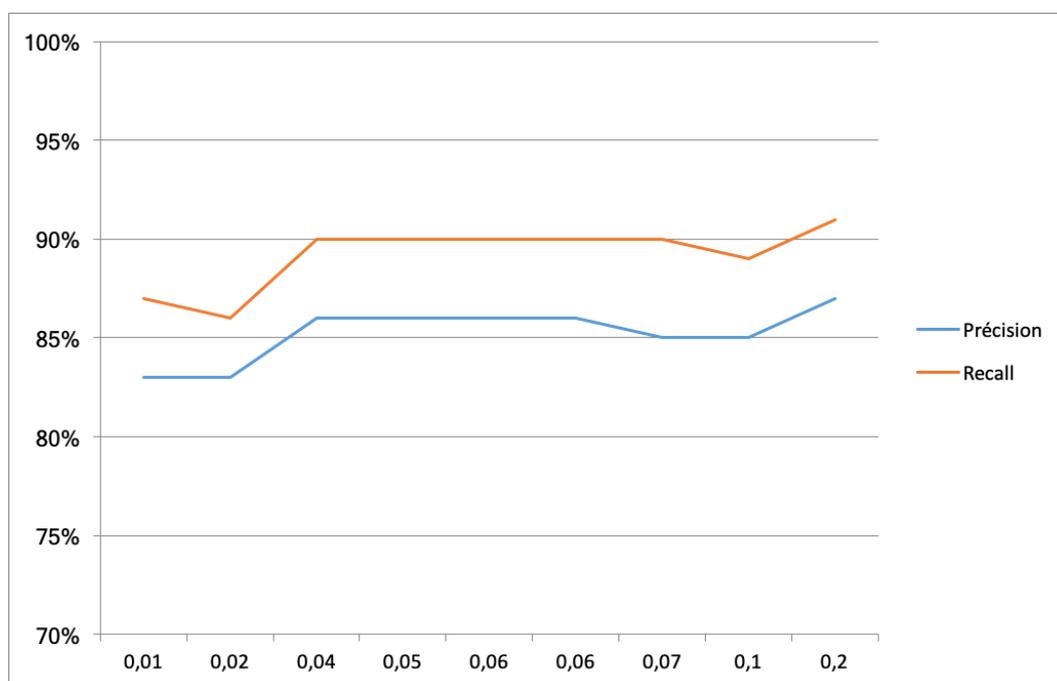


FIGURE 26 – Précision et Recall en fonction de gamma du noyau RBF pour la méthode SMO

La figure suivante représente l'évolution de la précision et le Recall en fonction de gamma en testant sur les données d'entraînement et en utilisant le noyau RBF de la méthode LibSVM :

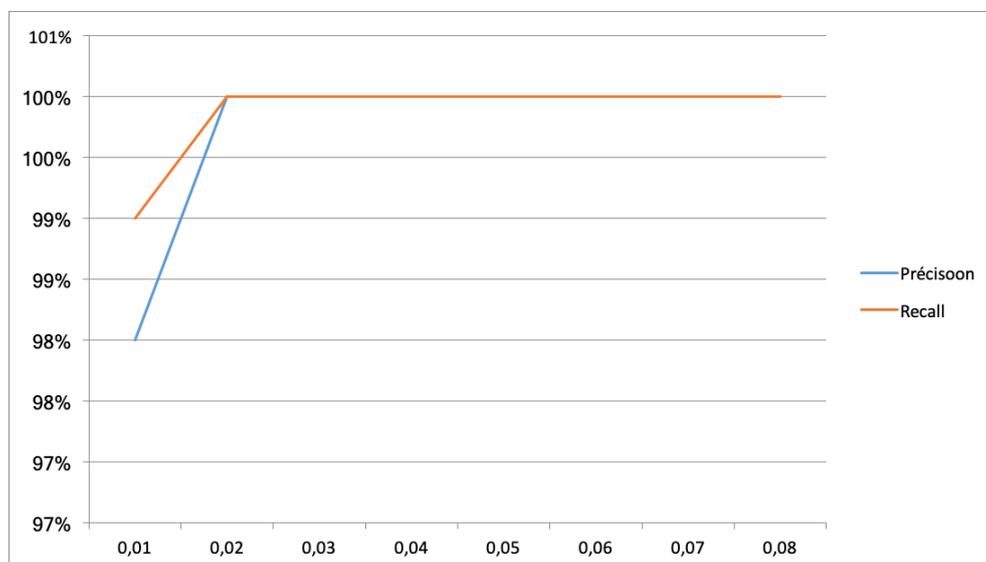


FIGURE 27 – Précision et Recall en fonction de gamma du noyau RBF pour la méthode LibSVM

La figure suivante représente l'évolution de la précision et le Recall en fonction de gamma en utilisant une validation croisée et en utilisant le noyau RBF de la méthode SMO :

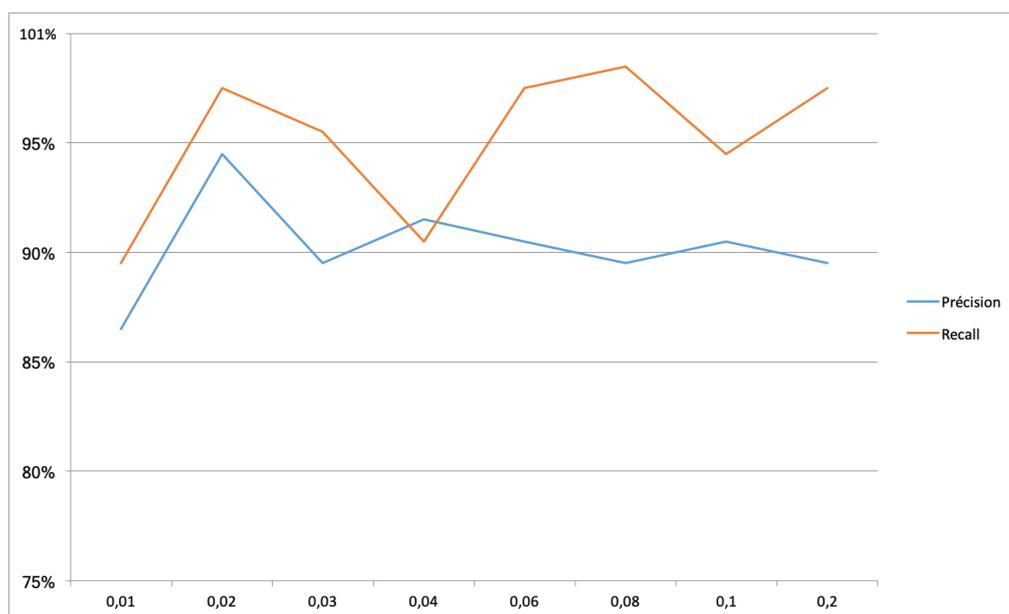


FIGURE 28 – Précision et Recall en fonction de gamma du noyau RBF pour la méthode SMO

La figure suivante représente l'évolution de la précision et le Recall en fonction de gamma en utilisant une validation croisée et en utilisant le noyau RBF de la méthode LibSVM :

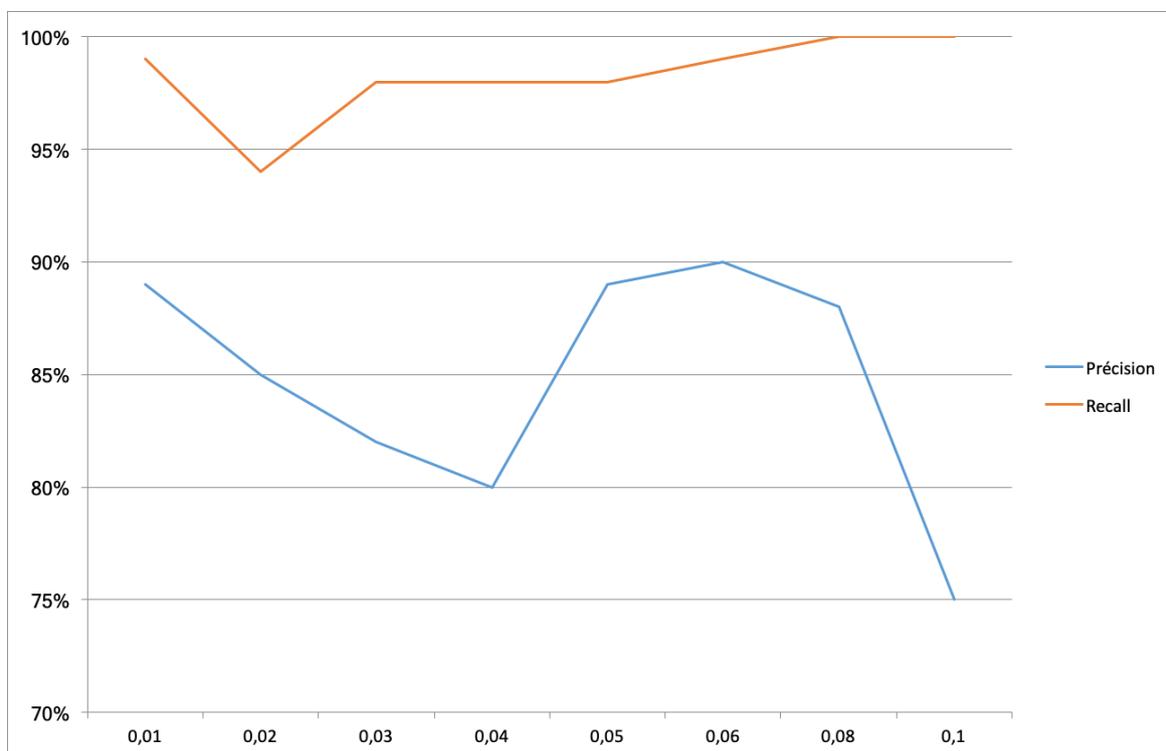


FIGURE 29 – Précision et Recall en fonction de gamma du noyau RBF pour la méthode LibSVM

4.4.4 Discussion

Des figures précédente, il s'avère que les meilleurs résultats pouvant être obtenus en terme de précision et recall sont celles utilisant le noyau RBF avec un $\gamma=0,08$ et la méthode LibSVM. La précision est alors de 90% et le recall est de 99% du fait du bon compromis qu'il présente entre la précision et le Recall.

A travers toutes nos expérimentations et résultats, nous pouvons dire que notre système proposé est capable d'identifier correctement de 99 %.

4.5 Conclusion

Dans ce dernier chapitre, nous avons représenté l'implémentation de notre système proposé : l'environnement, le langage de programmation et les outils de développement. Ensuite nous avons présenté quelques expérimentations effectuées et présenté les résultats obtenus et une discussion.

CONCLUSION GÉNÉRALE

Avec l'apparition de nouvelles technologies dans le domaine de la santé numérique, et la croissance de la population humaine, les systèmes de surveillance médicale sont devenus d'une importance capitale. Nous avons étudié dans ce travail ces systèmes et proposé une solution basée sur les montres intelligentes.

Nous avons conçu et réalisé dans ce travail, un système de surveillance médicale à distance mettant en collaboration trois applications principales : patient, médecin et administrateur.

Les applications du patient et du médecin sont localisées sur les montres intelligentes et sont chargées de la collecte des données des patients et leur notification avec les médecins. L'application de l'administrateur est localisées sur un serveur distant et chargé de construire et d'utiliser des modèles de décision automatique pour la détection précoce des maladies.

Nous avons à la fin validé notre proposition par une application sur le cas de suivi des maladies cardiovasculaires. Nous avons utilisé pour cela une base d'entraînement de plus 270 exemples représentant les maladies cardiaques. Le taux de reconnaissance obtenus sur les données d'entraînement dépassent les 97% ce qui est très encourageant et démontre l'efficacité de la méthode proposée. Un tel système offre des avantages certains, particulièrement :

- Diminution des coûts d'hospitalisation,
- Augmentation du degré d'accessibilité à l'aide des montres intelligentes et des réseaux,
- Réduction des coûts des soins ;

Pour les travaux futur, nous suggérons quelques idées qui peuvent améliorer notre système telles que :

- L'amélioration des application sur les montres pour un meilleur paramétrage du système,
- L'application du système dans des cas réels afin de construire des bases d'entra-

nement.

- Renforcer le système d'administration par d'autres méthodes de machine learning pour améliorer sa précision.

Bibliographie

- [1] Catherin Quantine ALIAN VENOT Anita Burgun. *Informatique médicale e_santéFondementetappli*. Springer-Verlag membre du groupe Springer Science Business Media, 2013.
- [2] *Android studio*. 28 mai 2019. URL : <http://www.smartwatch.me/t/tutorial-how-to-develop-android-wear-apps-for-beginners-part-1-setup/684>.
- [3] *Apache Web Server*. june 24, 2019. URL : <https://www.techopedia.com/definition/4851/apache-web-server>.
- [4] *Apple Watch Series 4*. URL : <https://www.apple.com/ca/fr/apple-watch-series-4/health/> (visité le 2019).
- [5] Yoshua BENGIO. *Une source pour la traduction en apprentissage profond*. Université de Montréal.
- [6] Michael BERRY et Gordon LINOFF. *Mastering data mining : The art and science of customer relationship management*. John Wiley, 1999.
- [7] BLUEGARD. “BlueGard : Braacelet de géoassistance”. In : (decembre 2018). URL : <http://www.bluelinea.com/Bluelinea/BlueGard.html>.
- [8] BOSCH. « *Système de télé-assistance Bosch* » : *Le poste de télé-assistance Bosch HTS3100*. Consulté le :2019.
- [9] Saraju P. Mohanty BY PRABHA SUNDARAVADIVEL Elias Kougianos et Madhavi K. GANAPATHIRAJU. “Everything you wanted to know about smart health care,Evaluating the different technologies and components of the Internet of Things for better health.” In : *IEEE Consumer Electronics Magazine* (january 2018).
- [10] *Cardiovascular disease*. 9 June 2019. URL : https://en.wikipedia.org/wiki/Cardiovascular_disease.
- [11] LAMICHE CHAABANE. “Fusion et fouille de données par les connaissances :Application a l’analyse d’image”. Thèse de doct. Universite Mohamed Khider, 2013.

- [12] Chih-Chung CHANG et Chih-Jen LIN. “Libsvm : A library for support vector machines.” Thèse de doct. Department of Computer Science National Taiwan University, Taipei, Taiwan, 2001.
- [13] Mme Martine CLAYER-FOUET. *Les technologies numériques au service de la santé*. la Commission ” Santé-Social” A/HRC/27/37. Version 23 février 2016. l’ouest de la France : Conseils économique sociale Environnemental Pays de la loire, 2016.
- [14] Christopher CLIFTON. *Encyclopædia Britannica : Definition of Data Mining*. 2010.
- [15] Y. Kodratoff CORNUÉJOLS L. Miclet. *Apprentissage artificiel*. Eyrolles.
- [16] Abd ElHamid DJEFFAL. *Fouille de données Avancé*. cours 2018/2019.
- [17] Abdelhamid DJEFFAL. “Utilisation des méthodes Support Vector Machine (SVM) dans l’analyse des bases de données”. Thèse de doct. Université Mohamed Khider-Biskra, 2012.
- [18] Dalibor D DVORSKI. “Installing, configuring, and developing with Xampp”. In : *Skills Canada* (2007).
- [19] *E-santé : la technologie au service de la santé*. Avr 5, 2018. URL : <https://www.dreaminzzz.com/fr/e-sante-la-technologie-au-service-de-la-sante>.
- [20] Nadia EL-MABROUK. *Introduction à la Bio-Informatique*. Universités IFT3295/IFT6291/BIN6000 Version 26 ouat 2013. Montréal : DIRO, Université de, 2013.
- [21] Usama FAYYAD, Gregory PIATETSKY-SHAPIO et Padhraic SMYTH. “From data mining to knowledge discovery in databases”. In : *AI magazine* 17.3 (1996), p. 37.
- [22] *Fichier technique Sumsung Gear S3*. 2019. URL : <https://www.frandroid.com/produits/samsung/montres-bracelet-connectes/6267-samsung-gear-s3/caracteristiques>.
- [23] FIOC. “Quelque technique de la fouille de données”. Master Pro. 2015/06.
- [24] FRÉDÉRIC-BOUSEFSAF. “MESURE SANS CONTACT DE L’ACTIVITÉ CARDIAQUE PAR ANALYSE DU FLUX VIDÉO ISSU D’UNE CAMÉRA NUMÉRIQUE”. Thèse de doct. UNIVERSITÉ DE LORRAINE, ÉCOLE DOCTORALE IAEM LORRAINE, LABORATOIRE DE CONCEPTION, OPTIMISATION ET MODÉLISATION DES SYSTÈMES, 2014.
- [25] James GOSLING et Henry MCGILTON. “The Java language environment”. In : *Sun Microsystems Computer Company* 2550 (1995).
- [26] Imane el HASSANI. “Svr avec boosting pour la prévision à long terme.” Thèse de doct. école polytechnique de l’université de tours, 2011-2012.
- [27] *how to develop android wear app for bignners*. juin 2019. URL : <https://www.commentcamarche.net/contents/577-javascript-introduction-au-langage-javascript>.

-
- [28] *Introduction to SQL*. 2019. URL : https://www.w3schools.com/sql/sql_intro.asp.
- [29] Rahman JAMAL et Herbert PICHLIK. *LabVIEW applications and solutions*. Prentice Hall PTR, 1999.
- [30] abdelbasset JARRAY. “conception et réalisation d’une application”. Thèse de doct. 2008.
- [31] J.L.FUNCK-BRENTANO. “L’informatique et la médecine de demain”. In : *journal de la sociétés statique Paris* 15 (1980), p. 1–9.
- [32] Braden Siempelkamp JOSH PIEMONTESE et Gordon MINAKER. “Smartwatches Biometrics for Personal Health Monitoring”. In : *SLIDO* (2009).
- [33] KODRATOFF. “techniques et outils de l’extraction de connaissances a partir des données”. In : *Revue Signaux* (1998).
- [34] Y. KODRATOFF. “techniques et outils de l’extraction de connaissances à partir des données”. In : 92 (Mars 1998), 38–43.
- [35] Frank Alexander KRAEMER. “Engineering android applications based on UML activities”. In : *International Conference on Model Driven Engineering Languages and Systems*. Springer. 2011, p. 183–197.
- [36] *La balise Aloïze*. [En ligne]. decembre 2018. URL : <http://www.iglfrance.com/alo45.php>.
- [37] Pierre LASBORDES. “La telesante :un nouvel atout au service de notre bien”. In : *Soins la revue de reference infirmiere* 750 (2010), p. 26.
- [38] *Les 10 meilleures applications médicales sur Android*. URL : <https://www.info24android.com/meilleures-applications-medicales-sur-android/>.
- [39] *Les meilleures montres connectées (2019)*. 7 févr 2019. URL : <https://www.techadvisor.fr/banc-essai/gadgets/meilleure-montre-connectee-3657340/#toc>.
- [40] Gordon S LINOFF et Michael JA BERRY. *Data mining techniques : for marketing, sales, and customer relationship management*. John Wiley, 2011.
- [41] Mokhinabonu MARDONOVA et Yosoon CHOI. “Review of Wearable Device Technology and Its Applications to the Mining Industry”. In : *Energy Review* (2018).
- [42] Henry C MCGILL JR, C Alex MCMAHAN et Samuel S GIDDING. “Preventing heart disease in the 21st century : implications of the Pathobiological Determinants of Atherosclerosis in Youth (PDAY) study”. In : *Circulation* 117.9 (2008), p. 1216–1227.

- [43] Shanthi MENDIS et al. *Global atlas on cardiovascular disease prevention and control*. Geneva : World Health Organization, 2011.
- [44] Michael MIKOWSKI et Josh POWELL. *Single page web applications : JavaScript end-to-end*. Manning Publications Co., 2013.
- [45] *Montre K'Watch Glucose et mesure du glucose en continu*. 19 Déc, 2018. URL : <https://diabete-infos.fr/kwatch-glucose-montre-mgc/>.
- [46] Marref NADIA. “Apprentissage Incrémental & Machines à Vecteurs Supports”. Thèse de doct. 2013.
- [47] ORANGE. “le bracelet-téléphone Columba : Retrouver les malades d’Alzheimer esituation d’errance”. In : (octobre 2018). URL : http://mobile.orange.fr/content/ge/high/rubriques/apropos/orange_care/bracelet.html.
- [48] *phpMyAdmin*. 10 mars 2019. URL : {<https://fr.wikipedia.org/wiki/PhpMyAdmin>}.
- [49] DIAF RADJA. “Système Embarqué Communicant pour la Surveillance Médicale des Personnes à Domicile”. Thèse de doct. Université Badji Mokhtar-Annaba, 2013-2014.
- [50] Marie-Odile SAFON. *La e-santé Télésanté, santé numérique ou santé connectée*. Institut de recherche et documentation en économie de la santé, Mars 2018.
- [51] Marie-Odile SAFON. *Manuel PHP*. sur PHP.net, 7 novembre 2007.
- [52] Gilbert SAPORTA. *Introduction au Data Mining et à l’apprentissage statistique*. Paris, mai 2019.
- [53] DRID ABOU BAKR SEDDIK. “PATTERN RECOGNITION FOR HEALTH-CARE ANALYTICS”. Thèse de doct. University Mohammed_khider – BISKRA, 2017.
- [54] Swasti SIGNAL et George D. MONIKA JENA. “A study on weka tool for data preprocessing classification and clustering”. In : *International Journal of Innovative technology and exploring engineering (IJTee)* 2.(6) (2013).
- [55] *Sleep as Android*. visité le Mai 2019. URL : <https://sleep.urbandroid.org/documentation/core/sleep-tracking/>.
- [56] *Statlog (Heart) Data Set*. July 15, 2018. URL : [https://archive.ics.uci.edu/ml/datasets/Statlog+\(Heart\)](https://archive.ics.uci.edu/ml/datasets/Statlog+(Heart)).
- [57] *Test : Watch Series 4, Apple améliore encore la meilleure des montres connectées*. URL : <https://www.01net.com/tests/test-watch-series-4-apple-ameliore-encore-la-meilleure-des-montres-connectees-6057.html>.
- [58] Stephane TUFFERY. *Data mining et statistique décisionnelle : l’intelligence dans les bases de données*. Editions Technip, 2005.

- [59] Mike VAN DRONGELEN. *Android studio cookbook*. Packt Publishing Ltd, 2015.
- [60] Lechevallier Y. *WEKA, un logiciel libre d'apprentissage et de data mining*". INRIA-Rocquencourt.
- [61] DA ZIGHED, Y KODRATOFF et NAPOLI. "Extraction de connaissance a partir d'une base de données". In : *Bulletin LFIA* 1 (2001).
- [62] Z.MAMMERI. *Modélisation de comportements de système en UML*.