



République du Sénégal
Un Peuple – Un But – Une Foi



Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation



ECOLE POLYTECHNIQUE DE THIES

DÉPARTEMENT DU GENIE INFORMATIQUE ET TÉLÉCOMMUNICATIONS

Projet de Fin d'Etudes en vue de l'obtention du Diplôme d'Ingénieur de Conception

Option : GIT

Présenté par : Cheikh Sadibou DIOUF

Sujet :

**CONCEPTION ET RÉALISATION D'UNE APPLICATION D'AIDE À LA
PRISE DE DÉCISION POUR LE DIAGNOSTIC DES MALADIES
CARDIOVASCULAIRES**

Soutenu publiquement le 03 août 2019 devant le jury composé de :

Président	Dr Abdoulaye	GUISSE	EPT
Examinateurs	Dr Samba	SIDIBE	EPT
	Dr Ibrahima	GUEYE	EPT
Superviseur	Pr Oumar	NIANG	ISEP Diamniadio
Encadreur	Dr Ndeye Fatou	NGOM	EPT
Invité	M. Ibou	SENE	ATOS
	M. Adoul-Dalibou	ABDOU	UT

ÉCOLE POLYTECHNIQUE DE THIES

Ce mémoire intitulé :

**Conception et réalisation d'un système intelligent d'aide à la prise de décision
pour le diagnostic des maladies cardiovasculaires**

présenté par **Cheikh Sadibou DIOUF**

en vue de l'obtention du diplôme d'ingénieur de conception en génie informatique et
télécommunications
a été dûment accepté par le jury d'examen constitué de :

Dr Abdoulaye GUISSE, président

Dr Samba SIDIBE, examinateur

Dr Ibrahima GUEYE, examinateur

Pr Oumar NIANG, superviseur

Dr Ndeye Fatou NGOM, encadreur

DÉDICACES

Louange à Allah, Seigneur de l'univers et que la paix et la bénédiction soient sur Son Messager et Serviteur, Muhammad, ainsi que sur sa famille, ses compagnons et tous ceux qui suivent le chemin qu'il a tracé, jusqu'au Jour de la Résurrection.

Je dédicace ce modeste travail :

À mes très chers parents,

Aucune dédicace ne saurait être assez éloquente pour vous témoigner mon amour profond et exprimer ce que vous méritez pour tous les sacrifices que vous n'avez cessés de consentir depuis mon enfance.

Je vous remercie pour tout le soutien et l'amour que vous me portez et je prie pour que votre bénédiction m'accompagne toujours.

Que ce modeste travail soit l'exaucement de vos voeux tant formulés, le fruit de vos innombrables sacrifices, bien que je ne vous en acquitterai jamais assez.

Puisse Dieu, le Très Haut, vous accorder santé, bonheur et longue vie.

À Ousmane et à Awa,

En témoignage de l'attachement, de l'amour et de l'affection que je porte pour vous.

À l'ensemble de ma famille,

Autant de phrases et d'expressions aussi éloquentes soit-elles ne sauraient exprimer ma gratitude et ma reconnaissance envers vous, particulièrement envers mes tantes : Oumy, Adama, Rokhaya, Aïssatou, Gnagna, Ndeye Racky ; et envers ma grand-mère, Mame Bigué. Ce travail est aussi dédicacé à Abdoulaye, ainsi qu'à tous mes cousins et cousines. Je tiens également à dédicacer ce travail à mon défunt grand-père, El Hadj Moussa Diagne.

À Batouly, Ndongo et à l'ensemble de mes amis,

Vos prières, vos encouragements et votre soutien ont toujours été une grande source de motivation.

À la famille polytechnicienne, en particulier à mes promotionnaires.

REMERCIEMENTS

Je tiens tout d'abord à écrire ces quelques lignes en signe de gratitude et de profonde reconnaissance à tous ceux qui, de près ou de loin, ont contribué à la réalisation et l'aboutissement de ce travail.

Je voudrais dans un premier temps adresser mes remerciements à mon encadreur, **Docteur Ndeye Fatou NGOM**, professeure au département de Génie Informatique et Télécommunications à l'École Polytechnique de Thiès, pour sa patience, sa disponibilité et ses judicieux conseils qui ont contribué à alimenter ma réflexion.

Je tiens à témoigner toute ma reconnaissance à l'ensemble des professeurs du département du Génie Informatique et Télécommunications de l'École Polytechnique de Thiès qui ont contribué à ma formation pendant ces dernières années en me fournissant les outils nécessaires à la réussite de mes études universitaires.

Je tiens à remercier spécialement l'infatigable **Amadou FALL**, étudiant en 5e année de médecine à l'École Militaire de Santé, qui ne ménageait aucun effort pour m'apporter son aide en me donnant des explications en rapport avec la cardiologie. Son inestimable soutien m'a permis de mener à bien ce projet de fin d'études.

Je ne saurais terminer sans remercier l'ensemble des membres du jury qui ont accepté d'évaluer la qualité de mon travail.

RÉSUMÉ

Le besoin de traçabilité des soins dans le domaine médical s'avère capital pour le patient mais aussi pour les établissements de santé. Il est primordial que chacune des observations, chaque acte puisse être recensé. L'informatisation du dossier patient permet un archivage facilité et une recherche d'informations rapide. De plus, cette transformation digitale engendre le stockage d'une grande quantité de données.

Les données médicales enregistrées dans des bases de données renferment un grand nombre d'informations de types divers. Il peut s'agir de textes, d'images, de signaux etc. Cependant, ces informations ne sont généralement pas exploitées pour faire une analyse qui vise à améliorer la qualité des soins.

Notre projet vise donc à concevoir et à réaliser un système basé sur des techniques d'intelligence artificielle visant à assister les cardiologues dans l'élaboration de leurs diagnostics. Ce système d'aide à la décision se base sur des données contenues dans les dossiers patients informatisés : les électrocardiogrammes.

Ce présent rapport est donc voué à retranscrire dans les détails le travail réalisé au cours de notre stage de fin d'études au sein du Laboratoire du Traitement d'Images et des Systèmes Intelligents de l'École Polytechnique de Thiès.

Mots-clés : Intelligence artificielle, Apprentissage automatique, Dossier Patient Informatisé, Électrocardiogramme

ABSTRACT

The need for traceability of care in the medical field is crucial for the patient but also for health care institutions. It is essential that each of the observations, each act can be identified. The computerization of the patient data allows for easy archiving and a quick search for information. In addition, this digital transformation results in the storage of a large amount of data.

Medical data recorded in databases contain a large amount of information of various types. These can be texts, images, signals, etc. However, this information is generally not used to make an analysis that aims to improve the quality of care.

Our project therefore aims to design and implement a system based on artificial intelligence techniques to assist cardiologists in the development of their diagnoses. This decision support system is based on data contained in computerized patient records: electrocardiograms.

This report is therefore devoted to detailing the work carried out during our final internship at the Laboratoire du Traitement d'Images et des Systèmes Intelligents of École Polytechnique de Thiès.

Keywords : Artificial intelligence, Machine learning, Electronic Health Record, Electrocardiogram

TABLE DES MATIÈRES

DÉDICACES	i
REMERCIEMENTS	ii
RÉSUMÉ	iii
ABSTRACT	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	vii
LISTE DES FIGURES	viii
LISTE DES SIGLES ET ABRÉVIATIONS	x
CHAPITRE 1 INTRODUCTION GÉNÉRALE	1
1.1 Présentation de la structure d'accueil	2
1.2 Présentation du sujet	3
1.2.1 Contexte et problématique	3
1.2.2 Objectifs du sujet	4
1.3 Plan du mémoire	4
CHAPITRE 2 CONCEPTS DE BASE	6
2.1 Dossier Patient Informatisé	6
2.1.1 Définition du dossier patient	6
2.1.2 Généralités sur le Dossier Patient Informatisé	7
2.2 Électrocardiogramme	9
2.2.1 Définition de l'électrocardiogramme	9
2.2.2 Interprétation de l'électrogramme	9
2.3 Machine learning	11
2.3.1 Historique et définition	11
2.3.2 Quelques méthodes de machine learning	12
2.3.3 Applications du machine learning	13
2.3.4 Algorithmes de classification	15
2.3.5 Évaluation de la performance des algorithmes	20

CHAPITRE 3 ÉTUDE CONCEPTUELLE DE LA SOLUTION	22
3.1 Analyse des besoins	22
3.1.1 Les besoins fonctionnels	22
3.1.2 Les besoins non fonctionnels	22
3.2 Modélisation UML	23
3.2.1 Diagramme des classes	23
3.2.2 Diagrammes des cas d'utilisation	25
3.2.3 Diagramme de séquences	28
3.3 Workflow des données utilisées pour la classification	29
CHAPITRE 4 IMPLEMENTATION DE L'APPLICATION	33
4.1 Modèle d'apprentissage automatique	33
4.2 Architecture logicielle	36
4.2.1 Front-End	37
4.2.2 Back-End	40
4.3 Présentation de l'application	42
4.3.1 Interface de la liste des dossiers	42
4.3.2 Interface de visualisation d'un dossier patient	43
4.3.3 Interface de visualisation des détails d'une consultation	44
4.3.4 Interface de visualisation des électrocardiogrammes	46
4.3.5 Interfaces de messagerie et de visioconférence	46
CHAPITRE 5 CONCLUSION	50
5.1 Synthèse des travaux	50
5.2 Limitations de la solution proposée	50
5.3 Améliorations futures	51
RÉFÉRENCES	52

LISTE DES TABLEAUX

LISTE DES FIGURES

Figure 1.1	Logo du LTISI	2
Figure 2.1	Allure d'un électrocardiogramme normal [1]	9
Figure 2.2	Allure normale d'un battement de coeur isolé [2]	10
Figure 2.3	Principe de l'apprentissage automatique [13]	12
Figure 2.4	Marge optimale dans le cas d'une classification binaire	15
Figure 2.5	Arbre de décision du jeu de données Iris [13]	16
Figure 2.6	Classification avec la méthode des k-plus proches voisins [3]	17
Figure 2.7	Réseau de neurone biologique [19]	18
Figure 2.8	Réseau de neurone artificiel [19]	18
Figure 2.9	Réseau de neurones à plusieurs couches [19]	19
Figure 2.10	Réseau de neurones à plusieurs couches [19]	20
Figure 3.1	Diagramme des classes proposé	24
Figure 3.2	Diagramme des cas d'utilisation proposé	26
Figure 3.3	Cas d'utilisation "Consultation patient"	27
Figure 3.4	Diagramme de séquences proposé	29
Figure 3.5	Workflow des données de classification	31
Figure 4.1	Précision et rappel de la classification avec SVM	34
Figure 4.2	Précision et rappel avec les arbres de décision [13]	34
Figure 4.3	Précision et rappel avec kNN	35
Figure 4.4	Matrice de confusion de la classification avec le CNN	35
Figure 4.5	Précision et rappel avec le réseau de neurones	36
Figure 4.6	Architecture logicielle de l'application	37
Figure 4.7	Logo du HTML et du CSS	38
Figure 4.8	Logo de la bibliothèque React	39
Figure 4.9	Logo de Django REST framework	40
Figure 4.10	Interface de la liste des dossiers	42
Figure 4.11	Interface de visualisation d'un dossier patient	43
Figure 4.12	Interface de visualisation d'un dossier patient	44
Figure 4.13	Interface de visualisation des détails d'une consultation	45
Figure 4.14	Interface de visualisation des détails d'une consultation	45
Figure 4.15	Interface de visualisation d'un ECG	46
Figure 4.16	Interface de messagerie	47
Figure 4.17	Choix du correspondant	48

LISTE DES FIGURES

ix

Figure 4.18	Message d'invitation	48
Figure 4.19	Interface de visioconférence	49

LISTE DES SIGLES ET ABRÉVIATIONS

LTISI	Laboratoire du Traitement de l'Information et des Systèmes Intelligents
DPI	Dossier Patient Informatisé
ECG	électrocardiogramme
IA	Intelligence Artificielle
NLP	Traitement du Langage Nature
SVM	Support vector Machines
kNN	k-Nearest Neighbors
UML	Unified Modeling Language
HTML	HyperText Markup Language
CSS	Cascading Style Sheets
API	Application Programming Interface
SGBD	Système de Gestion de Base de Données
AAMI	Association pour l'Avancement de l'Instrumentation Médicale

CHAPITRE 1 INTRODUCTION GÉNÉRALE

La tendance actuelle est au digital. Partout, l'outil informatique occupe une place à part entière dans la vie au quotidien de l'être humain aussi bien sur le plan personnel que professionnel. L'une des principales mutations de cette transformation digitale est donc l'augmentation continue des quantités de données produites et stockées. En effet, qu'il s'agisse d'échanger ou bien de traiter des données, tout est exécuté de façon informatisée pour faciliter la réalisation de ces tâches, et par la même occasion, gagner un temps non négligeable. Ainsi, la majorité des secteurs d'activités cherchent à tirer le plus grand profit des technologies de l'information. C'est dans cette optique que le secteur de la médecine connaît, depuis le début des années 2000, une évolution des dossiers patients informatisés, qui sont une numérisation des dossiers patients classiques [14].

Les dossiers patients informatisés sont constitués d'informations administratives et médicales qui forment une base de données dans le sens où il s'agit d'un recueil de données, ou d'autres éléments indépendants, individuellement accessibles par des moyens électroniques ou par tout autre moyen. Cependant, l'un des défis majeurs auquel le monde de la médecine fait face aujourd'hui est la valorisation des données présentes dans les systèmes de gestion de dossiers patients informatisés. De nos jours, l'intelligence artificielle, discipline visant à permettre aux machines d'imiter une forme d'intelligence réelle, se retrouve implantée dans un nombre grandissant de domaines d'application.

Au fur et à mesure de son évolution, l'intelligence artificielle passe du simple chabot générique à une évaluation des risques dans le domaine des prêts bancaires ou des assurances, et même à une aide au diagnostic en médecine. Elle se retrouve donc parmi les moyens susceptibles de valoriser tout type de données. L'objectif de l'utilisation de l'intelligence artificielle dans un système de gestion de dossiers patients informatisés serait donc de faciliter le travail des médecins en se basant sur des données de consultations effectuées dans le passé. L'élaboration diagnostics fiables ou la prescription de traitements adéquats pourraient résulter de ce processus. Dès lors, le contenu des dossiers patients informatisés ne sera pas seulement utilisé pour la consultation de l'historique d'un patient, mais aussi pour bénéficier à l'élaboration de diagnostics dans le secteur de la médecine.

Cette introduction générale consistera d'une part à présenter l'environnement et le contexte du stage. D'autre part, elle exposera l'importance du sujet étudié en mettant l'accent sur la problématique, les objectifs définis ainsi que le plan de ce rapport de projet.

1.1 Présentation de la structure d'accueil

La structure d'accueil dans laquelle nous avons effectué notre stage, en vue de l'obtention du diplôme d'ingénieur de conception en informatique et télécommunications, est le Laboratoire du Traitement de l'Information et des Systèmes Intelligents (LTISI). Ce laboratoire, affilié au département du Génie Informatique et Télécommunications de l'École Polytechnique de Thiès, a été créé en 2014. Le LTISI comprend en son sein des membres dont nous pouvons citer :

- 1 professeur de rang A
- 7 maîtres de conférences
- 8 doctorants
- des chercheurs appartenant à d'autres établissements d'enseignement supérieur



Figure 1.1 Logo du LTISI

Les thématiques principales de recherche sur lesquelles se penchent les membres du LTISI sont détaillées ci-dessous :

- | | |
|------------|---|
| A : | Ingénierie des connaissances |
| | Systèmes distribués dynamiques |
| | Acquisition et représentation des connaissances |
| | Bases de données et cloud computing |
| B : | Cryptographie et sécurité |
| | Systèmes embarqués et sécurité |
| | Sécurité et vie privée |
| | Cryptographie post-quantique |

C : **Signaux, images et télécommunications**

Modélisation mathématique et numérique

Traitement et analyse de signaux

Instrumentation et systèmes de transmission

C : **Science des données**

Analyse de données

Big data

Apprentissage automatique, apprentissage approfondi

Dans le cadre de notre immersion de quatre mois au sein du LTISI, notre mission a été de concevoir et de d'implémenter un système intelligent d'aide à la décision médicale, basé sur des signaux électriques émis par le cœur. Ce système a par la suite été intégré dans une application de gestion de dossiers patients informatisés.

1.2 Présentation du sujet

Cette partie du travail abordera le contexte, la problématique ainsi que les objectifs que nous nous sommes fixés lors de la réalisation du projet.

1.2.1 Contexte et problématique

Le monde actuel fait face au phénomène de la transformation digitale. La transformation digitale, aussi appelée transformation numérique, désigne le processus qui permet aux entreprises d'intégrer toutes les technologies digitales disponibles au sein de leurs activités.

Les technologies employées ont pour vocation d'améliorer de nombreux aspects au sein des entreprises et des organisations. L'objectif visé est de mieux répondre aux besoins des clients et surtout de délivrer les bonnes informations au bon moment. C'est dans ce sens que le domaine de la médecine a adopté des solutions informatiques tels que la numérisation des informations contenues dans les dossiers patients. Cette transformation digitale permet de développer des opportunités dans tous le secteur médical à bien des égards :

- La notion de distance n'existe plus, l'information peut se être consultée instantanément et n'est pas contrainte par une zone géographique
- L'information et les contenus dématérialisés peuvent toucher un plus grand nombre de personnes et sans réelle limite

- La collaboration entre personnes, les contenus partageables et modifiables en temps réel par tous permettent d'obtenir et de renseigner des informations sur un même patient bien plus facilement
- L'automatisation de tâches répétitives, telle que la recherche d'antécédants, permet une meilleure optimisation du temps de travail

La transformation digitale dans le secteur médical est alors un profond changement qui vient pour répondre aux besoins futurs. Cependant, un des gros enjeux est la capacité à faire parler les données que les systèmes de gestion de dossiers patients ont accumulé durant de nombreuses années. En effet, Si la récolte de données ne nécessite pas de but prédéfini, il faut souligner toutefois que la valorisation requiert en revanche un travail important de raffinage et d'extraction. L'objectif est alors de transformer la donnée en information, laquelle pourra produire de la valeur si délivrée au bon moment dans un contexte adapté.

Les données collectées sont aujourd’hui de divers types : images, signaux, textes descriptifs etc. Et, les progrès technologiques de nos jours permettent de valoriser ces gros volumes d’informations grâce notamment à l’intelligence artificielle.

Le rôle de l’intelligence artificielle dans les soins de santé a fait couler beaucoup d’encre ces derniers mois et il n’y a aucun signe que l’adoption de cette technologie ralentira. En effet, l’intelligence artificielle dans le domaine des soins de santé a un énorme et vaste potentiel du fait de toutes l’utilité des tâches pouvant être réalisées grâce à l’apprentissage automatique.

1.2.2 Objectifs du sujet

L’objectif de ce travail est de réaliser une étude conceptuelle avant d’implémenter une application de gestion de dossiers patients informatisés intégrant des outils d’aide à la décision. Dans notre cas de figure, l’outil d’aide à la décision sera réservé au diagnostic de pathologies cardiaques, à partir de données sous forme de signaux : les électrocardiogrammes.

En plus de faciliter la gestion des dossiers patients informatisés, l’application devra par ailleurs intégrer un espace d’échange textuel et téléphonique pour permettre la communication du personnel soignant.

1.3 Plan du mémoire

Notre mémoire sera divisé en trois parties. D’abord, dans le deuxième chapitre, nous ferons l’état de l’art des concepts qui ont servi à la réalisation du projet. Des concepts comme les dossiers patients informatisés l’électrocardiographie et l’apprentissage automatique (machine

learning en anglais) y seront abordés.

Ensuite, le troisième chapitre abordera l'étude conceptuelle de la solution proposée. Il permettra de décrire les différentes étapes de la réalisation du projet.

Enfin, le quatrième chapitre sera consacré à un exposé sur l'implémentation de la solution. Nous y définirons notre architecture logicielle tout en justifiant le choix des technologies utilisées.

CHAPITRE 2 CONCEPTS DE BASE

Dans ce chapitre, nous aborderons et expliquerons les notions relatives à ce projet. Nous nous pencherons essentiellement sur trois concepts :

1. Dossier patient informatisé
2. Electrocardiogramme
3. Machine learning : dans cette partie nous expliquerons brièvement les algorithmes utilisés dans notre application pour la classification de pathologies

2.1 Dossier Patient Informatisé

2.1.1 Définition du dossier patient

De tout temps, les médecins ont éprouvé la nécessité de conserver une « mémoire écrite de toutes les informations médicales concernant le patient ». Si les premiers écrits retrouvés en France datent du XVIIIème siècle, le papier s'est vite imposé dans les années 1970 pour devenir une pratique réglementaire.

Cette obligation s'explique par un besoin de traçabilité et d'amélioration de la qualité des soins. Le dossier patient devient un outil de preuve en cas de litiges entre patients et soignants. Il décrit l'ensemble des actes et opérations réalisées lors du séjour à l'hôpital. En effet, « le dossier du patient est le lieu de recueil et de conservation des informations administratives, médicales et paramédicales, formalisées et actualisées, enregistrées pour tout patient accueilli, à quelque titre que ce soit, dans un établissement de santé ».

Un dossier patient classique comporte des informations comme :

- L'état civil du patient : Nom de naissance, nom marital, prénom, sexe, date de naissance, lieu de naissance, nationalité, adresse...
- Les éléments de facturation : sécurité sociale, assurances, données bancaires...
- Les pièces justificatives à fournir lors de l'admission : carte vitale, carte de mutuelle, attestation de tiers payant, carte d'identité, justificatif d'adresse...
- La désignation de la personne de confiance et de la personne à prévenir ;
- Les autorisations d'opérer pour les mineurs etc [4].

Cependant, nous constatons que l'utilisation d'un support papier est la cause de survenue de plusieurs risques à différentes étapes de la prise en charge :

- Création ou sortie du dossier par les archives et mise à disposition dans les différents services
 - retard dans la mise à disposition du dossier
- Mise en place du dossier de soins et du classeur médicament
 - risque de perte d'information et de temps
- Rangement et tri du dossier
 - erreur d'identité
 - erreur de classement
- Rangement des documents produits lors du séjour du patient
 - risque de perte d'information et de temps

Ces risques peuvent aujourd’hui être évités grâce notamment à la numérisation des informations auparavant enregistrées par écrit. Le dossier patient classique devient progressivement informatisé. On parle alors du Dossier Patient Informatisé (DPI).

2.1.2 Généralités sur le Dossier Patient Informatisé

Gage de modernité, la numérisation du dossier patient offre de nouvelles perceptives aux établissements de santé. La transition depuis un modèle exclusivement papier induit de profondes modifications organisationnelles.

Un Dossier Patient Informatisé (DPI) est la collecte systématisée d’information sur la santé des patients et de la population stockée dans un format numérique. Ces dossiers peuvent être partagés dans différents milieux de soins de santé. Les documents sont partagés par l’entre-mise de systèmes d’information reliés au réseau. Les DPI peuvent comprendre un éventail de données, y compris les données démographiques, les antécédents médicaux, les médicaments et les allergies, l’état vaccinal, les résultats des tests de laboratoire, les images radiologiques, les signes vitaux, les statistiques personnelles comme l’âge et le poids, et les renseignements de facturation.

Il y a dix ans, les Dossiers Patients Informatisés (DPI) étaient présentés comme la clé de l’amélioration de la qualité des soins. Aujourd’hui, les prestataires de soins utilisent les données des dossiers des patients pour améliorer la qualité des résultats grâce à leurs programmes de gestion des soins. La combinaison de plusieurs types de données cliniques provenant des dossiers de santé du système a aidé les cliniciens à identifier et à stratifier les patients atteints d’une maladie chronique. Le DPI peut améliorer la qualité des soins en utilisant les données et les analyses pour prévenir les hospitalisations chez les patients à risque élevé.

Comparaison entre le dossier patient classique et le DPI

Les formulaires pré-imprimés, la normalisation des abréviations et les normes de calligraphie ont été encouragés afin d'améliorer la fiabilité des dossiers médicaux sur papier. La numérisation des formulaires facilite la collecte de données pour les études épidémiologiques et cliniques. Cependant, la normalisation peut créer des défis pour la pratique locale. Dans l'ensemble, ceux qui ont des DPI, qui ont automatisé les notes et les dossiers, la saisie des ordonnances et l'aide à la décision clinique, ont moins de complications, moins de mortalité et de coûts.

Les DPI peuvent être mis à jour en permanence. Cela facilite la coordination de la prestation des soins de santé dans les établissements de santé non affiliés. De plus, les données d'un système électronique peuvent être utilisées de façon anonyme pour la production de rapports statistiques dans des domaines comme l'amélioration de la qualité, la gestion des ressources et la surveillance des maladies transmissibles en santé publique.

Cependant, l'utilisation des Dossiers Patients Informatisés comporte quelques risques. Parmi eux, nous pouvons citer le risque d'accès aux données par des personnes non autorisées. En effet, la transparence et l'accessibilité acquises grâce à l'adoption des dossiers médicaux électroniques peuvent accroître la facilité d'accès des professionnels de la santé, mais aussi la quantité de renseignements volés par des personnes non autorisées ou des utilisateurs mal intentionnés. C'est pour cette raison qu'un tel système doit répondre à certaines exigences de sécurité lors de sa conception.

Quelques avantages de l'utilisation du DPI

L'informatisation du Dossier Patient Informatisé permet :

- De faciliter la coordination des soins entre les différents professionnels de santé.
- De faciliter l'exercice professionnel quotidien par la fourniture d'outils de classification permettant de retrouver les informations rapidement selon plusieurs critères
- au patient d'accéder à son dossier à n'importe quel endroit du monde et en plusieurs langues
- D'apporter une aide à la décision, en utilisant les gros volumes de données stockés dans les bases de données

C'est justement sur ce dernier point que nous nous pencherons dans le cadre de ce projet. C'est un aspect essentiel qui doit être pris en compte lors de l'implémentation de DPI. Il améliorerait en effet la qualité des soins.

2.2 Électrocardiogramme

2.2.1 Définition de l'électrocardiogramme

Un électrocardiogramme (ECG) est le résultat d'un examen médical réalisé afin de mesurer l'activité électrique qui est à l'origine des battements du cœur. Cet examen, appelé électrocardiographie, est réalisé en plaçant des électrodes sur les bras, les jambes ainsi que la poitrine du patient. Il enregistre les impulsions électriques étant à l'origine de la contraction des cavités du cœur, et donc de la propulsion du sang dans tout l'organisme. [5]

L'électrocardiogramme normal est une tracé caractérisé par des ondes qui se répètent à intervalles réguliers comme le montre la figure 2.1 :



Figure 2.1 Allure d'un électrocardiogramme normal [1]

C'est un examen rapide ne prenant que quelques minutes, indolore et dénué de tout danger. Il peut être fait en cabinet de médecin, à l'hôpital, voire à domicile. Son interprétation reste cependant complexe et requiert une analyse méthodique et une certaine expérience du clinicien. Il permet de mettre en évidence diverses anomalies cardiaques et a une place importante dans les examens diagnostiques en cardiologie.

2.2.2 Interprétation de l'électrogramme

La lecture et l'interprétation d'un ECG requièrent une grande habitude qui ne peut être acquise par le médecin que par une pratique régulière. Il existe des logiciels livrés avec certains électrocardiographes pouvant aider au diagnostic, mais ils ne peuvent se substituer au médecin.

Un ECG normal n'élimine en aucun cas une maladie du cœur. Un ECG anormal peut être également tout à fait anodin. Le médecin ne se sert de cet examen que comme un outil parmi d'autres, permettant d'apporter des arguments pour étayer son diagnostic.

Le tracé électrique comporte plusieurs accidents répétitifs appelés « ondes », et différents intervalles entre ces ondes. L'analyse d'un ECG correspond à l'étude de ces différentes ondes et intervalles. Les ondes (voir figure 4.15) qui composent un ECG sont :

- **l'onde P** : correspond à la dépolarisation (et la contraction) des oreillettes, droite et gauche. On analyse sa morphologie (positive ou diphasique en V1 voire V2 et monophasique dans toutes les autres dérivations), sa durée (qui est de 0,08 à 0,1 seconde), son amplitude (inférieure à 2,5 mm en D2 et 2 mm en V et V2), son axe et sa synchronisation avec l'onde QRS. ;
- **l'onde ou le complexe QRS** : correspond à la dépolarisation (et la contraction) des ventricules, droit et gauche. L'onde Q est la première onde négative du complexe. L'onde R est la première composante positive du complexe. L'onde S est la deuxième composante négative. Elle est généralement courte pour un battement normal ;
- **l'onde Q** : marque la repolarisation des ventricules. [2, 6]

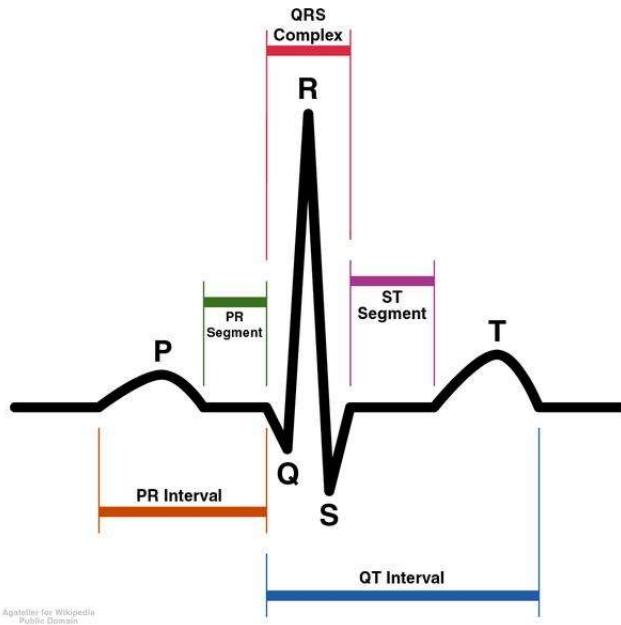


Figure 2.2 Allure normale d'un battement de coeur isolé [2]

Les électrocardiogrammes sont prescrits principalement pour la détection de troubles cardiaques. En effet, un ECG dont les caractéristiques sont anormales indique la présence d'une ou de plusieurs anomalies qui affectent le rythme cardiaque du patient. L'ECG permet de détecter de nombreuses anomalies :

- Du rythme cardiaque ;

- De la conduction de l'influx nerveux ;
- Du muscle cardiaque ;
- Des coronaires (artères irriguant le cœur) etc.

Le travail de détection est en général effectué par des spécialistes du domaine de la cardiologie. Le principal problème de l'analyse manuelle des signaux ECG réside dans la difficulté à catégoriser les différentes formes d'onde du signal. Pour un être humain, cette tâche prend beaucoup de temps et est sujette aux erreurs [15].

Pour résoudre ces problèmes, des études, telles que celles de Baloglu et al. [11], de Li et al. [17] et de Wang et al. [21], visant à analyser les ECG par le biais des ordinateurs ont été réalisées. Ces études reposent essentiellement sur des techniques basées sur l'intelligence artificielle, notamment le machine learning.

2.3 Machine learning

2.3.1 Historique et définition

Depuis l'antiquité, le sujet des machines qui pensent préoccupe les esprits. Ce concept est la base de pensées pour ce qui deviendra ensuite l'Intelligence Artificielle (IA), ainsi qu'une de ses sous-branches : le machine learning ou l'apprentissage automataque.

La concrétisation de cette idée est principalement due à Alan Turing et à son concept de la « machine universelle » en 19362, qui est à la base des ordinateurs d'aujourd'hui. Il continuera à poser les bases de l'apprentissage automatique, avec son article sur « L'ordinateur et l'intelligence » en 19503, dans lequel il développe, entre autres, le test de Turing.

Arthur Samuel, informaticien américain pionnier dans le secteur de l'intelligence artificielle, est le premier à faire usage de l'expression machine learning (en français, « apprentissage automatique ») en 1959 à la suite de la création de son programme pour IBM en 1952. Le programme jouait au Jeu de Dames et s'améliorait en jouant. À terme, il parvint à battre le 4e meilleur joueur des États-Unis.

Durant les années suivantes, les applications de l'apprentissage automatique se succèdent bien plus rapidement qu'auparavant.

L'apprentissage automatique est une application de l'intelligence artificielle (IA) qui permet aux systèmes d'apprendre automatiquement et de s'améliorer à partir de l'expérience sans être explicitement programmés. L'apprentissage automatique se concentre sur le développement de programmes informatiques qui peuvent accéder aux données et les utiliser pour apprendre par eux-mêmes. Tom Mitchell a donné une explication du concept en disant qu'

« un programme d'ordinateur apprend de l'expérience E par rapport à une tâche T et à une mesure de performance P, si sa performance sur T, telle que mesurée par P, s'améliore avec l'expérience E ». Le machine learning se base sur des méthodes de calcul utilisant les informations contenues dans un ensemble de données pour améliorer les performances dans la réalisation d'une tâche ou pour faire des prédictions précises.

Le processus d'apprentissage commence par des observations ou des données, telles que des exemples, des expériences directes ou des instructions, afin de rechercher des tendances dans les données et de prendre de meilleures décisions dans l'avenir en fonction des exemples que nous fournissons. L'objectif premier est de permettre aux ordinateurs d'apprendre automatiquement sans intervention ou assistance humaine et d'ajuster les actions en conséquence. Les algorithmes utilisés entraînent les modèles pour augmenter leur fiabilité en réduisant les erreurs de prédiction sur des données de test (figure 2.3).

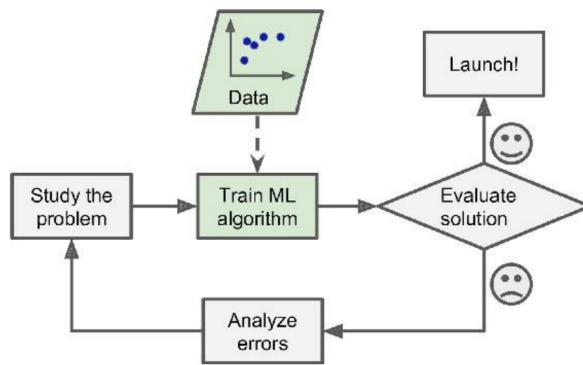


Figure 2.3 Principe de l'apprentissage automatique [13]

2.3.2 Quelques méthodes de machine learning

Les algorithmes d'apprentissage automatique sont souvent catégorisés comme étant supervisés ou non supervisés.

Les **algorithmes d'apprentissage automatique supervisés** peuvent appliquer ce qui a été appris dans le passé à de nouvelles données en utilisant des exemples étiquetés pour prédire les événements futurs. A partir de l'analyse d'un ensemble de données d'apprentissage connu, l'algorithme d'apprentissage produit une fonction inférée pour faire des prédictions sur les valeurs de sortie. Le système est en mesure de fournir des cibles pour tout nouvel intrant après une formation suffisante. L'algorithme d'apprentissage peut également comparer sa sortie avec la sortie correcte prévue et trouver des erreurs afin de modifier le modèle en conséquence.

En revanche, des **algorithmes d'apprentissage automatique non supervisés** sont utilisés lorsque l'information utilisée pour la formation n'est ni classifiée ni étiquetée. L'apprentissage non supervisé étudie comment les systèmes peuvent déduire une fonction pour décrire une structure cachée à partir de données non étiquetées. Le système ne trouve pas le bon résultat, mais il explore les données et peut tirer des inférences des ensembles de données pour décrire des structures cachées à partir de données non étiquetées.

Les **algorithmes d'apprentissage automatique semi-supervisés** se situent quelque part entre l'apprentissage supervisé et l'apprentissage non supervisé, puisqu'ils utilisent des données étiquetées et non étiquetées pour la formation - généralement une petite quantité de données étiquetées et une grande quantité de données non étiquetées. Les systèmes qui utilisent cette méthode sont capables d'améliorer considérablement la précision de l'apprentissage. Habituellement, l'apprentissage semi-supervisé est choisi lorsque les données étiquetées acquises nécessitent des ressources compétentes et pertinentes afin de les former et d'en tirer des enseignements. Sinon, l'acquisition de données non étiquetées ne nécessite généralement pas de ressources supplémentaires.

Les **algorithmes d'apprentissage automatique par renforcement** sont une méthode d'apprentissage qui interagit avec leur environnement en produisant des actions et découvre des erreurs ou des récompenses. La recherche d'essais et d'erreurs et la récompense différée sont les caractéristiques les plus pertinentes de l'apprentissage par renforcement. Cette méthode permet aux machines et aux agents logiciels de déterminer automatiquement le comportement idéal dans un contexte spécifique afin de maximiser ses performances. Une simple rétroaction de récompense est nécessaire pour que l'agent apprenne quelle action est la meilleure ; c'est ce qu'on appelle le signal de renforcement.

2.3.3 Applications du machine learning

L'apprentissage machine admet un très large éventail d'applications pratiques, dont les suivantes :

- Classification de texte ou de document : Cela inclut des problèmes tels que l'assignation d'un sujet à un texte ou à un document, ou la détermination automatique si le contenu d'une page Web est inapproprié ou trop explicite ; cela inclut également la détection du spam.
- Traitement du Langage Naturel (NLP) : La plupart des tâches dans ce domaine, y compris le marquage partiel de la parole, la reconnaissance d'entités nommées, l'analyse sans contexte ou l'analyse des dépendances, sont considérées comme des problèmes

d'apprentissage. Dans ces problèmes, les prédictions admettent une certaine structure. Par exemple, dans le balisage d'une partie de la parole, la prédiction d'une phrase est une séquence de balises de partie de parole qui étiquettent chaque mot. Dans l'analyse sans contexte, la prédiction est un arbre. Ce sont des exemples de problèmes d'apprentissage plus riches connus sous le nom de problèmes de prédiction structurée.

- Applications de traitement de la parole : Cela inclut la reconnaissance vocale, la synthèse vocale, la vérification du locuteur, l'identification du locuteur, ainsi que des sous-problèmes tels que la modélisation du langage et la modélisation acoustique.
- Applications de vision par ordinateur : Cela comprend la reconnaissance d'objets, l'identification d'objets, la détection de visages, la reconnaissance optique de caractères (ROC), la récupération d'images basées sur le contenu ou l'estimation de pose.
- Applications en biologie computationnelle : Cela comprend la prédiction de la fonction des protéines, l'identification des sites clés ou l'analyse des réseaux de gènes et de protéines.
- Beaucoup d'autres problèmes tels que la détection de fraude pour les sociétés de cartes de crédit, les compagnies de téléphone ou de services, l'intrusion dans le réseau, l'apprentissage de jeux comme les échecs, le backgammon ou le Go, le contrôle non assisté de véhicules comme les robots ou les voitures, le diagnostic médical, la conception de systèmes de recommandation, les moteurs de recherche ou les systèmes d'extraction d'information, sont abordés à l'aide de techniques d'apprentissage automatique.

Cette liste n'est en aucun cas exhaustive. La plupart des problèmes de prédiction rencontrés dans la pratique peuvent être considérés comme des problèmes d'apprentissage et le domaine d'application pratique de l'apprentissage machine ne cesse de s'étendre [18].

Dans le cadre de ce projet, nous appliquerons le machine learning en faisant une classification des électrocardiogrammes. Le rôle de l'algorithme utilisé sera d'attribuer à chaque nouvel ECG une classe parmi les différentes classes.

L'apprentissage machine est l'utilisation d'algorithmes pour acquérir des descriptions à partir d'exemples. Ces descriptions structurelles, aussi appelées modèles, sont construites pour contenir l'information extraite des données brutes, et sont utilisées pour prédire des résultats sur des données inconnues. Il existe plusieurs types de modèles que nous étudierons dans la section suivante.

2.3.4 Algorithmes de classification

Machine à vecteurs de support

Les machines à vecteurs de support ou Support vector Machines (SVM) sont un modèle d'apprentissage automatique polyvalent, capable d'effectuer des classifications linéaires et non linéaires, des régressions mais aussi des détections de valeurs aberrantes. [13]

Dans le cas d'une classification, l'optimisation des algorithmes de machines à vecteurs de support consiste en la recherche de l'hyperplan qui sépare le mieux les données. Cet hyperplan est appelé hyperplan optimal. Un hyperplan peut être défini comme étant un sous-espace d'une dimension inférieure à son espace ambiant. Par exemple, si les données sont contenues dans un plan (2 dimensions), l'hyperplan sera une droite contenue dans ce plan.

Les algorithmes implémentant les SVM sont aussi appelés séparateurs à vaste marge. En effet, afin de choisir le meilleur hyperplan, ils calculent une distance appelée marge. La marge représente la distance entre l'hyperplan de séparation et les échantillons de données les plus proches. Ces derniers sont appelés vecteurs supports. Dans les SVM, la frontière de séparation, ou encore l'hyperplan, est choisie comme celle qui maximise la marge (voir figure 2.4) [16].

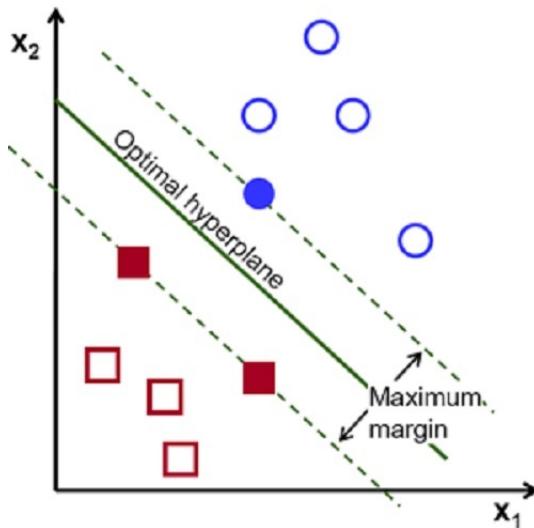


Figure 2.4 Marge optimale dans le cas d'une classification binaire

Arbres de décision

La méthodologie des arbres de décision a été introduite en 1984 par Leo Breiman, Jerome Friedman, Richard Olshen et Charles Stone, chercheurs à l'UC Berkley et Stanford. Le pro-

cessus des arbres de décision est structuré comme une séquence de questions simples. Les réponses à ces questions déterminent la question suivante. Le résultat est un réseau de questions qui forme une structure arborescente. Les extrémités de l'arbre sont des noeuds terminaux appelés feuilles, au-delà desquels il n'y a plus de questions [13].

Les deux algorithmes les plus populaires sont les suivants :

- Règles au niveau d'un "noeud" pour fractionner les données en fonction de leur valeur sur une variable
- Une règle d'arrêt pour décider quand un sous-arbre est complet.
- Affecter chaque noeud terminal (extrémité de l'arbre) ou feuille à un résultat de classe (prédiction)

Les arbres partitionnent récursivement les données, créant à chaque étape des groupes plus homogènes. Les règles résultantes sont les chemins qu'il faut emprunter pour aller du noeud racine à chaque noeud feuille.

Considérons l'arbre illustré à la figure 2.5, créé pour classer l'ensemble de données sur l'iris (un ensemble de données standard utilisé dans les discussions sur l'exploration de données). Deux variables seulement sont nécessaires pour classifier le type d'iris : la longueur et la largeur des pétales. Lorsque vous considérez deux des variables prédictrices, la longueur et la largeur des pétales, vous verrez comment l'algorithme des arbres de décision traite la première question au noeud 1.

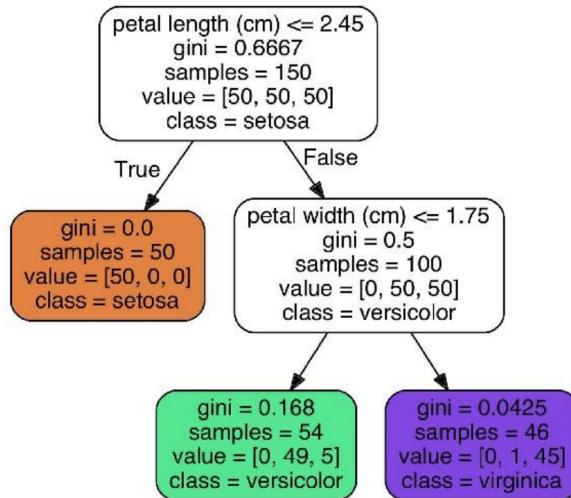


Figure 2.5 Arbre de décision du jeu de données Iris [13]

K-plus proches voisins

L'algorithme des k-plus proches voisins ou k-Nearest Neighbors (kNN) permet de réaliser des classifications binaires ou en classes multiples. Elle constitue sans doute l'une des approches de classification supervisées les plus simples. Les échantillons d'essai sont classés dans la classe la plus fréquente parmi les k voisins les plus proches dans l'espace multidimensionnel des paramètres. Par exemple, dans la figure 2.6, les deux valeurs du paramètre k attribuent au nouvel exemple deux classes différentes.

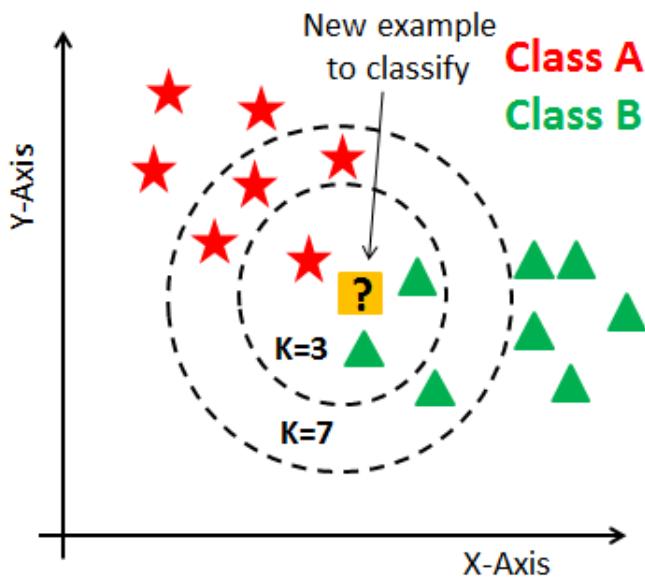


Figure 2.6 Classification avec la méthode des k-plus proches voisins [3]

La méthode ne nécessite que le choix de k, le nombre de voisins à prendre en compte pour la classification. Les petites valeurs de k sélectionneront les points d'entraînement les plus proches qui sont les mieux à même d'estimer la classification correcte au point de test.

Toutefois, en raison des faibles nombres, cette estimation sera sujette à d'importantes fluctuations statistiques. Inversement, des valeurs élevées de k réduisent les erreurs statistiques, mais permettent à des points éloignés de contribuer à la classification, ce qui peut lisser certains détails des distributions des classes. Habituellement, k est choisi comme la valeur qui minimise l'erreur de classification sur certaines données de validation indépendantes [12].

Réseaux de neurones

Les réseaux neuronaux artificiels sont des techniques d'apprentissage automatiques qui simulent le mécanisme de l'apprentissage dans systèmes nerveux biologiques. Le système nerveux humain contient des cellules, appelées neurones qui sont reliées les unes aux autres à l'aide d'axons et de dendrites (figure 2.7). Le mécanisme de fonctionnement de ce système est simulé dans des réseaux neuronaux artificiels (figure 2.8), qui contiennent des unités de calcul appelées neurones.

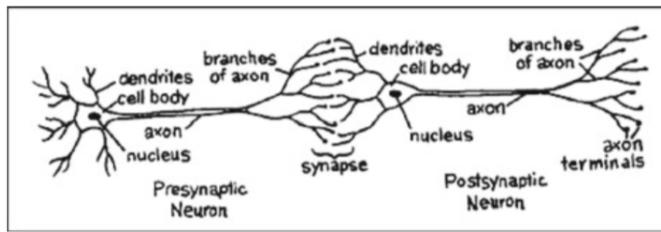


Figure 2.7 Réseau de neurone biologique [19]

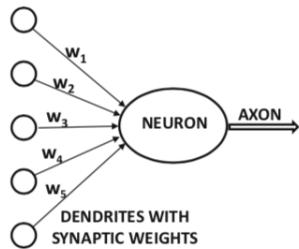


Figure 2.8 Réseau de neurone artificiel [19]

Les unités de calcul sont reliées les unes aux autres par des poids, qui jouent le même rôle que les forces des connexions synaptiques dans les organismes biologiques. Chaque entrée d'un neurone est mise à l'échelle avec un poids, ce qui affecte la fonction calculée à cette unité. Un réseau de neurones artificiel calcule une fonction des entrées en propageant les valeurs calculées des neurones d'entrée aux neurones de sortie et en utilisant les poids comme paramètres intermédiaires. L'apprentissage se fait en changeant les poids qui relient les neurones. [10]

Les réseaux de neurones sont composés d'une couche d'entrée (input layer), d'une ou de plusieurs couches intermédiaires (hidden layers) et d'une seule couche de sortie (output layer). Chaque couche peut avoir un nombre différent de neurones et chaque couche est entièrement connectée à la couche adjacente. Cette architecture est illustrée dans la figure 2.9.

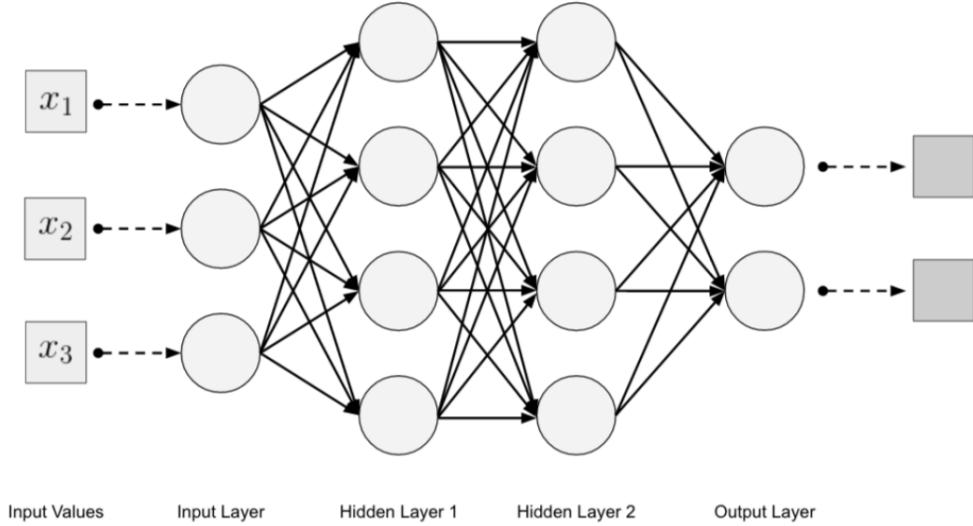


Figure 2.9 Réseau de neurones à plusieurs couches [19]

Le deep learning ou apprentissage profond pourrait être défini comme étant l'apprentissage en utilisant un réseau de neurones qui comprend un grand nombre de paramètres et de couches. Il existe quatre grandes architectures de que sont :

- Les réseaux de neurones à convolution
- Les réseaux de neurones récurrents
- Les réseaux de neurones récursifs
- Les réseaux pré-entraînés non supervisés [19]

Nous avons implémenté le **réseau de neurones à convolution** (CNN) défini dans l'article intitulé ECG Heartbeat Classification : A Deep Transferable Representation proposé par Kachuee et al. [15].

L'objectif d'un CNN est d'apprendre à partir d'un ensemble de données grâce des convolutions. Ils sont bien adaptés à la reconnaissance d'objets avec des images. Ils peuvent identifier des visages, des individus, des panneaux de signalisation, et bien d'autres aspects des données visuelles. Ils sont aussi utilisés pour l'analyse du son. Les couches convolutives sont considérées comme les éléments de base des architectures CNN. Ces couches transforment les données à travers l'opération de convolution. [19]

Une convolution est définie comme une opération mathématique décrivant une règle de fusion de deux ensembles d'informations. Il prend l'entrée, applique un noyau de convolution, et nous donne une carte de caractéristiques en sortie comme le montre la figure 2.10 ci-après :

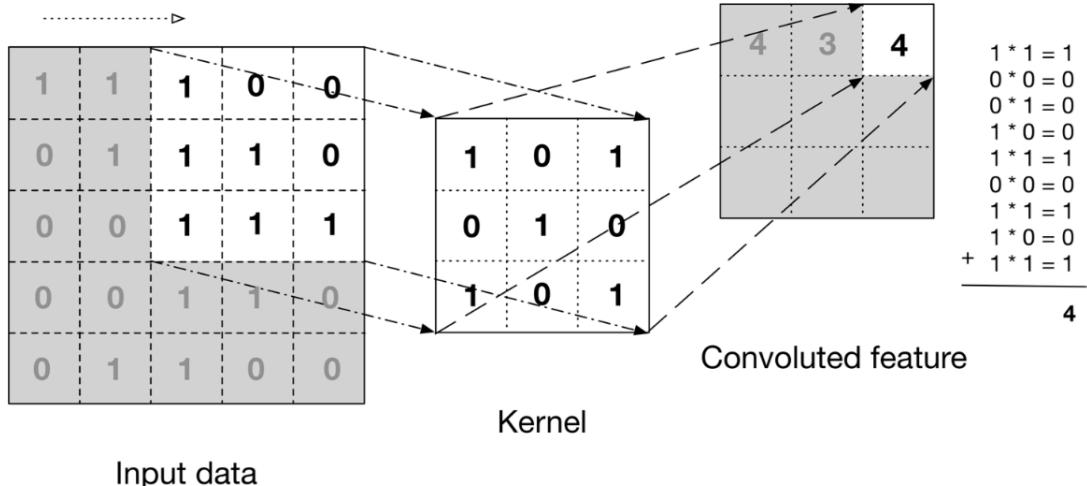


Figure 2.10 Réseau de neurones à plusieurs couches [19]

2.3.5 Évaluation de la performance des algorithmes

Pour évaluer un modèle de machine learning, certains indices de performances sont utilisés. C'est le cas par exemple de la précision et du rappel. Ce sont les valeurs de ces deux indicateurs que nous avons comparées pour chaque algorithme afin de choisir le modèle qui sera implémenté dans l'application d'aide au diagnostic.

Commençons par définir les notions de **vrai positif**, **vrai négatif**, **faux positif** et **faux négatif**. Ce sont les bases des critères que nous utiliserons pour évaluer les modèles de classification.

Un **vrai positif** est un résultat où le modèle prédit correctement la classe positive. De façon analogue, un **vrai négatif** est un résultat où le modèle prédit correctement la classe négative.

Un **faux positif** est un résultat où le modèle prédit incorrectement la classe positive. Et un **faux négatif** est un résultat où le modèle prédit incorrectement la classe négative.

La précision

La précision, aussi appelée valeur prédictive positive, correspond à la proportion d'identifications positives effectivement correcte. Autrement dit, le résultat de la précision est le pourcentage de réponses correctes pour la classe concernée.

$$\text{précision} = \frac{VP}{VP + FP}$$

Le rappel

Le rappel, ou la sensibilité du modèle, correspond à la proportion de résultats positifs réels correctement identifiés. En d'autres termes, il est défini par le nombre d'éléments positifs au regard du nombre d'éléments positifs que possède le jeu de données.

$$\text{rappel} = \frac{VP}{VP + FN}$$

Compromis entre précision et rappel

Pour évaluer les performances d'un modèle de façon complète, il faut analyser à la fois la précision et le rappel. Un modèle parfait devrait fournir des résultats égaux à 1 pour le calcul de la précision et du rappel (c'est-à-dire que l'algorithme trouve la totalité des documents pertinents et ne fait aucune erreur).

Malheureusement, précision et rappel sont fréquemment en tension. Ceci est dû au fait que l'amélioration de la précision se fait généralement au détriment du rappel et réciproquement.

CHAPITRE 3 ÉTUDE CONCEPTUELLE DE LA SOLUTION

Dans ce chapitre, nous aborderons les différentes étapes de conception de la solution proposée. La conception est la toute première étape du processus de développement. Nous commencerons par spécifier les besoins du projet avant d'expliquer les différents diagrammes réalisés à l'aide du Unified Modeling Language (UML).

3.1 Analyse des besoins

Notre solution doit pouvoir satisfaire les besoins fonctionnels qui sont les fonctionnalités mises à la disposition de l'utilisateur et les besoins non fonctionnels qui assurent la qualité logicielle du système.

3.1.1 Les besoins fonctionnels

La **classification des ECG** est le principal besoin fonctionnel de notre application. Le système doit pouvoir, en effet, classifier les ECG qu'on lui sont envoyés. C'est l'élément qui constitue l'outil d'aide à la décision pour les médecins.

En outre, notre application doit couvrir les besoins fonctionnels suivants :

- Consulter le dossier patient
- Modifier le dossier patient
- Ajouter une nouvelle consultation au dossier patient
- Confirmer ou modifier le diagnostic proposé par le système
- Accéder à l'interface de messagerie et de vidéoconférence

Nous verrons plus en détails ces fonctionnalités dans les diagrammes de cas d'utilisations proposés.

3.1.2 Les besoins non fonctionnels

Les besoins non fonctionnels peuvent être classifiés exigences qui ne concernent pas spécifiquement le comportement du système mais plutôt identifient des contraintes internes et externes du système. Ils ont pour but d'améliorer la qualité des services.

Parmi les besoins non fonctionnels, nous pouvons citer :

- l'**ergonomie** et la **convivialité** de l'application pour faciliter son utilisation au médecin
- la **sécurité de l'application** : seuls les personnes autorisés doivent pouvoir accéder aux informations confidentielles des patients
- la **maintenabilité** : Le code écrit doit respecter les bonnes pratiques de développement pour permettre des futures évolutions ou améliorations de la solution

3.2 Modélisation UML

La modélisation est la description de manière visuelle et graphique les besoins et, les solutions fonctionnelles et techniques d'un projet logiciel. Il existe plusieurs méthodes, dont UML, pour modéliser un logiciel.

UML, c'est l'acronyme anglais pour « Unified Modeling Language ». On le traduit par « Langage de modélisation unifié ». La notation UML est un langage visuel constitué d'un ensemble de schémas, appelés des diagrammes, qui donnent chacun une vision différente du projet à traiter. UML nous fournit donc des diagrammes pour représenter le logiciel à développer : son fonctionnement, sa mise en route, les actions susceptibles d'être effectuées par le logiciel, etc. UML offre donc un standard de modélisation, pour représenter l'architecture logicielle.

3.2.1 Diagramme des classes

Les diagrammes de classes sont l'un des types de diagrammes UML les plus utiles, car ils décrivent clairement la structure d'un système particulier en modélisant ses classes, ses attributs, ses opérations et les relations entre ses objets.

Une classe décrit les responsabilités, le comportement et le type d'un ensemble d'objets. Les éléments de cet ensemble sont les instances de la classe.

Nous proposons le diagramme des classes matérialisé à la figure 3.1 pour modéliser notre système.

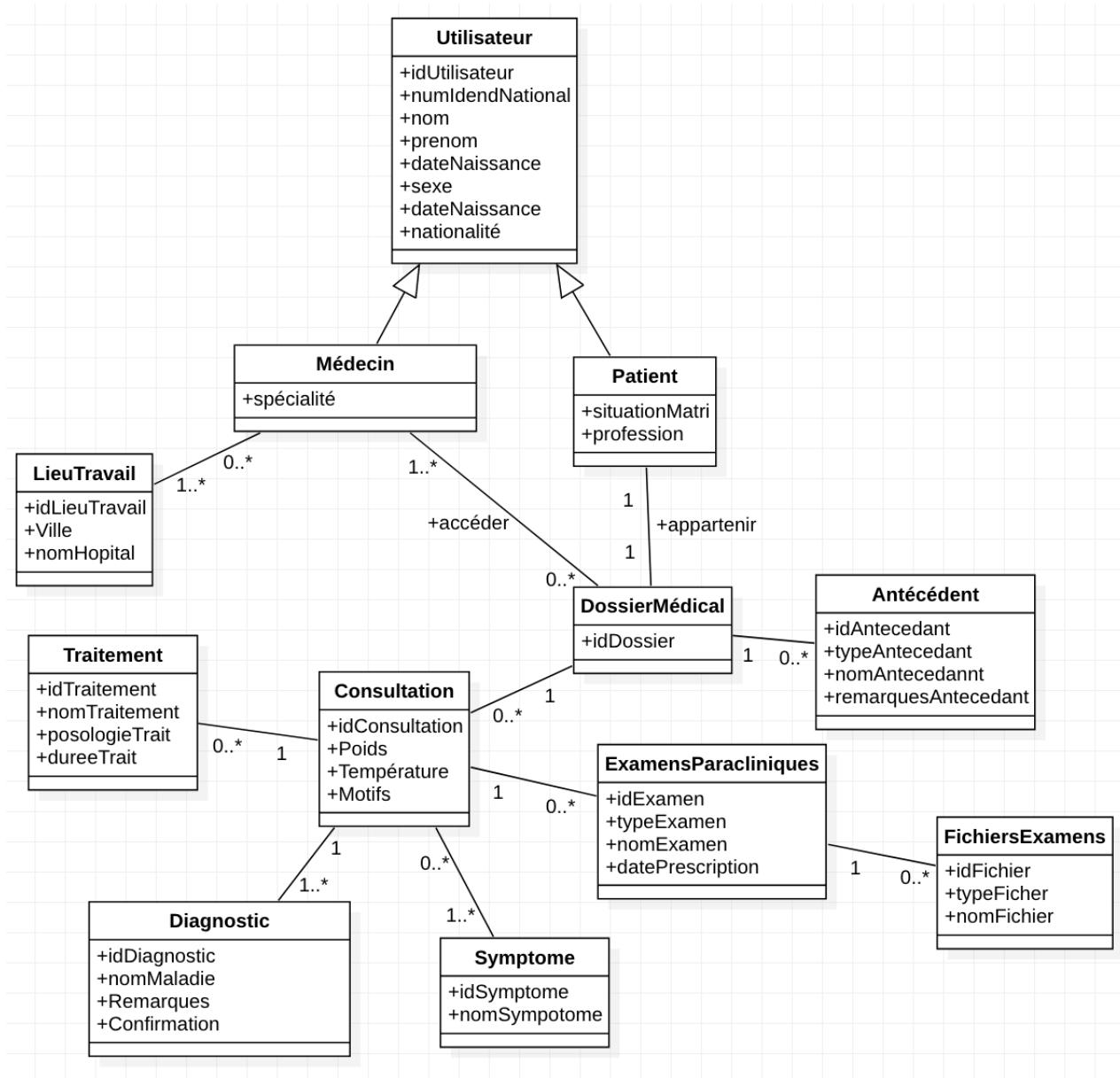


Figure 3.1 Diagramme des classes proposé

Le diagramme des classes ci-dessus comporte essentiellement les éléments constituant le Dossier Patient Informatisé (DPI). Le DPI est donc l'élément sur lequel repose notre application. Comme nous pouvons le constater, les éléments tels que les informations administratives du patient, la liste des antécédents, des examens paracliniques et des consultations sont représentés dans ce diagramme.

De plus, puisque l'intérêt de notre application est l'implémentation d'un outil d'aide à la

décision, il est primordial que les éléments sur lesquels reposent cette aide à la prise de décision apparaissent dans le schéma. Il s'agit, dans notre cas, des fichiers liés aux examens paracliniques.

Dans notre application, nous travaillerons exclusivement sur des ECG mais ces fichiers peuvent être de divers types : images radiographiques, résultats d'analyses etc.

Le diagramme de la figure 3.1 reflète le schéma de la base de données. Toutes les classes qui y figurent seront des tables dans notre base de données.

3.2.2 Diagrammes des cas d'utilisation

Les diagrammes des cas d'utilisation, ou diagramme des use cases sont utilisés pour donner une vision globale du comportement fonctionnel d'un système logiciel. Un cas d'utilisation représente une unité discrète d'interaction entre un utilisateur et un système. Dans un diagramme de cas d'utilisation, les utilisateurs du système, appelés acteurs, interagissent avec les cas d'utilisation(use cases).

La figure 3.2 illustre le diagramme des cas d'utilisation proposé dans le cadre de ce projet.

Le diagramme montre la présence de deux acteurs du système : le patient et le médecin. Le patient, qui est une généralisation du médecin, n'a à sa disposition qu'un nombre de fonctionnalités limité par rapport à son dossier.

Le médecin, quant-à lui, peut effectuer plusieurs actions, allant de la recherche de dossier, à la confirmation du diagnostic proposé. Nous décrirons plus en détails le cas d'utilisation qui répond à la problématique du sujet, à savoir la consultation du patient.

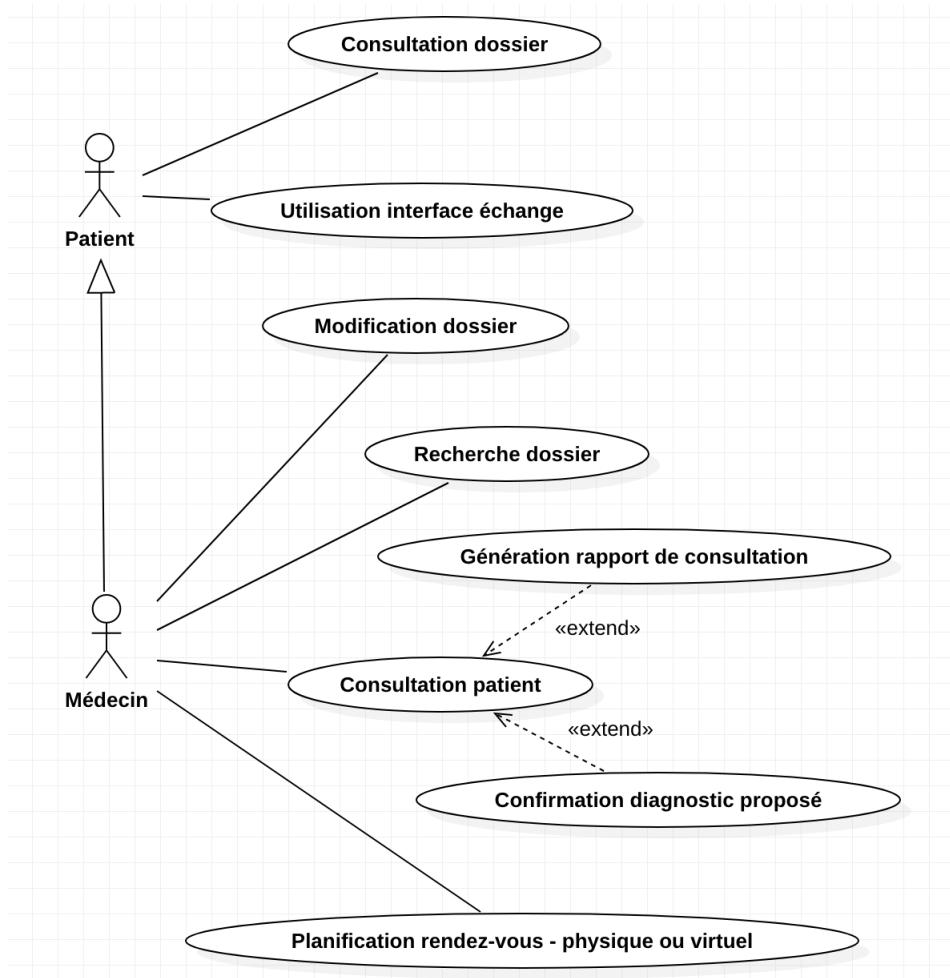


Figure 3.2 Diagramme des cas d'utilisation proposé

Description du cas d'utilisation "Consultation patient"

La consultation du patient se fait en plusieurs étapes et seul le médecin est habilité à modifier des informations concernant la consultation. La figure 3.3 fait office d'un schéma plus détaillé du cas d'utilisation : "Consultation patient".

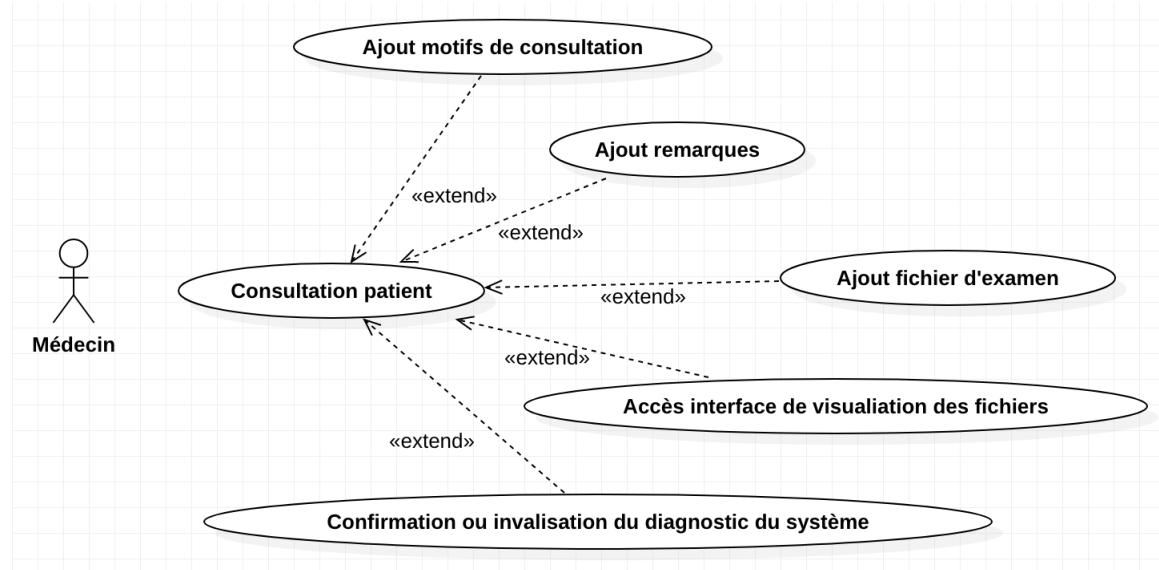


Figure 3.3 Cas d'utilisation "Consultation patient"

Le tableau 3.1 décrit de manière textuelle les étapes réalisées pour faire une consultation au niveau de notre application. À l'instar de tous les autres cas d'utilisation, la consultation requiert la l'authentification de l'utilisateur qui doit obligatoirement être un médecin.

Tableau 3.1 Fiche descriptive du cas d'utilisation "Consultation patient"

Cas d'utilisation	Consultation patient
Résumé	Procédure de consultation d'un patient
Acteurs	Médecin
Pré-conditions	Authentification à l'application d'aide au diagnostic
Post-conditions	Confirmation d'un diagnostic
Scénario normal	<ol style="list-style-type: none"> 1. Le médecin accède à la page de consultation à partir du dossier 2. Le système affiche la page de consultation 3. Le médecin choisit un formulaire à remplir 4. Le système affiche le formulaire demandé 5. Le médecin remplit et valide les informations entrées 6. Le médecin ajoute un fichier d'examen 7. Le système traite le fichier 8. Le système enregistre le diagnostic proposé après la classification 9. Le médecin accède à la page de visualisation du fichier 10. Le médecin confirme ou infirme le diagnostic 11. Le système enregistre le diagnostic

Ce tableau descriptif ne montre cependant pas une étape importante liée à l'outil d'aide au diagnostic. En effet, après l'enregistrement du diagnostic effectué par le médecin, le système

doit mettre à jour le jeu de données d'entraînement pour que le modèle s'améliore et fournisse des prédictions plus exactes dans le futur.

Les étapes manquantes de ce tableau seront matérialisés par le diagramme de séquences dans la partie suivante.

3.2.3 Diagramme de séquences

Le diagramme de séquences est la représentation graphique des interactions entre les acteurs et le système selon un ordre chronologique dans la formulation UML. Il permet de montrer les interactions d'objets dans le cadre d'un scénario d'un diagramme des cas d'utilisation. Dans un souci de simplification, on représente l'acteur principal à gauche du diagramme, et les acteurs secondaires éventuels à droite du système. Le but étant de décrire comment se déroulent les actions entre les acteurs ou objets.

Nous avons élaboré un diagramme des séquences pour le cas d'utilisation principal du système : la consultation d'un patient (voir figure 3.4). Le diagramme fait apparaître clairement les différentes phases dans lesquelles intervient le modèle de machine learning utilisé pour l'aide à la prise de décision.

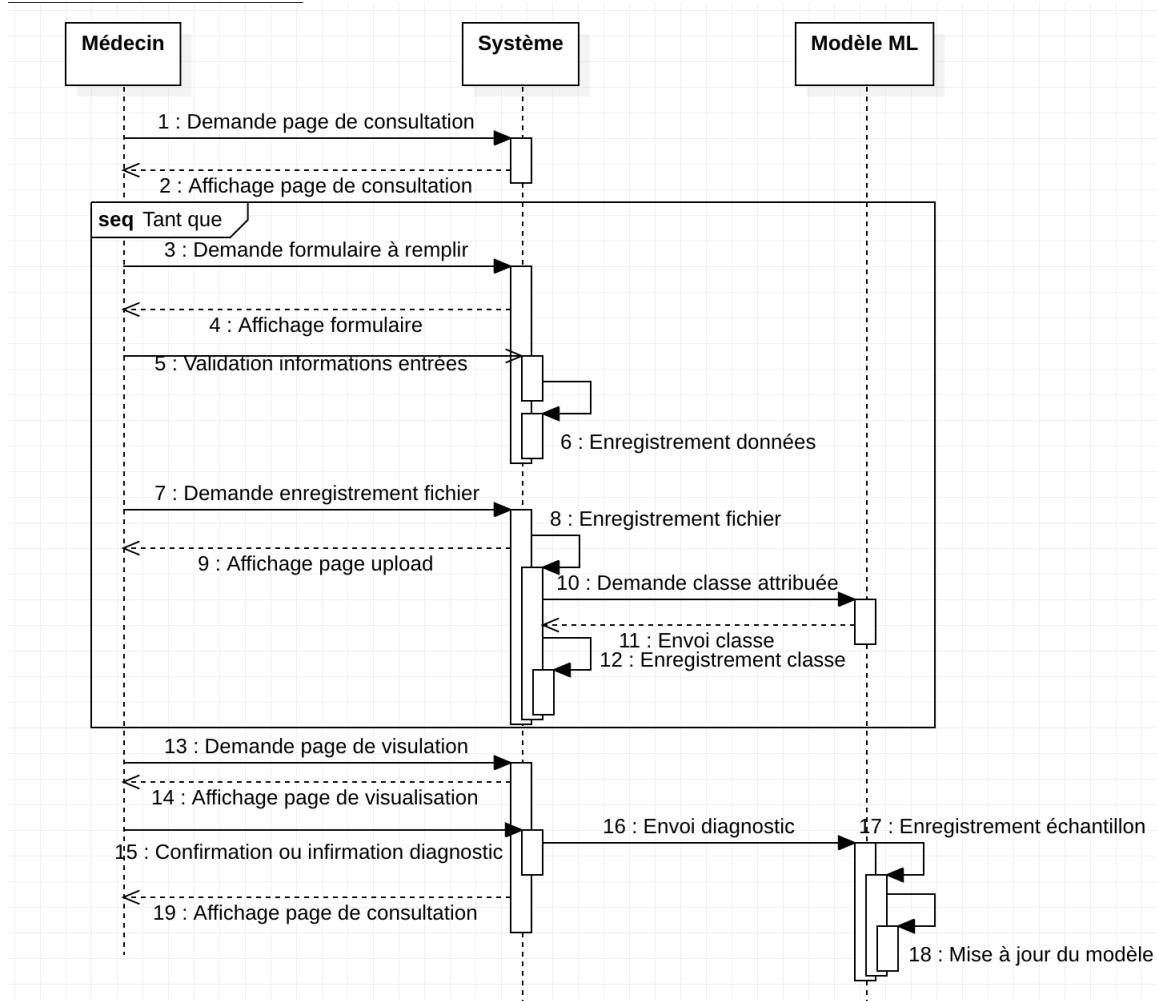


Figure 3.4 Diagramme de séquences proposé

3.3 Workflow des données utilisées pour la classification

Les données sur lesquelles le travail de classification a été effectué proviennent de la base de données du MIT-BIH sur l'arythmie. L'arythmie est une anomalie du système cardiaque qui affecte la fréquence des battements du cœur. Les différents types d'arythmie entraînent soit une fréquence de battement du cœur trop élevée, trop lente ou irrégulière. Ils sont généralement détectés et identifiés à travers une électrocardiographie. [7] Les données que contiennent cette base proviennent du Beth Israel Hospital (BIH) de Boston et du Massachusetts Institute of Technology (MIT). La base de données du MIT-BIH sur l'arythmie est composée de 48 extraits d'une trentaine de minutes chacun d'enregistrements d'ECG à deux canaux, obtenus de 47 sujets étudiés par le BIH entre 1975 et 1979. [8]

Les données ont déjà subi un travail de prétraitement et de segmentation conformément aux

étapes proposées par Kachuee et al. [15] dans leur article. Le jeu de données est composé de deux fichiers CSV : un jeu d'entraînement et un jeu de test. Chacun de ces fichiers contient une matrice dans laquelle chaque ligne représentant un exemple (un ECG) de l'ensemble de données. Le dernier élément de chaque ligne indique la classe à laquelle appartient cet exemple.

Chaque électrocardiogramme, en fonction du type de battement de cœur, appartient à une des cinq classes (voir Tableau 3.2) définies dans la norme EC57 par l'Association pour l'Avancement de l'Instrumentation Médicale (AAMI) [20].

Catégorie	Type de battement
N	<ul style="list-style-type: none"> • Normal • Left/Right bundle branch block • Atrial escape • Nodal escape
S	<ul style="list-style-type: none"> • Atrial premature • Aberrant atrial premature • Nodal premature • Supra-ventricular premature
V	<ul style="list-style-type: none"> • Premature ventricular contraction • Ventricular escape
F	<ul style="list-style-type: none"> • Fusion of ventricular and normal
Q	<ul style="list-style-type: none"> • Paced • Fusion of paced and normal • Unclassifiable

Tableau 3.2 Classes de battements de coeurs

La figure 3.5 illustre le processus par lequel les données utilisées passe pour rendre performant le modèle d'apprentissage automatique. Ce modèle permettra d'attribuer une des cinq classes

à chaque nouvel électrocardiogramme.

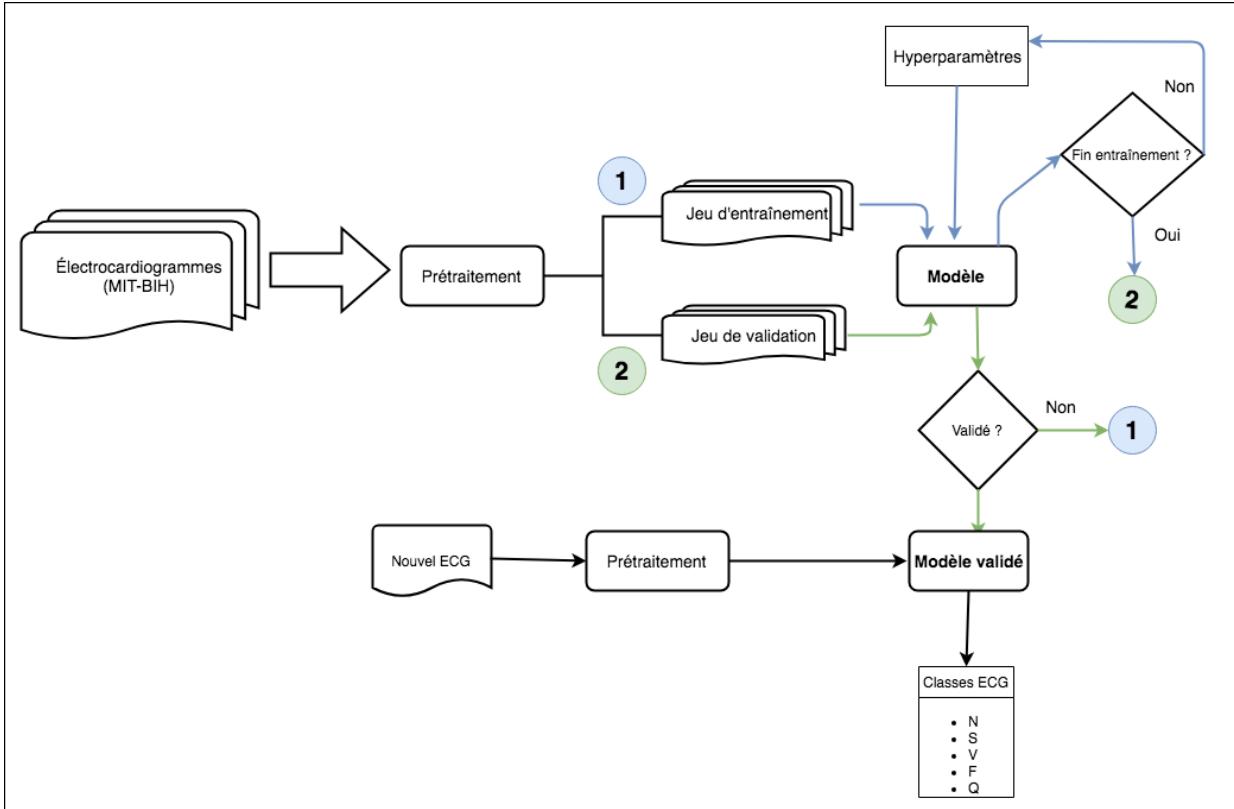


Figure 3.5 Workflow des données de classification

Comme le montre la figure, les enregistrements d'ECG [8] subissent d'abord une étape de prétraitement, visant à isoler chaque battement, conformément au processus décrit dans l'étude réalisée par Kachuee et al. [15] :

1. Diviser le signal ECG continu en fenêtres 10s et sélectionner une fenêtre de 10s à partir d'un signal ECG.
2. Normaliser les valeurs d'amplitude dans la plage comprise entre zéro et un.
3. Trouver l'ensemble de tous les maximums locaux basés sur les passages à zéro de la première dérivée.
4. Trouver l'ensemble des pics R des ECG
5. Trouver la médiane des intervalles de temps R-R en tant que valeur nominale de cette fenêtre (T).
6. Pour chaque pic R, sélectionner une partie de signal avec la longueur égale à 1,2T.

7. Rembourrer chaque pièce sélectionnée avec des zéros pour faire sa longueur égale à une longueur fixe prédéfinie.

CHAPITRE 4 IMPLEMENTATION DE L'APPLICATION

Dans ce chapitre, nous aborderons les différentes étapes de l'implémentation de la solution proposée. L'implémentation désigne la phase de réalisation du logiciel. Nous parlerons dans cette partie des outils et technologies qui nous ont permis de mettre en oeuvre l'application d'aide à la prise de décision pour les maladies cardiovasculaires. Il s'agira dans un premier temps de faire une analyse comparative des modèles d'apprentissage automatique avant de commenter l'architecture logicielle choisie. Nous terminerons par la présentation de quelques interfaces de l'outil proposé.

4.1 Modèle d'apprentissage automatique

Nous avons implémenté les trois premiers modèles d'apprentissage automatique présentés dans le chapitre 3 en utilisant les algorithmes de la bibliothèque Python scikit-learn. En ce qui concerne le réseau de neurones, nous nous sommes basés sur le travail de Kachuee et al. [15]. Pour ce faire, nous avons fait usage des bibliothèques pour le traitement des données :

- **scikit-learn**

Scikit-learn est une bibliothèque destinée à l'apprentissage automatique [9]. Elle fournit des outils simples et efficaces pour l'exploration et l'analyse de données.

- **NumPy**

numPy est paquet fondamental pour le calcul scientifique avec Python. NumPy est destinée à manipuler des matrices ou tableaux multidimensionnels ainsi que des fonctions mathématiques opérant sur ces tableaux.

- **Pandas**

Pandas est une bibliothèque open source sous licence BSD offrant des structures de données et des outils d'analyse de données performants et faciles à utiliser pour le langage de programmation Python.

- **Keras**

Keras est une bibliothèque open source écrite en python et permettant d'interagir avec les algorithmes de réseaux de neurones profonds et de machine learning, notamment Tensorflow.

Analysons maintenant les résultats obtenus pour chaque algorithme.

Machines à vecteur support

L'algorithme utilisé est un algorithme SVM non linéaire à noyau Gaussien radial (Radial Basis Function ou RBF en anglais), implémenté par la bibliothèque scikit-learn.

Les résultats de la classification des données contenues dans le jeu de test effectuée avec les machines à vecteurs de support ne sont pas satisfaisants. Les valeurs moyennes (voir figure 4.1) de la précision et du rappel étant respectivement de 0.56 et 0.48 sont trop faibles pour que ce modèle soit utilisé pour effectuer des classifications fiables sur des ECG.

	precision	recall
0.0	0.92	1.00
1.0	0.00	0.00
2.0	0.89	0.52
3.0	0.00	0.00
4.0	0.99	0.87
micro avg	0.92	0.92
macro avg	0.56	0.48

Figure 4.1 Précision et rappel de la classification avec SVM

Nous remarquons que les valeurs de ces indices de performance sont aussi faibles car l'algorithme n'a pu identifier aucun élément parmi ceux des classes **S** et **F**.

Arbres de décision

La classification des ECG avec les arbres de décision a produit des résultats plus satisfaisants que ceux obtenus avec les SVM. Contrairement aux SVM, les arbres de décision ont pu identifier des éléments dans chaque classe. Les indices de performances étudiés (précision et rappel) ont, quant à eux, des valeurs moyennes plus élevées comme le montre la figure 4.2.

	precision	recall
0.0	0.98	0.97
1.0	0.63	0.64
2.0	0.86	0.86
3.0	0.53	0.62
4.0	0.95	0.94
micro avg	0.95	0.95
macro avg	0.79	0.81

Figure 4.2 Précision et rappel avec les arbres de décision [13]

L'algorithme basé sur les arbres de décision classe donc mieux les électrocardiogrammes et commet moins d'erreurs que celui basé sur les SVM.

K-plus proches voisins

Les valeurs de la précision et du rappel calculées après avoir effectué la classification avec la méthode des k-plus proches voisins sont plus élevées que celles obtenues avec les deux algorithmes précédents. Par ailleurs, l'exécution de l'algorithme a été la plus rapide parmi toutes les méthodes que nous avons expérimentées. L'algorithme de classification n'a nécessité que quelques secondes pour être exécuté et a bien classé environ 91% des électrodiagrammes du jeu de test (voir figure 4.3).

	precision	recall
0.0	0.98	0.99
1.0	0.90	0.64
2.0	0.94	0.90
3.0	0.76	0.64
4.0	0.99	0.95
micro avg	0.97	0.97
macro avg	0.91	0.82

Figure 4.3 Précision et rappel avec kNN

Réseau de neurones

Le réseau de neurones proposé par Kachuee et al. [15] est un réseau de neurones à convolution qui comporte 13 couches cachées.

La classification des ECG avec le réseau de neurones proposé par Kachuee et al. [15] a fourni de meilleurs résultats que toutes les autres méthodes d'apprentissage automatique que nous avons essayées. La quasi-totalité des ECG du jeu de test a été classée correctement comme le montre la matrice de confusion obtenue :

```
[ [18095    7     9     6     1]
  [   38    506    10     2     0]
  [   4      0  1435     9     0]
  [   4      0      5  153     0]
  [   2      0      1     0 1605]]
```

Figure 4.4 Matrice de confusion de la classification avec le CNN

Les valeurs moyennes du rappel et de la précision sont, par conséquent, toutes les deux très élevées (voir figure 4.5). Le réseau de neurones proposé est donc plus performant, pour ce jeu de données, que les méthodes d'apprentissage utilisées précédemment. C'est la raison pour laquelle il a été retenu comme modèle pour la classification des ECG.

	precision	recall
0	1.00	1.00
1	0.99	0.91
2	0.98	0.99
3	0.90	0.94
4	1.00	1.00
micro avg	1.00	1.00
macro avg	0.97	0.97

Figure 4.5 Précision et rappel avec le réseau de neurones

4.2 Architecture logicielle

Une architecture logicielle est une représentation abstraite d'un système exprimée essentiellement à l'aide de composants logiciels en interaction. Elle met en exergue les outils et technologies utilisés pour mettre en oeuvre un système logiciel. Nous présenterons brièvement chaque outil présent dans notre architecture logicielle illustrée dans la figure 4.6. La figure montre l'acteur (le médecin) qui se servira du logiciel ainsi que les deux parties faisant fonctionner l'application : le **Front-End** et le **Back-End**.

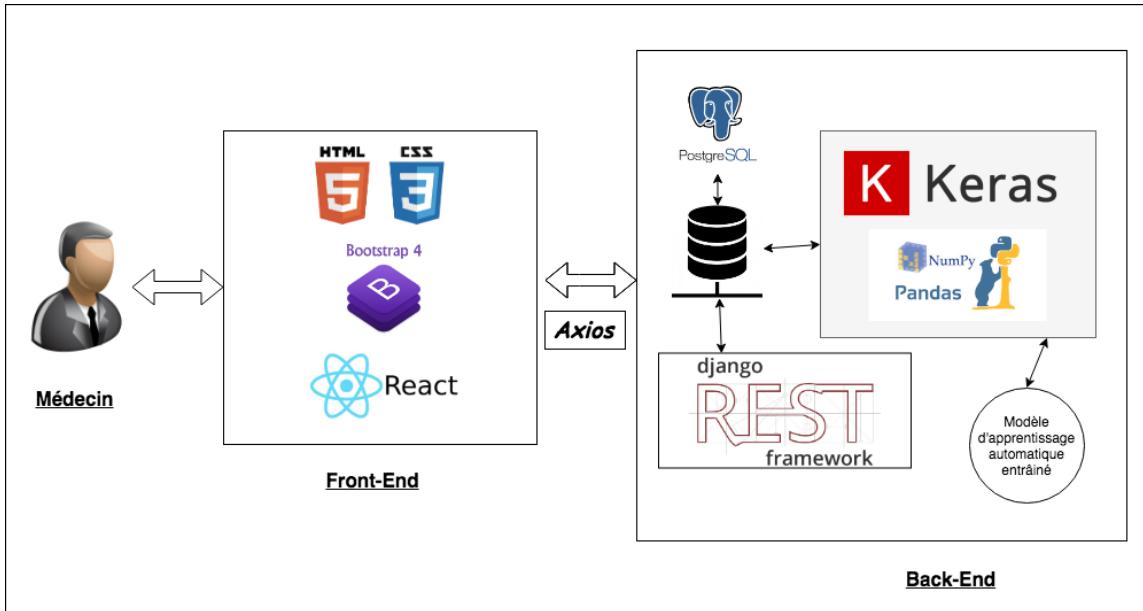


Figure 4.6 Architecture logicielle de l'application

4.2.1 Front-End

Le Front-End correspond à la partie visible de l'application. En d'autres termes, il correspond aux productions HTML, CSS et JavaScript d'une page internet ou d'une application qu'un utilisateur peut voir et avec lesquelles il peut interagir directement.

Nous avons utilisé principalement quatre outils pour faire le développement front-end de notre application (voir figure 4.6) :

- HyperText Markup Language (HTML)
- Cascading Style Sheets (CSS)
- Bootstrap
- React

HyperText Markup Language (HTML)

Le HyperText Markup Language, généralement abrégé HTML décrit le contenu et définit la structure dans une page Web. HTML permet également de structurer sémantiquement et logiquement et de mettre en forme le contenu des pages, d'inclure des ressources multimédias dont des images, des formulaires de saisie et des programmes informatiques.

Il est souvent utilisé conjointement avec le langage de programmation JavaScript et des feuilles de style en cascade (CSS).

Cascading Style Sheets (CSS)

Les feuilles de style en cascade, généralement appelées CSS de l'anglais Cascading Style Sheets, forment un langage informatique qui décrit la présentation des documents HTML et XML. Les standards définissant CSS sont publiés par le World Wide Web Consortium (W3C). Introduit au milieu des années 1990, CSS devient couramment utilisé dans la conception de sites web et bien pris en charge par les navigateurs web dans les années 2000.



Figure 4.7 Logo du HTML et du CSS

Le CSS permet d'avoir un rendu plus élégant en organisant mieux la mise en forme des pages Web et en offrant des fonctionnalités permettant le responsive design.

React

React, aussi appelé React.js ou ReactJS, est une bibliothèque JavaScript pour créer des interfaces utilisateurs. C'est une bibliothèque libre développée par Facebook depuis 2013. Son but principalest de faciliter la création d'application web monopage, via la création de composants dépendant d'un état et générant une page (ou portion) HTML à chaque changement d'état.

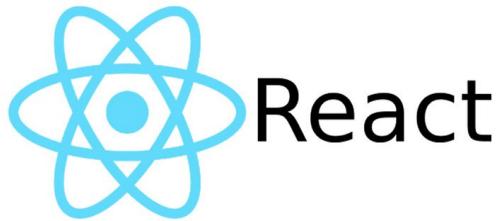


Figure 4.8 Logo de la bibliothèque React

React est une bibliothèque qui ne gère que l'interface de l'application, considéré comme la vue dans le modèle MVC. Elle peut ainsi être utilisée avec une autre bibliothèque ou un framework MVC comme AngularJS. La bibliothèque se démarque de ses concurrents par sa flexibilité et ses performances, en travaillant avec un DOM virtuel et en ne mettant à jour le rendu dans le navigateur qu'en cas de nécessité.

React est une librairie basée sur les composants. Elle permet de créer des composants autonomes qui maintiennent leur propre état, puis de les assembler pour créer des interfaces utilisateurs complexes.

React offre également de nombreux avantages aux développeurs parmi lesquels nous pouvons citer :

- La facilité d'apprentissage
- La rapidité grâce au DOM virtuel
- La maintenabilité du code qui est bien structuré en composants
- La possibilité de créer des applications mobiles avec React Native
- La grande communauté utilisant la librairie

Pour communiquer avec les éléments du back-end, nous avons utilisé Axios. Axios est un client HTTP léger caractérisé par sa facilité d'utilisation. Il permet de faire des requêtes de type GET, POST, DELETE, de l'application React à une base de données. C'est donc la passerelle entre notre front-end et notre back-end comme le montre la figure 4.6.

4.2.2 Back-End

Le back-end est un terme désignant la partie d'un logiciel devant produire un ou plusieurs résultats. C'est dans cette partie que sont effectués les calculs, les requêtes vers les bases de données, les traitements de données. Les résultats produits après traitements du back-end sont envoyés à la partie front-end qui se chargera de les afficher.

Nous avons utilisé plusieurs outils suivants pour réaliser le back-end de notre application (voir figure 4.6) :

- Django REST framework
- un Système de Gestion de Base de Données (SGBD) PostgreSQL
- des bibliothèques pour l'analyse de données : Keras, numPy et Pandas

Django REST framework

Django REST Framework est une boîte à outils puissante et flexible qui vous facilite la création d'application web API . Il est basé sur le langage de programmation Python caractérisé par son typage dynamique fort, sa gestion automatique de la mémoire et son système de gestion d'exceptions.

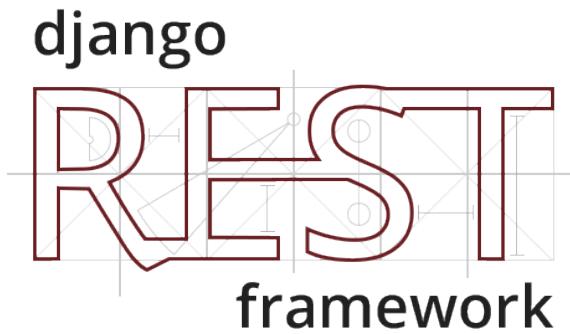


Figure 4.9 Logo de Django REST framework

Une API ou interface de programmation d'application est un ensemble normalisé de classes, de méthodes ou de fonctions qui sert de façade par laquelle un logiciel offre des services à d'autres logiciels.

Nous avons donc opté pour le choix d'un framework créant des API pour prévoir les futures évolutions de l'application. En effet, si une application mobile devait être développée pour ce même projet, seule la partie front-end aurait été implémentée à nouveau, les fonctions

et méthodes étant déjà disponibles en back-end. Il suffirait donc d'intégrer les requêtes au nouveau code front-end.

PostgreSQL

PostgreSQL est un système de gestion de base de données relationnelle et objet.

PostgreSQL a la particularité de stocker plus de types de données que les types simples traditionnels entiers, caractères, etc. L'utilisateur peut créer des types, des fonctions, utiliser l'héritage de type, etc. PostgreSQL est aussi reconnu pour son comportement stable, mais aussi pour ses possibilités de programmation étendues, directement dans le moteur de la base de données.

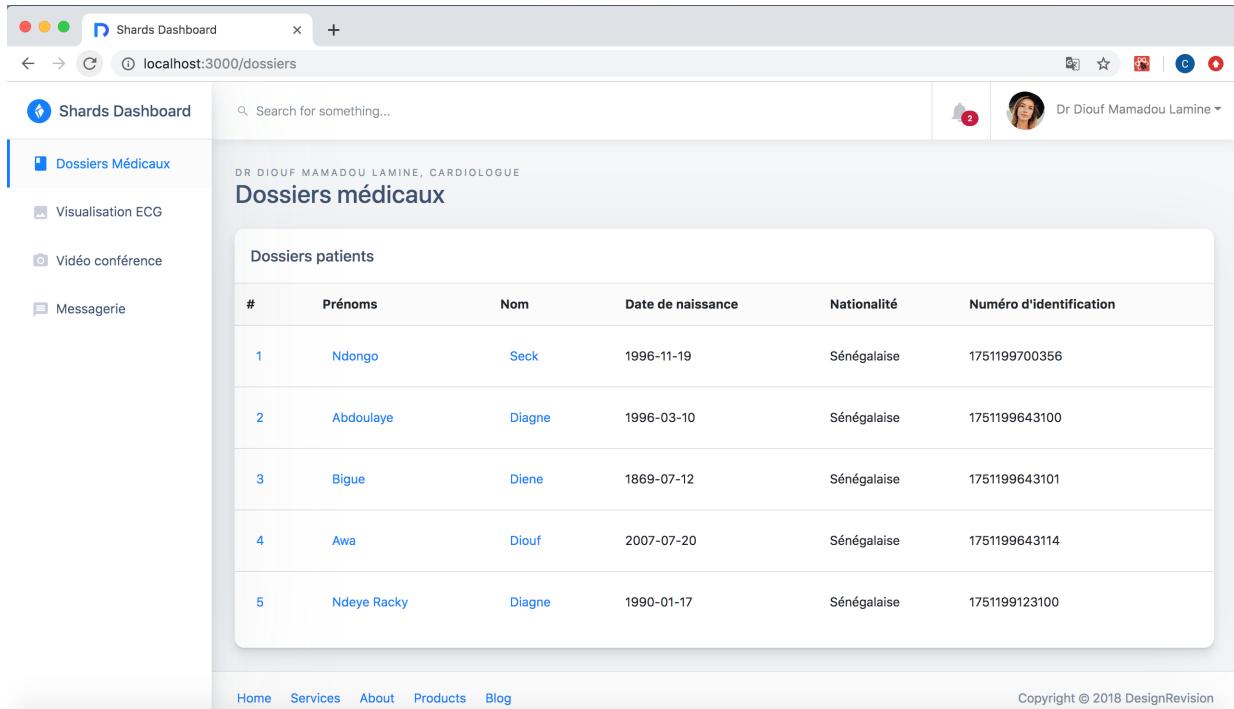
Son interface d'administration, pgAdmin facilite considérablement le travail du développeur.

4.3 Présentation de l'application

Dans cette partie, nous présenterons quelques interfaces de l'application réalisée.

4.3.1 Interface de la liste des dossiers

La figure 4.10 illustre l'interface de l'application qui présente au médecin la liste des dossiers patients auxquels il a accès.



The screenshot shows a web browser window titled "Shards Dashboard" at the URL "localhost:3000/dossiers". The page has a header with a search bar and a user profile for "Dr Diouf Mamadou Lamine". On the left, there is a sidebar menu with options: "Dossiers Médicaux" (which is selected and highlighted in blue), "Visualisation ECG", "Vidéo conférence", and "Messagerie". The main content area is titled "Dossiers médicaux" and "Dossiers patients". It displays a table with the following data:

#	Prénoms	Nom	Date de naissance	Nationalité	Numéro d'identification
1	Ndongo	Seck	1996-11-19	Sénégalaise	1751199700356
2	Abdoulaye	Diagne	1996-03-10	Sénégalaise	1751199643100
3	Bigue	Diene	1869-07-12	Sénégalaise	1751199643101
4	Awa	Diouf	2007-07-20	Sénégalaise	1751199643114
5	Ndeye Racky	Diagne	1990-01-17	Sénégalaise	1751199123100

At the bottom of the page, there is a footer with links to "Home", "Services", "About", "Products", and "Blog", and a copyright notice "Copyright © 2018 DesignRevision".

Figure 4.10 Interface de la liste des dossiers

Sur le côté gauche de la figure se trouve le menu qui permet au médecin de naviguer aisément dans les différentes sections de l'application. À droite de ce menu se situe la liste des dossiers présentée sous forme de tableau contenant des informations qui permettent d'identifier rapidement un patient. Il s'agit des prénoms, du nom, de la date de naissance, de la nationalité ainsi que du numéro d'identification nationale.

4.3.2 Interface de visualisation d'un dossier patient

La figure 4.11 illustre l'interface de l'application qui présente au médecin les informations contenues d'un dossier patient.

The screenshot shows a web browser window titled "Shards Dashboard" at "localhost:3000/dossier-patient". The left sidebar has links for "Dossiers Médicaux", "Visualisation ECG", "Vidéo conférence", and "Messagerie". The main content area displays a "Dossier médical de M. Ndongo Seck" for "DR DIOUF MAMADOU LAMINE, CARDIOLOGUE". It is divided into two main sections: "Fiche Patient" and "Antécédents du Patient".

Fiche Patient:

- Nom & Prénoms: Seck Ndongo
- Date de naissance: 1996-11-19
- Nationalité: Sénégalaise
- Situation matrimoniale: Célibataire
- Profession: Étudiant

Antécédents du Patient:

- Type : Allergies
- Intitulé: Allergies polliniques
- Remarques: Crises d'éternuement, rhinite, respiration nasale difficile, yeux rougis / Prise régulière d'un collyre anti-allergique
- Intitulé: Intolérance au gluten
- Remarques: Diarrhées fréquentes après l'ingestion de certaines céréales. Aucun traitement suivi.

Figure 4.11 Interface de visualisation d'un dossier patient

L'interface de visualisation comporte des éléments du dossier patient définis dans le diagramme des classes. Cette figure montre la fiche patient qui renferme les informations administratives du patient. De plus, le médecin a la possibilité de visualiser la liste des antécédents qui sont de trois types : chirurgicaux, familiaux ou les allergies.

En outre, cette page permet au médecin d'avoir un aperçu sur l'historique des consultations du patient (voir figure 4.12). En plus de pouvoir voir la liste des consultations passées, il peut accéder aux pages qui présentent les consultations en détail, mais aussi ajouter une nouvelle consultation.

The screenshot shows a web-based dashboard titled "Shards Dashboard" at the URL "localhost:3000/dossier-patient". On the left, there is a sidebar with links: "Dossiers Médicaux", "Visualisation ECG", "Vidéo conférence", and "Messagerie". The main content area has a search bar at the top. Below it, a card displays the patient's information: "Profession" (Étudiant), "Intitulé" (Intolérance au gluten), and "Remarques" (Diarrhées fréquentes après l'ingestion de certaines céréales. Aucun traitement suivi.). A large section titled "Historique des consultations" lists three consultations:

#	Date de consultation	Motifs	Hypothèses diagnostiques	Confirmation
3	2019-06-18	Souffle court, suées importantes, nausée	Arrythmie cardiaque	Non
2	2019-04-23	Essoufflement, nausées, vomissements	Essai hypothèse	Oui
1	2019-02-21	Maux de tête, vertiges, fatigue	Paludisme, Malaria	Oui

At the bottom of the page, there are links for "Home", "Services", "About", "Products", and "Blog", along with a copyright notice: "Copyright © 2018 DesignRevision".

Figure 4.12 Interface de visualisation d'un dossier patient

4.3.3 Interface de visualisation des détails d'une consultation

La figure 4.13 illustre l'interface de l'application qui présente au médecin les informations relatives à une consultation particulière.

En plus de la fiche patient, l'interface fait apparaître les résultats de l'examen médical effectué. Les informations qui y sont consignées sont : les motifs de la visite du patient, les remarques faites par le médecin, les examens paracliniques prescrits ainsi que le diagnostic confirmé du médecin.

La figure 4.14 montre les fichiers liés aux examens paracliniques. L'application offre la possibilité au médecin d'en ajouter ou de visualiser le contenu du fichier. Dans notre cas, il s'agit d'électrocardiogrammes.

The screenshot shows a web application interface titled "Shards Dashboard". The main content area displays the "Détails de consultation de M. Ndongo Seck". On the left, there is a sidebar with links to "Dossiers Médicaux", "Visualisation ECG", "Vidéo conférence", and "Messagerie". The main content is divided into two sections: "Fiche de consultation" and "Examen médical".

Fiche de consultation:

- Nom & Prénoms du patient: Seck Ndongo
- Date de consultation: 2019-06-18
- Médecin: Dr Diouf Mamadou Lamine
- Poids: 71

Examen médical:

- Motifs:** Souffle court, suées importantes, nausée
- Remarques médecin:** Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed non risus. Suspendisse lectus tortor, dignissim sit amet, adipiscing nec, ultricies sed, dolor. Cras elementum ultrices diam. Maecenas ligula massa, varius a, semper congue, euismod non, mi. Proin porttitor, orci nec nonummy molestie, enim est eleifend mi, non fermentum diam nisl sit amet erat. Duis semper. Duis arcu massa, scelerisque vitae, consequat in, pretium a, enim. Pellentesque congue.
- Hypothèses diagnostiques:**
 - Arrythmie cardiaque 2
 - Arrythmie cardiaque
- Examens paracliniques:**
 - Electrocardiographie
- Diagnostic confirmé:**
 - Atrial premature

Figure 4.13 Interface de visualisation des détails d'une consultation

This screenshot shows the same dashboard interface as Figure 4.13, but with a different section visible. The "Examens paracliniques et fichiers liés" section is now the active one, displaying a table of linked files.

#	Nom de l'examen	Date de prescription	Type fichier	Nom fichier
1	Electrocardiographie	2019-06-18	CSV	patient.csv

At the bottom of the page, there is a navigation bar with links to "Home", "Services", "About", "Products", and "Blog". The footer contains the copyright information: "Copyright © 2018 DesignRevision".

Figure 4.14 Interface de visualisation des détails d'une consultation

4.3.4 Interface de visualisation des électrocardiogrammes

La figure 4.15 illustre l'interface de l'application qui présente au médecin le résultat de l'électrocardiographie prescrite.

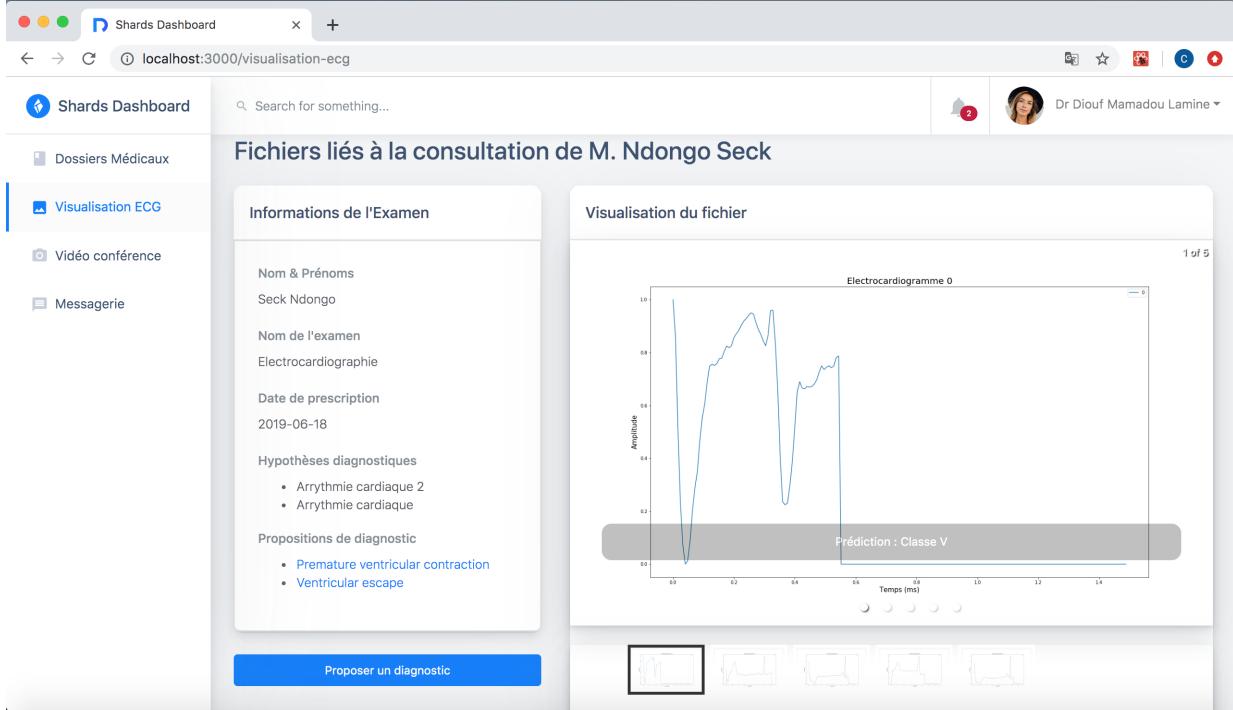


Figure 4.15 Interface de visualisation d'un ECG

L'application affiche également le diagnostic effectué par notre modèle d'apprentissage automatique. Le médecin a la possibilité de le confirmer ou de l'invalider en proposant lui-même son diagnostic.

4.3.5 Interfaces de messagerie et de visioconférence

Parmi les fonctionnalités implémentées se trouvent une espace de messagerie, ainsi que l'intégration d'un service de visioconférence.

L'interface de messagerie (voir figure 4.16) permet aux médecins de communiquer entre eux. Les médecins ont la possibilité de lire, de répondre et d'archiver leurs messages.

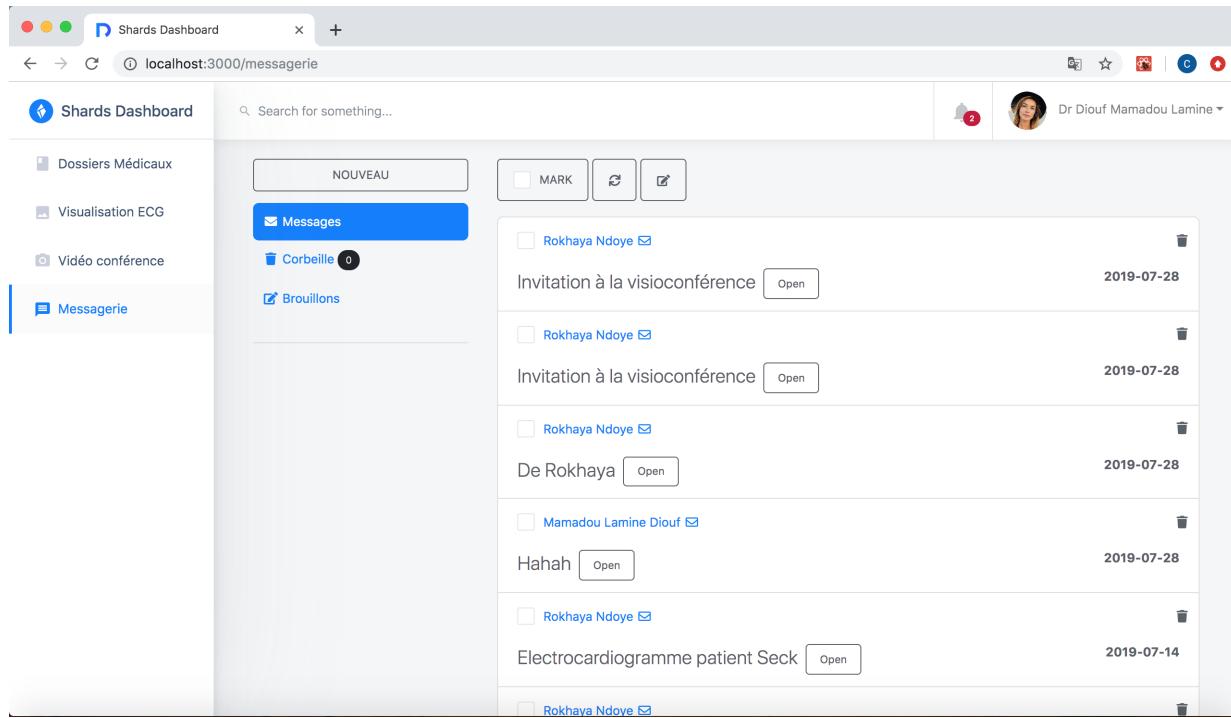


Figure 4.16 Interface de messagerie

En ce qui concerne l'interface de visioconférence, le médecin choisit d'abord celui avec qui il veut entrer en communication comme le montre la figure 4.17. Un message est ensuite envoyé dans l'espace de messagerie du correspondant. qui peut rejoindre l'appel à partir du lien qui lui sera envoyé (voir figure 4.17).

Dossiers médicaux

#	Prénoms	Nom	Spécialité	Hôpital	Ville
2	Rokhaya	Ndoye	Généraliste	Dakar	Hôpital Principal
3	Mamadou	Ba	Pédiatre	Dakar	Clinique Pasteur
4	Fadel	Ndiaye	Généraliste	Dakar	Clinique du Cap
5	Mohamed	Diop	Cardiologue	Thies	Hôpital Régional
6	Ousmane	Diouf	Généraliste	Dakar	Clinique de la Madeleine

Home Services About Products Blog Copyright © 2018 DesignRevision

Figure 4.17 Choix du correspondant

OBJET
Invitation à la visioconférence

DE
Mamadou Lamine Diouf

ENVOYÉ
2019-07-29

Bonjour,

Je vous invite à rejoindre la visioconférence en vous rendant à l'adresse suivante :

<https://localhost:3005/Diouf-Mamadou-Lamine>

Merci.

Replay Close

Mamadou Lamine Diouf [✉](#)
lien [Open](#)
 Mamadou Lamine Diouf [✉](#)

Navigation privée (6)

Figure 4.18 Message d'invitation

Nous avons utilisé l'outil SimpleWebRTC pour intégrer la visioconférence à notre application. La figure 4.19 illustre l'interface de visioconférence :

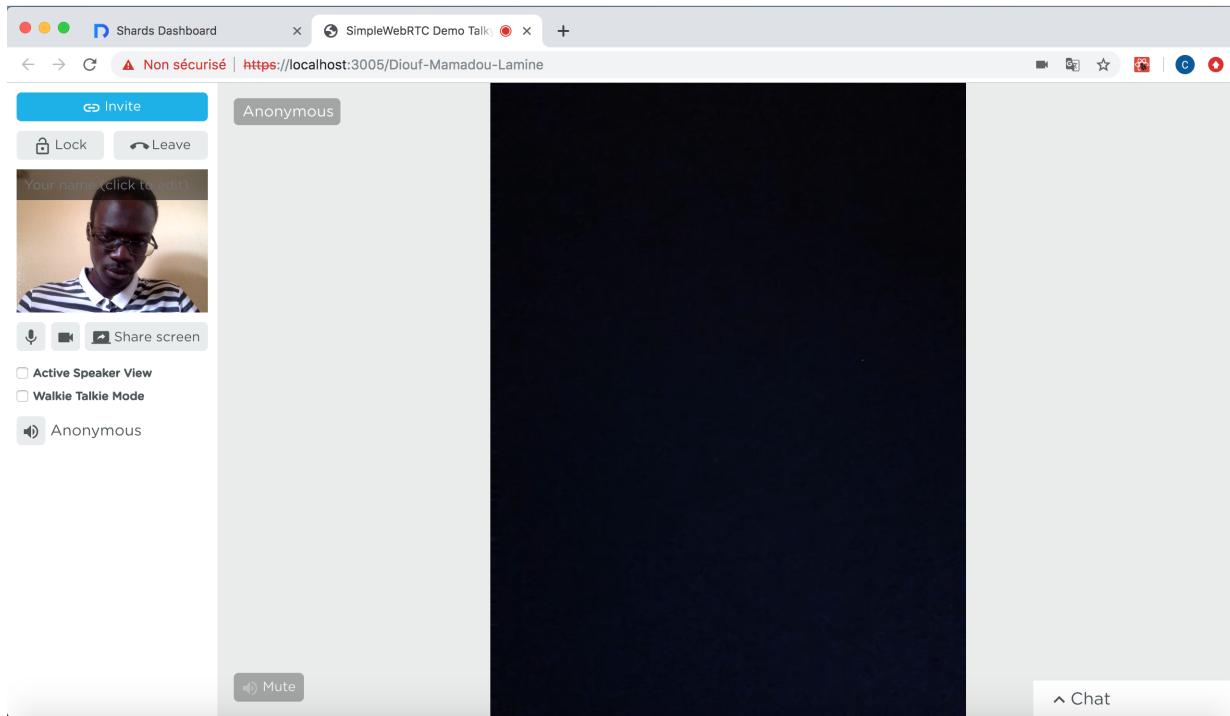


Figure 4.19 Interface de visioconférence

CHAPITRE 5 CONCLUSION

L'apport des nouvelles technologies de l'information et de la communication dans un domaine tel que la médecine ne devrait en aucun cas se limiter à la manipulation, au stockage et à la recherche de données. Les technologies dérivées de l'informatique devraient être utilisées pour faire valoir ces données en tirant le maximum d'informations qu'elles nous offrent.

C'est dans cette perspective que la numérisation des systèmes d'informations s'accompagne de la mise en place d'outils liés à l'intelligence artificielle pour tirer profit des grandes masses de données produites. Images, signaux et données textuelles doivent davantage faire l'objet d'analyse dans le but de faciliter et de rendre fiable certaines décisions.

Il faut cependant noter qu'il n'est pas question de concevoir un système qui remplacera un médecin, mais plutôt de mettre à sa disposition un outil visant à l'assister dans ses tâches quotidiennes.

5.1 Synthèse des travaux

Dans ce rapport composé de trois parties, nous avons d'abord fait une étude des notions qui ont servi à l'implémentation de la solution en expliquant les concepts de Dossier Patient Informatisé, d'électrocardiogramme ainsi que de machine learning. Ensuite, nous avons fait un exposé de l'étude conceptuelle complète de la solution. Enfin, la dernière partie du rapport décrit les différentes phases de l'implementation du logiciel.

5.2 Limitations de la solution proposée

Étant donné que la médecine est un domaine vaste, nous n'avons pas pu prendre en compte toutes les spécialités. La solution proposée ne se limite qu'à l'aide à la décision pour les maladies cardiovasculaires. Cette limitation est la plus importante dans la mesure où un grand volume de données reste inexploité. C'est le cas par exemple des remarques des médecins, du dialogue médecin-patient, des images radiographiques etc.

De plus, le logiciel proposé, étant une application web, oblige les utilisateurs à faire usage d'un navigateur pour y accéder. La saisie de l'URL à chaque besoin d'utilisation est une limite du projet. Or cette application n'est utilisé par un nombre restreint de personnes (les médecins). Son accès pourrait donc se faire via une application de bureau pour exemple pour faciliter le travail du médecin.

5.3 Améliorations futures

Au regard des limitations soulevées, la première amélioration proposée est l'intégration d'autres outils d'aide à la décision pour les autres spécialités de la médecine. Ainsi, en plus de servir aux cardiologues, l'application pourrait servir à d'autres spécialistes.

De plus, chaque patient devrait pouvoir avoir accès et consulter les informations contenues dans son dossier. Cette fonctionnalité permettrait aux patients de suivre également leur état de santé. Il serait donc pertinent de mettre en oeuvre une application mobile puisque les terminaux mobiles tels que les smartphones et les tablettes sont les plus utilisés.

En outre, les évolutions futures de l'application devraient permettre aux cardiologues de récupérer les ECG à partir de l'électrocardiographe qui est l'appareil dont ils se servent pour réaliser l'examen. Ainsi, il ne sera plus nécessaire de charger le fichier contenant les données manuellement.

Pour finir, un système de recommandation de traitements, basée sur les prescriptions antérieure des médecins, pourrait être ajouté au système.

RÉFÉRENCES

- [1] [online]URL : <http://tpe-electrochimie-1s4.over-blog.com/article-de-la-batterie-au-pacemaker-64895854.html> [cited 10 mai 2019].
- [2] [online]URL : <https://ecgwaves.com/ecg-normal-p-wave-qrs-complex-st-segment-t-wave-j-p/> [cited 10 mai 2019].
- [3] [online]URL : <https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn> [cited 14 mai 2019].
- [4] [online]URL : <https://www.caducee.net/DossierSpecialises/systeme-information-sante/dmi.asp#ctabs5> [cited 24 juillet 2019].
- [5] [online]URL : <https://www.ottawaheart.ca/fr/examen-intervention/electrocardiographie> [cited 10 mai 2019].
- [6] [online]URL : <https://www.futura-sciences.com/sante/definitions/medecine-electrocardiogramme-3353/> [cited 10 mai 2019].
- [7] [online]URL : <https://www.ottawaheart.ca/fr/maladie-du-c\T1\oeur/les-arythmies-anomalies-du-rythme-cardiaque> [cited 13 mai 2019].
- [8] [online]URL : <https://physionet.org/physiobank/database/mitdb/> [cited 12 mai 2019].
- [9] [online]URL : <https://scikit-learn.org/> [cited 13 mai 2019].
- [10] Charu C. Aggarwal. *Neural Networks and Deep Learning*. Springer International Publishing AG, 2018.
- [11] Ulas Baran Baloglu, Muhammed Talo, Ozal Yildirim, Ru San Tan, and U Rajendra Acharya. Classification of myocardial infarction with multi-lead ecg signals and deep cnn. *Pattern Recognition Letters*, 2019.
- [12] Timothy M.D. Ebbels. Non-linear methods for the analysis of metabolic profiles. *The Handbook of Metabonomics and Metabolomics*, 2007.
- [13] Aurélien Géron. *Hands-On Machine Learning*. O'reilly, 2017.
- [14] Tracy D Gunter and Nicolas P Terry. The emergence of national electronic health record architectures in the united states and australia : Models, costs, and questions. *J Med Internet Res.*, 2005.
- [15] Mohammad Kachuee, Shayan Fazeli, and Majid Sarrafzadeh. Ecg heartbeat classification : A deep transferable representation. 2018.

- [16] Alexandre Kowalczyk. *Support Vector Machines Succinctly*. Syncfusion, 2017.
- [17] Yaoguang Li and Wei Cui. Identifying the mislabeled training samples of ecg signals using machine learning. *Biomedical Signal Processing and Control*, 2019.
- [18] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Yalwalkar. *Foundations of Machine Learning*. The MIT Press, 2018.
- [19] Josh Patterson and Adam Gibson. *Deep Learning A practitioner's approach*. Oreilly, 2017.
- [20] Association pour l'Avancement de l'Instrumentation Médicale. *Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms*. 2012.
- [21] Jeen-Shing Wang, Wei-Chun Chiang, Yu-Liang Hsu, and Ya-Ting C. Yangb. Ecg arrhythmia classification using a probabilistic neural network with a feature reduction method. *Neurocomputing*, 2013.