



REPUBLIQUE DEMOCRATIQUE DU CONGO
ENSEIGNEMENT SUPERIEUR ET UNIVERSITAIRE
**INSTITUT SUPERIEUR D'ENVIRONNEMENT
SOLIDAIRE ET DE DEVELOPPEMENT
DURABLE**

ISESOD-GOMA

**CONTRIBUTION À L'AMÉLIORATION
DES PERFORMANCES
DES SERVICES MÉDICAUX
URGENTS : INTELLIGENCE ARTIFICIELLE,
BIGDATA ET RÉALITÉ VIRTUELLE**

MEMOIRE

- Mémoire en vue de l'obtention du Diplôme de Licence en Réseau Informatique et Télécommunication
- Présenté par **BUSANGA SHAMAMBA Seth**
- Dirigé Par **CT. Josué KALEMA**

■ Année Académique

2021 - 2022

ÉPIGRAPHE

« Quand un homme est blessé, il découvre la haine,
Quand il blesse quelqu'un il découvre la culpabilité ainsi que la rancœur des victimes.
C'est par ce que l'on connaît ces différents sentiments
Qu'on peut se montrer bon envers autrui.
Connaitre la souffrance, nous aide à grandir.
Tu verras avec l'expérience ont fini par apprendre à
A Prendre des décisions seul
Avec les peines, les souffrances
On finit par trouver des réponses. »

(Motivation, 2020)

DEDICACE

Ce présent travail nous le dédions à la dignité et l'honnêteté incarné : notre très chère père
KAMBERE NYARUSUMBA Désire :

À nos frères et sœurs ;

À nos amis et connaissances ;

À tous les prestataires de services sanitaire ;

Et à toute la communauté de développeurs et chercheurs en Intelligence Artificiel.

REMERCIEMENT

La fin d'une chose vaut plus que son début dit-on. Ce mémoire en Réseau Informatique et Télécommunications, un mérite de ces quatre années d'études Universitaires ; est une expression de divers efforts physiques, moraux et intellectuels. C'est ainsi que, nous tenons à exprimer nos vifs et sincère remerciements à tous ceux qui par leurs travaux, leurs assistances, leurs idées, leurs collaborations ou leurs expériences, ont participé de près ou de loin à la réalisation de ce travail ; qu'ils trouvent ici, l'expression de notre profonde gratitude.

Nous remercions l'Eternel Dieu Tout Puissant, pour sa protection et ses bienfaits qu'il ne cesse de renouveler en notre égard sans oublier ce souffle de vie qu'il nous accorde au jour le jour enfin d'arriver à la réalisation du présent travail.

Nos remerciements s'adressent à toutes les autorités académiques de l'Institut Supérieure d'Environnement Solidaire et de Développement Durable, ISESOD en sigle pour ce bon cadre où nous avons pu cultiver notre intellect dans un climat scientifique digne de sa renommée.

Nous tenons à exprimer nos remerciements avec un grand respect à notre Directeur le CT. Josué KALEMA pour son investissement dans tous les aspects de mon travail et de ma rédaction, des fois malgré la distance qui nous sépare. Il m'a soutenu à chaque instant de ce mémoire et m'a gardée à un moment où je ne me serais pas gardée moi-même. Si j'ai pu conclure mes travaux de mémoire, un grand mérite lui revient qui a répondu présent à chaque fois où j'en avais besoin. Grâce à ses remarques pertinentes et des encouragements sans cesse renouvelés, j'ai pu avancer en toute assurance.

Ce mémoire a été réalisé dans le cadre d'une cotutelle entre des laboratoires, des officines pharmaceutiques et certaine clinique. Je remercie leurs directeurs ainsi que tous les chercheurs de ces laboratoires pour l'accueil et les conditions de travail privilégiées qui m'y ont été offertes. Je n'oublie pas le personnel administratif, qui n'ont ménagé aucun effort pour me faciliter les démarches administratives tout au long de ce mémoire.

Mes remerciements vont aussi à ma famille et à mes amis qui m'ont permis de ne jamais dévier de mon objectif final. Ces remerciements ne peuvent s'achever, sans une pensée pour mes parents qui ont tout sacrifié pour leurs enfants, n'épargnant ni santé ni efforts ni argent : Mon père Désiré KAMBERE, son soutien constant est pour moi le pilier fondateur de ce que je suis et de ce que je fais, son ambition illimitée m'a toujours inspirée. Maman Grace qui a su m'écouter et pris le temps d'entendre mes histoires de biométrie et d'intelligence artificiel. Elle a su être là dans tous les moments, les bons comme les difficiles. Merci Maman.

Nous pensons à tous ceux qui ont contribué sur le plan tant affectif qu'amical et le goût du savoir pour l'arrivée à terme de notre travail de mémoire. Que tous mes frères et sœurs qui ont contribué de près ou de loin à notre formation trouvent leur part dans cette œuvre ; mercis pour vos encouragements.

Nous n'oublions pas de penser à tous les amis et camarade, nous saluons votre participation. Que ceux-là, qui ont contribué de près comme de loin et qui ne sont pas nommément cité, trouvent ici nos remerciements le plus chaleureux et fraternel en Jésus-Christ.

RESUME

Pour planter le décor avec un quadruple anglicisme, chatbot, deeplearning, computer vision et le extend reality sont radicalement devenus les maîtres mots de notre siècle.

Notre mémoire vise à apporter une contribution dans le domaine de agents conversationnels médicales qui est devenu un facteur clés de l'automédication, dans le domaine de l'authentification biométrique par l'iris (l'une des modalités les plus précises et difficiles à pirater), le visage (l'une des modalités les moins intrusives et les moins coûteuses) et celui d'assurer une prise en charge appropriée des patients victimes d'un accident vasculaire cérébral (AVC). Ceci implique des mesures immédiates lors de l'admission du patient, des procédures de revascularisation, la mise en route de la prévention et de l'analyse des complications aigues, un bilan des facteurs étiologiques et le début de la rééducation tout en considérant le patient dans sa globalité.

A travers ce travail, nous abordons plusieurs aspects importants de la biométrie mono et multimodale. Nous commençons par dresser un état de l'art sur la biométrie monomodale par l'iris et par le visage et sur la multimodalité iris/visage, avant de proposer plusieurs approches personnelles de reconnaissance d'individus pour chacune des deux modalités. Nous abordons, en particulier, la reconnaissance faciale par des approches classiques reposant sur des combinaisons d'algorithmes et des approches bio-inspirées émulant le mécanisme de la vision humaine. Nous démontrons l'intérêt des approches bio-inspirées par rapport aux approches classiques à travers deux méthodes. La première exploite les résultats issus de travaux neuroscientifiques indiquant l'importance des régions et des échelles de décomposition utiles à l'identification d'un visage. La deuxième consiste à appliquer une méthode de codage par ordre de classement dans la phase de prétraitement pour renforcer le contenu informatif des images de visage.

Pour ce qui est de la prise en charge des patients victimes d'un accident vasculaire cérébral, nous élargissons notre champs d'étude pour aborder plus la prévention et l'analyse des complications aigües enfin de déterminer les facteurs de risque et ainsi envisager une prise en charge précoce. Des logiciels de traitement, des algorithmes de diagnostic, des interfaces utilisateurs, et des systèmes de collecte et de diffusion d'informations entre tous les acteurs concernés. Les dispositifs proposés sont pensés pour être les moins invasifs et intrusifs possibles. Pour une meilleure interaction avec le patient un chatbot accompagnerons nos interfaces comme agent conversationnel vu que dans les domaines de l'intelligence artificielle (IA) et du bigdata, ce dernier permettent de collecter et d'utiliser un maximum de données pour offrir un service plus performant et plus personnalisé. On loue également leur meilleure accessibilité en comparaison des outils existants, car ils nous permettent d'accéder à un service sans avoir à ouvrir une application mobile ou un navigateur Internet. Enfin, des chercheurs ont observé que les individus sont disposés à parler à des robots-assistants plus longtemps qu'à des êtres humains, et aussi plus tentés de leur confier leurs secrets intimes.

Mots-clefs :

Chatbot, deeplearning, computer vision, authentification biométrique, accident vasculaire cérébral, algorithme, reconnaissance faciale, bio-inspirer, neuroscientifique, intelligence artificiel, bigdata, robot-assistant.

ABSTRACT

To set the scene with a quadruple Anglicism, chatbot, deeplearning, computer vision and extend reality have radically become the watchwords of our century.

Our thesis aims to make a contribution in the field of medical conversational agents which has become a key factor in self-medication, in the field of biometric authentication by the iris (one of the most precise and difficult to hack modalities), face (one of the least intrusive and least expensive modalities) and that of ensuring appropriate management of stroke patients. This involves immediate measures upon admission of the patient, revascularization procedures, initiation of prevention and analysis of acute complications, assessment of etiological factors and initiation of rehabilitation while considering the patient in its entirety.

Through this work, we address several important aspects of mono and multimodal biometrics. We begin by drawing up a state of the art on monomodal biometrics by the iris and by the face and on the multimodality iris/face, before proposing several personal approaches to individual recognition for each of the two modalities. We approach, in particular, facial recognition by classical approaches based on combinations of algorithms and bio-inspired approaches emulating the mechanism of human vision. We demonstrate the interest of bio-inspired approaches compared to classical approaches through two methods. The first exploits the results from neuroscientific work indicating the importance of regions and decomposition scales useful for identifying a face. The second is to apply a rank-order coding method in the pre-processing phase to enhance the information content of face images.

With regard to the management of patients who have suffered a stroke, we will broaden our fields of study to address more prevention and analysis of acute complications, finally to determine the risk factors and thus consider taking in early care. Processing software, diagnostic algorithms, user interfaces, and systems for collecting and disseminating information between all the actors concerned. The proposed devices are designed to be the least invasive and intrusive possible. For better interaction with the patient, a chatbot will accompany our interfaces as a conversational agent, since in the fields of artificial intelligence (AI) and big data, the latter make it possible to collect and use as much data as possible to offer a better service, efficient and more personalized. We also praise their better accessibility compared to existing tools, because they allow us to access a service without having to open a mobile application or an Internet browser. Finally, researchers have observed that individuals are willing to talk to robot assistants longer than to human beings, and also more tempted to confide their intimate secrets to them.

Keywords:

Chatbot, deeplearning, computer vision, biometric authentication, stroke, algorithm, facial recognition, bio-inspire, neuroscientist, artificial intelligence, bigdata, robot-assistant.

SOMMAIRE

Epigraphie	
Dédicace	ii
Remerciement.....	iii
Résumé	iv
Abstract	v
Notations	ix
1. Minuscules Latines	ix
2. symboles mathématiques	ix
3. Abréviations.....	x
Listes des figures	xi
Listes des équations.....	xiii
Chapitre I. INTRODUCTION GENERALE.....	- 1 -
I.1. contexte de l'étude	- 1 -
I.2. Problématique.....	- 3 -
I.3. Objectifs.....	- 8 -
I.4. Choix et intérêts	- 9 -
I.5. méthodologie	- 9 -
I.6. Délimitation spatio-temporelle.....	- 11 -
I.7. Considération éthique	- 11 -
I.8. ossature du travail	- 11 -
Volet de généralité	- 13 -
CHAPITRE. II. GÉNÉRALITÉ ET CHAMPS D'APPLICATION	- 14 -
II.1. INTRODUCTION	- 14 -
II.2. Première section : aperçu thématique	- 14 -
A. Le système de santé	- 14 -
B. la E-Santé.....	- 15 -
C. Les ressources médicales	- 15 -
D. L'intelligence artificielle	- 15 -
E. Le chatbot.....	- 16 -
F. La réalité virtuelle.....	- 16 -
G. Le big-data	- 17 -
II.3. Deuxième section : champs d'application	- 17 -
II.3.1. Automatisation de la détection	- 18 -
II.3.2. La santé numérique et le chatbot.....	- 30 -
II.3. conclusion	- 33 -

CHAPITRE III. ETAT DE L'ART.....	- 35 -
III.1. Introduction.....	- 35 -
III.2. Section première : De quoi il était question ?	- 36 -
III.3. Section Deuxième : Ligne de démarcation	- 38 -
III.4. CONCLUSION	- 40 -
CONCLUSION DU PREMIER VOLET	- 42 -
Volet d'analyse et de technicité	- 43 -
Chapitre IV. APPROCHES, ANALYSEs ET TECHNICITEs.....	- 44 -
IV.1. Introduction.....	- 44 -
IV.2. Section première : approches techniques	- 45 -
IV.2.1. Mapping des chatbots existant	- 45 -
IV.2.2. Mathématique des réseaux de neurones	- 49 -
IV.3. Section Deuxième : Analyse de l'objet d'étude	- 55 -
IV.3.1. Maladie cardio-vasculaire	- 55 -
IV.3.2. Cardiopathie hypertensive	- 55 -
IV.3.3. Infarctus aigu du myocarde.....	- 56 -
IV.3.4. Angine de poitrine.....	- 57 -
IV.3.5. Arythmies cardiaques.....	- 57 -
IV.3.6. Insuffisance cardiaque	- 58 -
IV.3.7. Analyse épidémiologique	- 58 -
IV.3.8. Les Facteurs risques	- 60 -
Analyse de morbidité	- 61 -
IV.4 conclusion	- 62 -
Conclusion du deuxième volet.....	- 63 -
Volet de résultats	- 64 -
Chapitre V. Mise en œuvre et Modélisation.....	- 65 -
V.1. introduction	- 65 -
V.2. mise en œuvre et AQUISITION de données	- 65 -
V.2.1. Base de données médicale MIMIC III	- 65 -
V.2.2. Sélection des sujets et des signaux	- 66 -
V.2.3. Présentation de l'approche	- 68 -
V.3. Modélisation d'un système	- 74 -
V.3.1. Le Processus Unifié.....	- 74 -
V.3.2. Language de modélisation	- 76 -
V.4. Présentation des Architecture du nouveau système	- 77 -
V.4.1. Exigences fonctionnelles	- 77 -
V.5. Conclusion.....	- 79 -

CHAPITRE VI. Le resultat : Classification multi-label par ensemble pour le suivi médical	- 80 -
VI.1. Introduction.....	- 80 -
VI.2. Approche proposée : COLEX	- 80 -
VI.2.1. La méthode de RAKEL	- 80 -
VI.2.2. Le coefficient de corrélation binaire ϕ	- 81 -
VI.2.3. Facteur de corrélation globale	- 81 -
VI.2.4. Stratégie d'échange : Méthode COLEX	- 82 -
VI.3 Résultats et discussions	- 83 -
VI.3.1 Application de la méthode Colex à la base de données médicale	- 83 -
VI.3.2 Illustration des performances de la méthode proposée sur des bases de données réelles multi-label- 84 -	
VI.3.4 Impact des paramètres k et n	- 85 -
VI.3.5. Discussion.....	- 87 -
VI.4. Implementation : technologie, Architectures, Modeles et dataset view.....	- 89 -
VI.4.1. Etude des Technologie	- 89 -
VI.4.2. Graphique et Dataset	- 92 -
VI.4.3. presentations de models et DES INTERFACES graphiques	- 96 -
Conclusion et perspective.....	- 101 -
Bibliographie	- 104 -

NOTATIONS

1. MINUSCULES LATINES

(m,n)	: représentation matriciel
act()	: fonction d'activation
agr()	: fonction d'agrégation
b	: le paramètre de biais
cos	: cosinus
f	: fonction
grad	: gradient
h	: hyperbolique
i	: unité d'incrémentation
k	: constante
l	: log loss(fonction cout)
ln	: logarithme népérien
log	: logarithme
max	: maximum
min	: minimum
n	: nombre d'itération
neur()	: neurones
sin	: sinus
sum	: somme arithmétique
tan	: tangente
w	: indice pour le poids des neurones
X	: indice pour la entrée des neurones

2. SYMBOLES MATHEMATIQUES

<	: signe de comparaison d'infériorité
=	: affectation
Δ	: dérivée
∂	: dérivée partielle
π	: élément trigonométrique Pi
≡	: est identique à
∫	: intégrale
μ	: micro
+	: opérateur arithmétique d'addition
/	: opérateur arithmétique de division
*	: opérateur arithmétique de multiplication
-	: opérateur arithmétique de soustraction
±	: plus ou moins
%	: pourcentage
√	: radicale
≤	: signe de comparaison d'infériorité ou d'égalité
>	: signe de comparaison de supériorité
≥	: signe de comparaison de supériorité ou d'égalité
∞	: signe représentative de l'infinié
Σ	: sommation de plusieurs élément

3. ABREVIATIONS

ANSSI	: Agence Nationale De La Sécurité Des Système D'information
API	: Application Programming Interface
APP	: Application
AVC	: Accident Vasculaire Cérébral
CMD	: Connected Medical Device
CNN	: Convolution Neural Network
DL	: Deeplearning
DME	: Dossier Médical Informatisé
GAFAM	: Google Apple Facebook Amazon Microsoft
GAN	: Generative Adversity Network
HIS	: Hospital Information System
HTA	: Hypertension Artérielle
IA	: Intelligence Artificiel
IDE	: Integrated Developpement Kit
IOMT	: Internet Of Medical Thing
IOS	: Iphone Operating System
IOT	: Internet Of Thing
IRM	: Imagerie Par Résonnance Magnétique
AMEN	: Assistante Médicale Extra Numérique
ML	: Machine Learning
NLP	: Natural Language Processing
NN	: Neural Network
OMS	: Organisation Mondiale De La Santé
RF	: Random Forest
ROS	: Robot Operating System
SAAS	: Soft As A Service
SAL	: Structured Query Language
SGBD	: Système De Gestion De Base De Données
SI	: Système D'information
SIG	: Système D'information De Gestion
SIH	: Système D'information Hospitalier
SII	: Système Informatique Informatisé
MVC	: maladies cardio-vasculaires
SIO	: Système D'information Organisationnel
UML	: Unified Model Language
URL	: Uniform Ressource Locator
WWW	: World Wide Web

LISTES DES FIGURES

Figure 1 : En 1956, le séminaire de Dartmouth.....	- 19 -
Figure 2 : une équivoque à levée : le Deep Learning, Le Machine Learning, L'intelligence Artificiel	- 20 -
Figure 3 : modèle représentative : apprentissage superviser	- 21 -
Figure 4 : modèle représentative : apprentissage non-superviser.....	- 22 -
Figure 5 : modèle représentative : apprentissage par renforcement	- 23 -
Figure 6 : représentation d'un neurone naturel.....	- 25 -
Figure 7 : Neurone Artificiel complet	- 26 -
Figure 8 : représentation schématique d'un réseau de neurones	- 26 -
Figure 9 : neurone complet	- 26 -
Figure 10 : réseau de neurones profond	- 27 -
Figure 11 : schéma complet d'un réseau de neurone convolutifs	- 28 -
Figure 12 : composante et structure de la e-santé	- 31 -
Figure 13 : L'organisation technique d'un chatbot peut être représentée de la manière suivante :	- 32 -
Figure 14 : internet des objets connectée lier à la e-santé.....	- 44 -
Figure 15 : représentation graphique tangentielle d'une dérivée	- 50 -
Figure 16 : méthode numérique pour les calcul du gradient.....	- 51 -
Figure 17 : recherche du minimum pour une fonction	- 52 -
Figure 18 :réseau de neurones pour classification	- 53 -
Figure 19 : la différence entre une artère normale et une artère rétrécie	- 55 -
Figure 20 : Répartition des décès par maladies cardio-vasculaires standardisée selon l'âge par 100000 hab en 2000	- 59 -
Figure 21 : Répartition des décès par maladies cardio-vasculaires standardisée selon l'âge par 100000 hab en 2017	- 59 -
Figure 22 : Incidence des maladies cardio-vasculaire en 2017	- 61 -
Figure 23 : Prévalence des maladies cardio-vasculaire en 2017.....	- 62 -
Figure 24: Extrait de la table "ADMISSIONS" de la base de données clinique	- 66 -
Figure 25 : Histogrammes des durées de séjour pour les patients avec et sans FA.	- 67 -
Figure 26 : Statistiques de disponibilité des signaux physiologiques.....	- 68 -
Figure 27 : Processus de prédiction de la FA.....	- 69 -
Figure 28 : Répartition des fenêtres selon les bases d'apprentissage et de test.	- 71 -
Figure 29 : Exemple de distribution des masses.	- 73 -
Figure 30 : algorithme de la matrice avec la stratégie d'échange	- 82 -
Figure 31 : algorithme de la matrice avec la stratégie d'échange optimisée.....	- 83 -

Figure 32 : – Comparaison entre RAkEL et COLEX pour différentes valeurs de n avec k = 3 pour la base de données Genbase.....	- 85 -
Figure 33 : –Comparaison entre RAkEL et COLEX pour différentes valeurs de n avec k = 10 pour la base de données Genbase.	- 86 -
Figure 34 : Comparaison entre RAkEL et COLEX pour différentes valeurs de n avec k = 2 pour la base de données Yeast.....	- 88 -
Figure 35 : Architecture du système AMEN	- 89 -
Figure 36 : Application web présentation	- 90 -
Figure 37 : représentation d'un api	- 91 -
Figure 38 :mimic-III data base	- 93 -
Figure 39 : data mining : jonction de la table Drug_prescription et admission.....	- 93 -
Figure 40 : étude sur la validité de données par colonne row_id, subject_id : authentification	- 94 -
Figure 41 : étude sur la validité de données par colonne startdate, enddate : séjour	- 94 -
Figure 42 : étude sur la validité de données par colonne pour les médicaments	- 94 -
Figure 43 : relation entre toute les variable du heart attack dataset	- 95 -
Figure 44 : représentation de notre modèle par tri des hyper paramètre par model pré entraîner	- 95 -
Figure 45 : installation de packages	- 96 -
Figure 46 : entraînement pour le chatbot	- 97 -
Figure 47 : test en console de notre chatbot.....	- 98 -
Figure 48 : Détection des maladie basée sur ML diabets	- 98 -
Figure 49 : Détection des maladie basée sur ML Heart Disease Prediction	- 99 -
Figure 50 : optical character Image to Text to audio	- 99 -

LISTES DES EQUATIONS

Équation 1 : formalisation d'une fonction d'apprentissage	- 24 -
Équation 2 : définition de la dérivée.....	- 50 -
Équation 3 : équation tangentielle d'une dérivée	- 50 -
Équation 4 :le vecteur gradient	- 50 -
Équation 5 : fonction d'erreur	- 53 -
Équation 6 : produit de convolution	- 54 -
Équation 7 : liaison de point d'un réseau de neurones	- 54 -
Équation 8 : moyenne de donnée extraite	- 70 -
Équation 9 : Ecart-type	- 70 -
Équation 10 : Skewness	- 70 -
Équation 11 : Kurtosis	- 70 -
Équation 12 : Valeur absolue de la différence	- 70 -
Équation 13 : Écart-type de la différence	- 70 -
Équation 14 : Skewness de la différence	- 70 -
Équation 15 : Kurtosis de la différence	- 70 -
Équation 16 : Moyenne de la valeur minimale sur une fenêtre glissante.....	- 71 -
Équation 17 : démonstration de la distribution normale.....	- 72 -
Équation 18 : la fonction masse	- 73 -
Équation 19 : coefficient de corrélation binaire.....	- 81 -
Équation 20 : coefficient de corrélation binaire entre label.....	- 81 -
Équation 21 : Facteur de corrélation globale.....	- 82 -

CHAPITRE I. INTRODUCTION GENERALE

I.1. CONTEXTE DE L'ETUDE

L'intelligence Artificielle (IA) existe depuis des décennies, initialement inventée en 1956 par John McCarthy Professeur adjoint de l'université de Dartmouth. Toutefois, les progrès étaient limités parce qu'il était extrêmement complexe de programmer à la main les algorithmes. Les activités sophistiquées, comme les diagnostics médicaux, la prédiction de résultats, implique des milliers de jeux de données et des relations non linéaires entre les variables. C'est l'accélération de la puissance du calcul informatique et de l'apparition d'une massive quantité de données qui ont permis de rentrer dans l'âge d'or de l'IA. (M. Durand, 2020)

Cette discipline profite en particulier du Machine Learning (ML) où les progrès sont rapides et importants, le ML est une branche de l'IA basée sur l'idée que le système apprend un modèle à partir d'une base de données à grande échelle en utilisant des outils probabilistes et statistiques et en prenant des décisions ou des prévisions sur les nouvelles données. C'est l'imagerie médicale et le service cardiovasculaire, en particulier en oncologie, qui dispose du plus grand nombre de données nécessaires à l'exploitation de l'IA et de la ML. Le ML pourrait ainsi bénéficier aujourd'hui des nombreuses données obtenues par l'imagerie IRM haute résolution. Il existe plus de 15 méthodes d'apprentissage en ML dont chacune utilise une structure algorithmique différente pour optimiser des prédictions basées sur les données reçues. Une approche particulière, Deep Learning (DL) ou apprentissage profond donne actuellement des résultats particulièrement révolutionnaires. (M. Durand, 2020)

L'enjeu du DL est de modéliser le cerveau pour ne pas traiter les informations avec des règles exhaustives mais plutôt en utilisant la pratique et la rétroaction, à la manière d'un enfant qui découvre le monde, le DL utilise une approche équivalente, inspiré du modèle du cortex visuel. Le DL fonctionne de façon quasi similaire à la fonction des neurones du cerveau humain en formant un réseau neuronal convolutif profond. L'application du DL s'est particulièrement illustrée dans le domaine de l'imagerie médicale pour la détection des tumeurs. Différentes architectures de réseaux de neurones convolutif sont de plus en plus étudiées en tant que solutions prometteuses pour concevoir des outils de diagnostic assisté par ordinateur plus précis pour la détection du cancer. C'est la convergence de toutes ces innovations associant réseaux de neurones artificiels multicouches, algorithmes d'analyse discriminante et croissance de la puissance de calcul informatique permettant le traitement de données massives, qui profite à l'applicabilité de l'IA en médecine.

Le chatbot est un nouveau canal de communication digitale, au même titre que les sites internet et les applications mobiles en leur temps. Comme ces derniers, il se développe à la faveur des révolutions. L'une des fonctions premières du chatbot est de répondre aux questions les plus fréquentes sur un sujet donné, c'est pourquoi il suscite un fort intérêt dans le milieu médical. Il peut prendre en charge les interrogations des patients qui aspirent de plus en plus à devenir acteurs et responsables de leur santé. (e-santé, 2019). D'autant que les traitements n'ont de cesse de se complexifier, et que les professionnels de santé ne trouvent plus le temps d'expliquer et de dialoguer avec les patients. L'assistant conversationnel peut permettre de renforcer le lien patient-professionnels de santé, suscitant ainsi de grands

espoirs pour l'efficience du suivi. En effet, la continuité et la qualité de la relation entre le patient et le professionnel de santé, même à distance, constituent un élément majeur de la réussite du traitement, en particulier pour les pathologies chroniques.

Les chatbots peuvent aussi servir à questionner l'usager de santé et collecter des informations pour faciliter son admission à l'hôpital ou pour réaliser des bilans de santé basiques et ainsi alléger le travail des médecins, généralistes ou non. En amont, ils sont même utilisés dans le cadre d'études cliniques pour gagner du temps sur le recrutement des patients et sur la compréhension de la partie législative du consentement. Ainsi, c'est tout le parcours de soin du patient qui pourrait, à terme, être impacté par cette interaction entre l'homme et le robot : non seulement pour le bien-être et la prévention, mais aussi pour le curatif et l'accompagnement dans le traitement. Le chatbot, assistant conversationnel doté d'une compréhension du langage naturel, apprenant, capable à terme de détecter des émotions, et disponible 24h/24, 7j/7, devrait par nature répondre à l'ensemble de ces besoins, bien que le caractère récent de cette technologie ne permette ni encore de l'affirmer avec certitude, ni encore moins de le prouver scientifiquement. (e-santé, 2019)

D'autre part, l'avènement de cet outil dans le champ de la santé n'est pas sans susciter des craintes : de concurrence faite au médecin, de nouvelles fractures numériques, de mauvais usage voire de piratage des données de santé. A cette défiance s'ajoutent les nombreuses questions d'ordre éthique auxquelles il est essentiel de consacrer son attention.

La réalité virtuelle permet à un utilisateur d'interagir en temps réel, au moyen de diverses interfaces, avec un monde tridimensionnel engendré par l'ordinateur. Les thérapeutes se sont saisis des possibilités de ce nouveau mode d'interaction, et ont contribué au développement d'environnements virtuels ciblant le diagnostic, la thérapie, la réhabilitation, et l'évaluation. Nous avons choisi de nous concentrer sur la prise en charge de troubles cognitifs et comportementaux rencontrés dans les domaines de la psychiatrie et de la neuropsychologie. Les technologies de la réalité virtuelle offrent de nouveaux moyens d'évaluer et de réhabiliter les fonctions cognitives. Elles permettent de surmonter certaines difficultés inhérentes aux méthodes traditionnelles et proposent des solutions plus adaptées. Typiquement, les tâches « papier-crayon » sont remplacées par des mises en situation proches des activités de la vie quotidienne. La contribution de cette recherche concerne ainsi l'adéquation de l'approche réalité virtuelle au diagnostic et à la thérapie, et les apports conséquents. (Klinger, 2006)

La réalité virtuelle est une discipline qui se situe à la croisée de plusieurs domaines tels que l'infographie, la conception assistée par ordinateur, la simulation, la tel opération, l'audiovisuel, le travail collaboratif, etc. L'interaction 3D est la composante motrice de la réalité virtuelle. Elle permet à l'utilisateur d'être un acteur capable d'interagir avec les entités de l'environnement virtuel. Il existe différentes techniques d'interaction 3D qui sont d'éditions aux différentes tâches d'interaction 3D à savoir la navigation, la sélection, la manipulation et le contrôle d'application. Dans la plupart des cas, l'interaction 3D est traitée d'un point de vue général et la technique utilisée cherche à accomplir une tâche d'interaction 3D sans prendre en considération les exigences des environnements complexes tels que les environnements pour la téléportation et les environnements collaboratifs. Les Tâches de téléportation sont souvent des tâches complexes et doivent respecter un certain nombre de contraintes, qui sont : des saisies très précises des objets, des sélections et des manipulations sécurisées et une faible charge cognitive pendant l'exécution de la tâche.

Les environnements collaboratifs quant à eux, sont des mondes peuplés, dans lesquels les utilisateurs interagissent entre eux pour réaliser des objectifs ensemble. La complexité de ces environnements est liée à l'interaction d'un groupe d'utilisateurs avec des entités partagées. (OURAMDANE, 2008)

Le monitoring et Bigdata, Dans le cadre de suivi distant la médecine moderne dans les pays développés tels que les États-Unis, la France, La Suisse, etc., a connu de formidables évolutions notamment grâce aux progrès de la recherche et aux innovations technologiques. Les modes de prises en charges des patients ont également été modifiés, tant dans la forme que dans la durée d'hospitalisation. Le progrès de la médecine a permis d'éradiquer ou de soigner un grand nombre de maladies graves, mais de nouvelles maladies apparaissent issues de la mutation des éléments pathogènes, devenue plus résistant, des changements de comportement et de l'application de risque de la vie moderne (Maes, 2018). Les avancées les plus prometteuses de la médecine d'aujourd'hui s'observent dans les domaines de la génétique et de la biologie moléculaire ou biotechnologie. La génétique, grâce au séquençage progressif du génome humain, doit permettre non seulement de soigner mais aussi de mettre en œuvre une véritable politique de prévention issue de la possibilité de cibler les porteurs de risque. Ces innovations permettent de cultiver des souches d'organes vivants pouvant remplacer des organes malades (Eroukhmanoff J, 2018). Les évolutions technologiques ont révolutionné la médecine notamment en matière d'information, de télématique et d'imagerie. L'imagerie médicale connaît un bouleversement technologique. Désormais avec un ordinateur, il est possible de reconstituer le corps humain par le scanner et de le faire revivre sur l'écran. La chirurgie non invasive constitue aujourd'hui l'avancée la plus spectaculaire car le chirurgien n'ouvre plus le corps. Il peut circuler dans les organes, le cœur, le cerveau grâce à des mini robots dont il surveille la progression sur un écran. Dans le domaine des prothèses, des progrès ont été effectués, qu'il s'agisse de la miniaturisation, des matières utilisées ou du recours à l'électronique.

I.2. PROBLEMATIQUE

Nous sommes à l'aube de l'ère des nouvelles techniques d'acquisition d'images médicales, des modèles médicaux prédictifs et des simulations réalistes d'images. Le suivi des patients en temps réel et les bases de données médicales à grande échelle deviennent omniprésentes et offrent la possibilité et l'occasion d'améliorer encore la santé des patients aussi que notre compréhension de ces maladies. Les données contenues dans ces bases de données n'ont d'utilité que par les questions auxquelles elles peuvent aider à répondre, les idées qu'elles peuvent générer et les décisions qu'elles permettent de prendre. Des études cliniques sont menées sur de grandes populations afin de recommander un traitement, d'ajuster la thérapie en examinant des patients similaires, précédemment traités, en comparant leurs résultats et en prédisant ce qui est susceptible de se reproduire (Rjiba, 2022).

Les systèmes de santé sont confrontés, depuis des décennies, à des changements historiques. Ces changements comprennent les maladies chroniques, l'évolution des segments démographiques, l'augmentation des coûts, le vieillissement des populations, etc. Tous ces éléments mettent les systèmes de soins de santé, privés et publics, du monde entier sous pression et drainent toutes les ressources à mettre au service de toute la population. L'un de ces indicateurs de pression est la charge de travail des cliniciens. Le flux de travail des cliniciens s'est considérablement transformé, tant en termes d'augmentation de la charge de

travail que de changement d'orientation. Les cliniciens travaillent, de plus en plus, avec une productivité et une efficacité accrue tout en fournissant des soins médicaux à tous les patients. En outre, les cliniciens doivent faire face à des défis tels que les nouvelles exigences émergentes en matière de conformité réglementaire, l'augmentation de la charge de travail clinique et la croissance rapide de la base de dispositifs médicaux et de connaissances, qui entraînent tous une adaptation très intense et difficile et peuvent conduire à l'épuisement des médecins. Un clinicien stressé entraîne une diminution de la productivité et de l'efficacité du travail et, dans certains cas, une augmentation des erreurs médicales qui peuvent aller de l'oubli de remplir une autorisation de sortie à la négligence "involontaire" des schémas de certaines maladies. Cela conduira à un mauvais diagnostic qui, dans certains cas, s'il n'est pas corrigé à temps, peut entraîner la mort. En outre, les hôpitaux ne parvenant pas à faire face à la demande croissante de soins aigus, l'engorgement des services d'urgence en est une manifestation essentielle (Karboub, 2022).

L'engorgement des urgences, dans de nombreux pays est considéré comme une crise nationale. En 2017, les États-Unis ont enregistré plus de 22 millions de visites aux urgences et un temps d'attente moyen de 40 minutes, dont plus de 16% ont répondu avoir attendu plus d'une heure. Contrairement aux pays européens, les patients peuvent attendre en moyenne quatre heures pour voir un médecin au service des urgences. Cette crise est due à différents facteurs qui peuvent être regroupés en quatre éléments principaux : 1) la manière dont les soins primaires et les traitements sont organisés, car certains cas non prioritaires peuvent demander de l'aide aux urgences ; 2) l'existence et l'efficacité de modèles organisationnels et de parcours cliniques pour les patients dont les maladies sont chroniques ; 3) la présence de goulets d'étranglement liés à la dotation en personnel ou en matériel des urgences ; 4) la manière dont les urgences sont organisées et leur connexion avec le reste de l'hôpital. En revanche, pour les patients en attente d'une intervention chirurgicale ou d'une procédure facultative, considérés comme des cas moins urgents, les patients peuvent attendre plus longtemps. Aux États-Unis, un record de 6 jours comme temps d'attente médian concernant les rendez-vous disponibles pour les soins primaires, et 6,3 jours pour les rendez-vous chirurgicaux en 2017. Alors qu'en Inde, les patients doivent attendre 4 semaines pour être vus par un médecin traitant. Aux Pays-Bas, les patients peuvent consulter un spécialiste pour un diagnostic primaire en 4 semaines et être admis si nécessaire en 7 semaines, et le Royaume-Uni a enregistré un temps d'attente de 18 semaines pour 92% des patients pour commencer leur traitement (Karboub, 2022).

Face à cette énorme pénurie de **ressources médicales**, les prestataires de soins médicaux sont pressés de suivre des politiques d'une nouvelle gestion des ressources, efficaces et en constante rénovation. Cependant, en tant que professionnels voués à la protection des intérêts de leurs patients, les médecins ont donc la responsabilité de mettre leur expertise au service de l'élaboration de politiques d'attribution qui soient équitables et préservent le bien-être de ces patients. Individuellement et collectivement par le biais de la profession, les médecins doivent plaider en faveur de politiques et de procédures qui répartissent équitablement entre les patients les ressources limitées en matière de soins de santé, en respectant de nombreux critères tels que : premièrement, dans certaines circonstances, il peut être approprié de prendre en considération le nombre de ressources nécessaires à la réussite du traitement. Il n'est pas approprié de fonder les politiques d'allocation sur la valeur sociale, les obstacles perçus au traitement, la contribution du patient à la maladie, l'utilisation antérieure des ressources ou d'autres caractéristiques non médicales. Deuxièmement, donner la priorité aux patients pour lesquels le traitement évitera

un décès prématuré ou des résultats extrêmement médiocres, puis aux patients qui connaîtront le plus grand changement de qualité de vie, lorsqu'il existe des différences très importantes entre les patients qui doivent avoir accès à la ou aux ressources rares.

Si l'**activité hospitalière** augmente en termes d'admissions, le nombre de journées d'hospitalisation est en diminution constante. La durée moyenne de séjours au sein des établissements de santé, n'a cessé de décroître au cours des années (Képénérian). Les soins ambulatoires, c'est-à-dire une prise en charge du patient pendant une durée inférieure à vingt-quatre heures dont la chirurgie de jour est une réalité médicale incontournable. C'est une alternative à l'hospitalisation puisqu'il doit s'agir d'une véritable chirurgie substitutive. Toutes les disciplines sont concernées. Le concept de chirurgie ambulatoire est né d'une triple exigence : améliorer la qualité des interventions chirurgicales, réduire le cout de ces interventions et rendre possible l'innovation en matière chirurgicale en partant d'un concept d'organisation radicalement nouveau qui bouleverse les modes de fonctionnement traditionnels (Pierre-louis, 2009). Nous ne manquerons pas de signaler tous ce prodigieux progrès technologique ne datent pas de l'antiquité. La France, l'une de sommité aujourd'hui dans la recherche médicale tire sa force dans la récolte massive des données des malades, des innovations grâce à la combinaison des outils informatiques et de l'électronique. Ces progrès technologiques offrent des techniques de soins plus avantageuses pour le patient. Ces modifications engendrent des bouleversements dans l'organisation hospitalière qui doit sans cesse s'adapter aux nouvelles techniques de soins et modes de prise en charge pour répondre aux demandes des usagers.

Prenant en considération tous ces contraintes, un autre problème se pose lors de **la priorisation des patients** une fois admis à l'hôpital. De nombreuses techniques sont utilisées pour classer les patients ayant un besoin urgent de matériel de sauvetage, en salle de soins spéciaux ou dans les unités de soins intensifs. L'une de ces techniques est l'utilisation de systèmes de scorage de la gravité. En général, ces systèmes de scorage sont utilisés dans les unités de soins intensifs, ils permettent d'évaluer la gravité de la maladie et fournissent une estimation de la mortalité hospitalière. Cette estimation est réalisée en rassemblant les données mesurées régulièrement et spécifiques au patient, telles que les conditions préexistantes, les mesures physiologiques (rythme cardiaque, fréquence respiratoire, température et score de coma de Glasgow), les indices biomédicaux et d'autres données relatives au patient, comme le sexe et l'âge (Bali, 2018). Ces variables sont ensuite pondérées et ajoutées les unes aux autres, ce qui donne un score de gravité. Les conditions médicales préexistantes, l'âge et d'autres facteurs augmentent le risque de mortalité à l'hôpital. Une fois admis le premier jour, les médecins calculent des scores spécifiques pour trier leurs patients, comme l'APACHE¹ (Acute Physiologic Assessment and Chronic Health Evaluation), le SAPS (Simplified Acute Physiology Score) et le MPM (Mortality Prediction Model), qui est le plus utilisé. Ensuite, ils continuent à mesurer d'autres systèmes de scorage répétitifs tels que le SOFA (Sequential Organ Failure Assessment. Les variables appliquées peuvent être regroupées en cinq catégories : âge, comorbidités, anomalies physiologiques, diagnostic aigu et interventions.

En cas de maladie, l'être humain devient un demandeur de soins. De ce fait, l'hôpital est un lieu de prestation de service qui permet la prise en charge des besoins de l'individu et/ou de la population toute entière. Il représente tout établissement sanitaire, public ou

¹ APACHE: Acute Physiologic Assessment and Chronic Health Evaluation
<https://www.accjournal.org/journal/view.php?viewtype=pubreader&number=1119>

privé, où sont effectués les soins médicaux et chirurgicaux ainsi que les accouchements. Il a pour mission de soigner les patients dans le respect des principes du service public : principes d'égalité, de neutralité et d'efficacité (GUILIANO, 2009). Les hôpitaux publics sont également tenus d'assurer un service continu qu'ils devraient être en mesure d'accueillir les personnes dont l'état requiert. Dans ce contexte, tout patient devant être hospitalisé doit suivre un parcourt précis au sein des établissements de santé. En premier lieu, le(la) malade atterri(e) au niveau du service des urgences, ensuite une demande d'hospitalisation est rédigée par le médecin qui décide de l'admission du patient. A la présentation de cette demande d'hospitalisation, le patient est inscrit dans le registre des admissions et un billet d'hôpital est rempli. À partir de ce point, le cheminement commence. Le service public hospitalier a pour fonction principale d'assurer des diagnostics et des soins. Mais il a nécessairement une mission sociale et doit, en outre, concourir aux actions de prévention ainsi qu'à la formation des médecins et du personnel paramédical. Cette mission de formation incombe aux Centre Hospitalo-universitaires en République Démocratique du Congo (RDC). Il existe, par ailleurs, des instituts chargés de former le personnel paramédical. Depuis de nombreuses années, la prise en charge du patient s'appuie sur une succession d'actes de soins morcelés, isolés, réalisés par divers professionnels, sans visibilité. Cette situation mène à revoir les modalités de prise en charge avec plus de transversalité afin d'améliorer la qualité des soins et afin de structurer le parcours de soins du malade avec efficience (CLAUDE, 2014). Face aux besoins du malade, il y a lieu de faire face à diverses contraintes au sein d'un établissement de santé ce qui engendre des dysfonctionnements à plusieurs niveaux de son entrée à sa sortie, donc cet établissement n'arrive pas à répondre aux besoins des individus et/ou de la population entière. Cependant, il est important de s'inscrire dans une démarche d'optimisation de parcours de soins du malade.

En observant **les infrastructures et institutions médicale** nous avons pu constater que notre médecine est très agonissant. Ces deux facteurs estropient notre médecine : le manque d'un équipement unique pouvant effectuer un diagnostic complète se basant sur la prise d'une grande quantité d'information vital du patient de son arrivé et en continue, cela conduit à une en manque des données suffisantes sur le patient. Ces facteurs poussent à une perte de confiance progressive de la part de la population dans les institutions médicales (Flory A., 2021). En se présentant aux près des prestataires de services sanitaire, l'observation a révélé que le patient a du mal à trouver une très bonne assistance, à trouver un infirmier ou un soignant pouvant rester en permanences à son côté. C'est pourquoi un patient souffrant a du mal à entre en contact directement et instantanément avec le service hospitalier ou même demander de l'assistance lorsque son état s'aggrave. Le patient n'arrive pas à joindre un infirmier ou un médecin quand il se retrouve seul dans sa chambre pour lui faire part de ces inquiétudes. Le problème de la mobilité corporelle bienveillante dans les différentes pièces selon les blocs de chambres suscite le mécontentement des patients en difficulté car la tâche de suivi des patients dans leurs chambres respectives pose toujours problème (Eroukhmanoff J, 2018). Le manque de système d'informatique dans nos institutions est à la base de diverse perturbations et trouble en milieux médicale.

Le bureau des entrées est le cœur même des services administratifs de l'hôpital, c'est une structure administrative sur laquelle s'appuie toute la gestion de l'établissement hospitalière. Son rôle ne se limite pas seulement aux tâches basiques de la gestion mais vise également l'évolution et l'exploitation d'un certain nombre d'information et statistique, liées à la comptabilité et à l'évolution sanitaire des patients des journées d'hospitalisation (Belabbes Amina, 2015). Vu que nos institutions sanitaires sont démunies de système

information hospitalier, le bureau d'entrée n'existe que sur des organigrammes cependant son vrai fonctionnement est inexistant. Cela se remarque par la difficulté d'obtentions des informations liées au patient. Les dossiers médicaux des services hospitaliers constituent la principale source d'information dans l'activité hospitalier et mais aussi dans la recherche clinique. Ces dossiers contiennent également des données écrites et des examens imprimés en image.

Dans la pratique quotidienne l'enregistrement des patients hospitalisés, opérés ou bien même vus en consultation, ainsi que les différentes pathologies et interventions sont réalisées manuellement avec les différents aléas fautes, d'oubli et de redondances. En plus, l'organisation de l'hôpital fait que parfois les activités de consultation et d'hospitalisation se déroulent dans deux ou plusieurs blocs différents ce qui rend toute tentatives de faire sortir le dossier médical de patient à chaque consultation tout à fait complexe sur le plan organisationnel. L'exploitation des données de chaque dossier médical et la recherche d'éléments précis sont hasardeuses, longues et surtout non fiables en raison des aléas d'enregistrement, d'archivage voire la perte des dossiers (Flory A., 2021). Avec la multitude des patients, nous avons observé qu'arrivé à identifier le bloc, la chambre du patient n'est vraiment pas chose facile. Le fait de consulter les dossiers du patient fréquemment, vu que ce dernier est sur format papier comme support de stockage cela cause une détérioration des archives et peut même aller jusqu'à occasionner des pertes des documents. La codification et le stockage pose soucis vu la diversité des intervenant pour cette tâche, cela ne facilite pas l'établissement des statistiques médicaux. Le temps de réaction est aussi à déplorer dans la prestation des services sanitaire (Fiorini, 2018).

Cette problématique est aujourd'hui au centre de plusieurs recherches scientifiques, depuis l'avènement de la pandémie covid-19. La politique internationale de la santé vise des soins de qualité ; qui permettrait la prise en charge globale des patients en répondant à leurs besoins. L'espoir du patient est que ses besoins, ses soins soient satisfaits par le soignant. Il entre à hôpital car il vise la présence continue et permanente d'infirmier qui lui prodigue des soins et ces soins doivent être de qualité. La surveillance doit être sûr dans sa période de convalescence pour une réinsertion dans son milieu social le plus tôt possible. De nos jours, la technologie se développe d'une manière remarquable dans tous les domaines et cela grâce à l'outil informatique. Ce dernier est l'art, la technique ou la science qui consiste à traiter de façon relationnelle et automatique les informations communément appelle les données. L'informatique apporte une aide au développement et au fonctionnement des institutions médicales, donc sa gestion quotidienne, par exemple la gestion stratégique, opérationnelle et la prise de décision rationnelle définitive. En effet, dans les institutions médicales des pays développés, la diffusion de l'information se fait via des réseaux informatiques dont le monitoring constitue le point focal. C'est ainsi qu'en associant les techniques de télémédecine à l'informatique facilite largement l'administration et la gestion de patients en proposant ses services télématiques et bureautiques (Alexandre, 2017).

Notre problématique s'efforce de construire un thème de recherche d'actualité, mais aussi de répondre à une série d'interrogations relatives à la place de l'intelligence artificielle et de la réalité virtuelle dans la médecine. Nous avons de ce fait, axé la problématique de notre recherche sur la question de savoir : quelle approche technologique sera plus contributif et performant pour une thérapie digitale moins encombrant et cela en améliorant les performances de service médicaux urgent.

Certes, le domaine d'intelligence artificielle (IA) et celui des agents conversationnels (chatbot) sont maintenant un aimant d'attraction pour plusieurs chercheurs chacun conjuguant ces efforts pour se démarquer et apporte une pierre de plus dans le monde scientifique. Plusieurs recherches ont déjà été effectuées dans différents domaines à l'occurrence de :

KAHUSI JOSEPH Jamal qui venait de traiter sur: « the setting in place of a patient control and monitoring system in the room of a hospital » (Jamal, 2015). Ce chercheur a traité sur la notion du monitoring au sein d'un hôpital en se focalisant sur le suivi de l'état vital du patient et de contrôler ces mouvements pendant que le patient est interné à l'hôpital. Il a aussi souligné que pour améliorer la vie économique de la population africaine, nous devrions faire de la santé et de son amélioration une priorité. Ce dans ce contexte de résultat de la recherche que le chercheur a été insisté à faire des recommandations au ministère de la santé, aux soignants pour des services de meilleure qualité à donner aux patients fréquentant les infrastructures. Dans cette brillante recherche, effectuer minutieusement le chercheur a omis de prendre en compte l'état du patient après sa sortie de l'hôpital.

RONALD NACUA qui a mené sa recherche en parlant sur :« Conception et développement d'un système ambulatoire pour la mesure de l'activité du Système Nerveux Autonome pour la surveillance de personnes âgées. » (Nacua, 2009). Il a pris pour population d'étude les personnes âgées de plus de 55 ans. Il nous a montré que le nombre d'infrastructures dédiées pour la surveillance des personnes âgées ne sera donc pas suffisant et une de solutions qui permettra de soulager ces institutions est celle de surveiller ces personnes dans leur maison de retraite. Cependant, pour assurer un suivi plus approfondi de l'état de la personne (suivi de l'activité et de l'état de santé) il est indispensable de réaliser ce suivi avec des dispositifs portables sur la personne. Les travaux de cette thèse consistent au développement d'un dispositif portable pour le suivi de l'activité du système nerveux autonome de façon ambulatoire. La prise en compte des personnes dont l'âge est inférieur serait un atout majeur dans cette thèse.

A l'heure actuelle, trop de piste des solutions on était proposé apportant chacune une pierre sur la grande muraille toujours en pleine construction : l'amélioration de conditions de vie sanitaire. Ce travail de mémoire est de nature scientifique qui vient compléter ces acquis en proposant comme thème : **contribution à l'amélioration des performances des services médicaux urgents : intelligence artificielle, Bigdata et réalité virtuelle.** Soulignons que notre système portera le nom d'**Amen**.

I.3. OBJECTIFS

La ligne directrice de cette étude étant d'identifier la meilleure méthodologie et de proposer des approches originales de classification adaptées au contexte médical, avec des données manquantes, hétérogènes et imprécises pour un système de suivi médical automatique ayant la capacité d'anticiper la progression de certaines pathologies chez des patients vivants hors de infrastructures hospitalier sans pour autant leur imposer des actions recommandées ou une invasive potentiellement gênante pour eux. Ce système repose entièrement sur la mesure de signaux physiologiques en temps réel et leur analyse. Cela suppose une analyse des besoins afin :

- * D'adapter les technologies au quotidien des patients pour rendre le suivi discret et continu ;
- * D'aider les professionnels de santé à maintenir une prise en charge des patients au meilleur niveau en proposant des systèmes qui assurent un niveau de sécurité supplémentaire en institution pour une intervention rapide en cas de danger.
- * De familiariser les utilisateurs à cette nouvelle approche thérapeutique pour instaurer un climat de confiance aux près de prestataire sanitaire et des patients pour la certification et l'amélioration de l'authenticité des informations dans le collecte de données.
- * D'installer un système de son et d'affichage des détails des données du patient pour permettre d'établir des statistiques de l'amélioration ou de la dégradation de l'État de santé du patient.

I.4. CHOIX ET INTERETS

Le choix de la thématique de ce présent travail a été motivé par la volonté de fournir un système simple d'utilisation qui permettra aux patients de dialoguer avec le personnel soignant pour lui faire part de l'évolution de son état ou partager des inquiétudes et d'aider les prestataires de services sanitaires à avoir un outil d'aider à l'observance. Cela nous a poussé en tant que chercheur, d'élargir des connaissances et d'essaie de comprendre qu'elles sont les différents dysfonctionnements liés au cheminement du malade en milieux hospitalier dès son entrée, à sa prise en charge jusqu'à sa sortie, mais aussi pour faire des découvertes complémentaires sur nos acquis tout au long de notre cursus académique en réseau informatique et télécommunication.

Les résultats de ce présent travail apporteront à notre humble personne la satisfaction et la fierté scientifique de nous savoir désormais figurant parmi ceux qui par leurs labeurs ont construit l'édifice de l'amélioration de condition de vie sanitaire grâce à leur recherche. À son plafond, ce travail nous donne également un aperçu du système de santé en nous montrant les différents défis qui peuvent être résolue par l'intelligence artificielle et les agents conversationnels.

Ce travail est également d'une importance capitale car il va permettre au personnel médical de conserver des documents numériques liées au patient d'une façon sécurisé, une manière plus sûr que la conservation actuelle se faisant sur papier. Les pharmacologues et les urgentistes aurons enfin un outil qui le fournira un échantillonnage et une aperçue en temps réel des antécédents médicaux du patient avant de le faire une prescription ou avant de lui apporte les premières des soins.

L'implémentation de l'IA dans le système de monitoring, un des grands défis de notre siècle en exécutant cette tache nous aurons apporté une contribution significative dans le domaine de la recherche en intelligence artificiel. La communauté de développeurs, de programmeurs et chercheurs trouveront dans ces pages une approche algorithmique pouvant facilement s'intègre dans n'importe quel système de gestion des institutions hospitalière.

I.5. METHODOLOGIE

Notre approche opérationnelle est conçue comme des études de cas multiples, conformément aux ensembles de données utilisés et aux contraintes prédéfinies. Le

phénomène contemporain particulier de ce mémoire est la recherche des utilisations possibles de l'intelligence artificielle, la réalité virtuelle et les agents conversationnels pour aider les patients atteints des maladies chroniques tel que les maladies cardiovasculaires dans divers environnements. Plus précisément, comment les interventions technologiques peuvent être utilisées pour accroître l'efficacité et améliorer les décisions médicales individualisées.

Les solutions techniques proposées associent des algorithmes d'apprentissage, des logiciels de traitement, des algorithmes de diagnostic, des interfaces utilisateurs, et des systèmes de collecte de données et de diffusion d'informations entre tous les acteurs concernés. Avec ces approches ; nous cherchons à résoudre chaque problème indépendamment et à rechercher l'homogénéité avec un modèle de simulation incluant toutes les caractéristiques de la solution.

Pour nous, nous recherchons un filtre capable d'extraire des caractéristiques de haut niveau à partir du volume. La rareté des données rend le domaine de l'imagerie médicale peu adapté aux techniques d'apprentissage profond. Un petit nombre d'échantillons limite la capacité d'un réseau de neurones à apprendre les caractéristiques représentatives dans une image ou un volume. La nature 3D des scans CTA nécessite un modèle 3D pour apprendre les corrélations les plus importantes entre les différentes structures du volume.

Tout d'abord, une étape de prétraitement en vue d'amélioration des vaisseaux est utilisée pour atténuer les structures non-vasculaires, puis un modèle CNN 3D est appliqué pour détecter les artères dans le volume CTA. En post-traitement, les limites des structures cibles sont définies en extrayant l'aorte et la région cardiaque grâce au modèle d'extraction des sous-régions du cœur. Enfin, une étape d'autocorrection est proposée par une analyse intelligente des composants connectés pour rectifier les erreurs de détection.

La finalité de ce travail est donc de proposer des outils originaux de suivi actimétrique qui seront à la portée des patients répondant à leurs attentes et à celles du personnel soignant. En effet, nous montrerons que nos approches se démarquent par son originalité vis-à-vis des systèmes existants sur le marché, notamment au niveau de leur intégration dans la vie courante et leur grande facilité d'utilisation, sur le plan médical en proposant des mesures pertinentes centrées sur des critères d'activités cliniques, et sur le plan d'aide à l'observance qui implique le médecin pour fixer les objectifs. Ces travaux seront illustrés par des approches algorithmique de thérapie digital comme outils d'aide aux patients aux seins et même en dehors des institutions médicales :

- * L'étude des différentes variations de constantes du patient avec pour objectif de mettre en place une prise en charge précoce de ces personnes pré-dépendantes et faire de la prévention. En effet, des études récentes sur les causes de la perte d'autonomie ont mis en évidence le rôle du syndrome de la fragilité comme précurseur spécifique de la perte d'autonomie (Carré, 2008). Le suivi est prévu sur la base d'un système d'un assistant vocal : cette tâche représente la contribution originale de ce présent travail scientifique avec la proposition d'un algorithme de compréhension du langage et de l'émotion humaine.
- * La récolte, l'étude et l'analyse de données vu que la médecine moderne est devenue inconcevable sans l'utilisation des données de santé, volumineuses et hétérogènes, issues de la relation patient-médecin (dossiers des patients, résultats de

biologie et d'imagerie, de la e-santé, de la télémédecine, de l'ensemble des NBIC² (nanotechnologies, biotechnologies, informatique et sciences cognitives), de la littérature médicale et de l'expression personnelle dans les médias sociaux, grâce au approches propose par le bigdata pour la bonne conservation de données et la biométrie pour la sécurité nous déployerons une base de donnée.

I.6. DELIMITATION SPATIO-TEMPORELLE

La certification d'un travail scientifique réside dans sa précision à être située dans le temps ainsi que dans l'espace. Pour que notre travail puisse faire preuve de cette dernière qualité nous ne prétendons pas avoir à couvrir toutes les notions liées à l'intégration de l'intelligence artificielle dans le domaine de la santé, ni même d'avoir épuisé l'intégralité des champs application de ce dernier. C'est le pourquoi de la focalisation de notre mémoire sur les performances de service médicaux urgent sous l'approche d'interaction patient-médecin.

Par ailleurs, Vu l'immensité de notre sujet par rapport au temps mis à notre disposition pour l'accomplissement de notre recherche, nous nous sommes attelé à analyser et traiter les données mise à notre disposition par différentes communautés scientifiques (PhysioBank, Kaggle, CDC, etc....) tout le long de la rédaction et le développement de notre Platform.

I.7. CONSIDERATION ETIQUE

Depuis les origines, les hommes doivent s'être aperçus qu'on ne peut pas se conduire n'importe comment. Comme le monde est ordonné, il doit y aussi avoir un ordre moral. Car il y a une correcte façon de se conduire. Par conséquent, l'éthique a son origine dans les activités quotidiennes comme vivre, se marier, élever les enfants, étudier, enseigner, travailler dans une entreprise ou organisation, ... En effet, il y a une manière correcte de faire ces activités ordinaires : il y a des règles, des indications ou lignes directrices qui indiquent, déterminent comment on doit agir ou procéder dans telle ou telle autre activité. Ces règles et lignes directrices sont devenues des coutumes et données pour quelques pratiques (Innocent HAKUZIMANA KABUMBA, 2020 - 2021).

Ayant connaissance de ce qu'est l'éthique, lors de cette recherche, nous avons dû respecter toutes les lois et normes éthique requises. Toutes les informations de recherche ont été collectées et utilisées avec sécurité et l'utilisation des données ont respecté la confidentialité des archives et elles ont été utilisées à des fins académiques.

I.8. OSSATURE DU TRAVAIL

Le présent manuscrit n'est pas rédigé sous forme d'un bloc, il est structuré par un certain nombre des chapitres pour le rendre plus claire et faciliter sa lecture. Ainsi, il est subdivisé et organisé comme suit :

Le premier chapitre présente la partie introductory de notre mémoire en mettant en lumière notre problématique, le champs d'action de la recherche, la méthodologie de

² NBIC : nanotechnologies, biotechnologies, informatique et sciences cognitives

travail et les objectifs poursuivies. Tout cela dans l'optique de faire sortir les points saillant de notre recherche.

Pour une organisation minutieuse, la suite de la rédaction sera groupée en trois volet : le volet de généralité, le volet d'analyse et de technicité et en dernier le volet de résultat.

Dans le volet de généralités nous traiterons de l'aspect basique et conceptuel de nos différentes thématiques. Il est composé du deuxième chapitre intitulé généralité et champs d'application. Le présent chapitre comprend deux sections dont la première est constitué des approches conceptuelles où nous prélèveront les définitions des différents mots clés liés de près à notre sujet de recherche et la seconde section est celle de l'étude et de la compréhension du domaine hébergent en son sein notre sujet de recherche.

Le troisième chapitre qui présentera les solutions déjà proposer autour de notre problématique sera intitulé état de l'art. Puis nous finirons ce point par une partie dédié aux points de démarcation d'où la scientificité même de ce travail. À la fin de ce volet une synthèse conclusive sera proposer en guise de brève rappelle de la thématique débattue tout le long de ce dernier.

Le volet d'analyse et de technicité : Comme le titre le fait sous-entendre, dans cette partie nous allons dans le quatrième chapitre nous allons présenter les différentes approches technologiques et les modèles algorithmique déjà développé par d'autre chercheurs et nous finirons par faire un choix bien justifier, ce que nous appellerons l'analyse de l'objet d'étude

Dans sa seconde section nous allons décortiquer les manquements des infrastructures et système hospitalière tout en gardant notre angle de recherche pour éviter le débordement.

Le résultat et la nouvelle architecture sera présenté au niveau du volet de résultat au travers d'un projet AMEN. Ce projet vise la validation d'une plateforme complète de surveillance, d'analyse des données et d'un agent conversationnel. La manière dont nous avons exploiter les données pour aboutir à cet algorithme sera présenter au chapitre cinq. En guise, nous montrerons la mise en place du projet par quelques interfaces graphiques réalisées pour interpréter les résultats au chapitre six.

Nous conclurons et nous tirerons finalement les leçons des retours d'expérience par rapport à l'objectif ultime de l'intégration, dans notre système de santé, d'une instrumentation médicalisée de l'habitat de manière plus généralisée pour une meilleure prise en charge de patients à risque des accidents vasculaire cardiaque. Nous donnerons des perspectives de recherches à court et moyen terme sur les dispositifs développés mais également quelques recommandations sur les orientations à suivre.

VOLET DE GENERALITE

La santé est la condition normale et naturelle de tous les êtres vivants, car cette capacité est inhérente à la vie. Donc, elle est perçue comme une ressource de la vie quotidienne, et non comme le but de la vie. La maladie désigne un ensemble d'altération qui engendre un mauvais fonctionnement de l'organisme en redondance, une influence négative sur la santé d'une personne. Cette dernière qui souffre, nécessitant un traitement thérapeutique. Dans ce contexte, le suivi de patients constitue aujourd'hui la majeure partie des activités médicales, que ce soit en ambulatoire comme en établissements de santé qui est un lieu de prise en charge de ses besoins en matière de santé. Le parcours de soins de ces patients s'étale dans la durée et fait appel à l'ensemble des compétences des différents professionnels s'intervenant dans la prise en charge.

Ce volet sera consacré sur la présentation des généralités des concepts de base. Il est composé du deuxième chapitre intitulé généralité et champs d'application. Le présent chapitre comprend deux sections dont la première est constitué des approches conceptuelles où nous prélèveront les définitions des différents mots clés ou concepts de base liés de près ou de loin à notre sujet de recherche et la seconde section est celle de l'étude et de la compréhension du domaine hébergeant en son sein notre sujet de recherche. Le troisième chapitre qui présentera les solutions déjà proposer autour de notre problématique sera intitulé état de l'art et sera clôturer par une section concentré à la démarcation. À la fin de ce volet une synthèse conclusive sera proposer en guise de brève rappelle des thématiques débattue tout le long de ce volet.

CHAPITRE. II. GÉNÉRALITÉ ET CHAMPS D'APPLICATION

II.1. INTRODUCTION

L'importance des travaux de recherches dans le domaine de santé a pris une nouvelle dimension au cours des deux décennies passées. En effet, confrontés à un contexte socio-économique difficile, et des insatisfactions observées auprès de demandeur des soins, la majorité des établissements hospitaliers du monde entier doivent se plier à de nouvelles règles de gestion afin de minimiser les coûts engendrés et de maximiser le confort et le soin des patients. Voilà la raison poussant un bon nombre de chercheur à construire des thématiques autour du monitoring, de l'e-santé, du healthCare, du système informatique hospitalière, etc. tentant d'apporter des nouvelles stratégies d'organisation et de planification dédiées en milieu hospitalier.

Aujourd'hui, le système hospitalier n'est plus abordé d'un seul revers ou même en utilisant une seule technologie mais c'est en combinant deux, trois voire même quatre disciplines qui nous redonne un peu cette lueur d'espoir de courir à une condition de vie sanitaire stable.

L'incompréhension et le non familiarisation de certains concepts cités ou non dans ce paragraphe sera bien définie dans la section ci-dessous.

II.2. PREMIERE SECTION : APERÇU THÉMATIQUE

A. LE SYSTEME DE SANTE

Un système de santé désigne l'ensemble des organisations, institutions et ressources consacrées aux soins. En France, il regroupe les activités sanitaires, sociales et médico-sociales et s'organise aux niveaux national, régional et local à travers une grande diversité d'acteurs (vie-publique-fiche thématique, 2023).

Dans son rapport intitulé "Des systèmes de santé renforcés sauvent plus de vies", l'Organisation mondiale de la santé (OMS) définit les systèmes de santé comme "*la totalité des organisations, institutions et ressources consacrées à la production d'actions visant principalement à améliorer, maintenir ou restaurer la santé*". En France, le système de santé s'appuie sur différents types de structures :

- Sanitaires (qui s'occupent de la prise en charge hospitalière) ;
- Médico-sociales et sociales (à destination de publics âgés, handicapés ou considérés comme fragiles) ;
- Ambulatoires (qui dispensent des soins dits "de ville").

Il implique ainsi plusieurs types d'acteurs :

- Les offreurs de soins ;
- Les producteurs de biens et services en santé ;
- Des institutions publiques ;
- Des financeurs ;
- Les bénéficiaires (patients, proches aidants), qui font vivre la démocratie sanitaire.
- L'ensemble de ces acteurs s'organisent aux niveaux national, régional et local (OMS, 2023).

B. LA E-SANTE

Selon l'Organisation mondiale de la santé (OMS), la e-santé se définit comme « les services du numérique au service du bien-être de la personne » c'est-à-dire comme l'application des technologies de l'information et de la communication (TIC) au domaine de la santé et du bien-être. La télémédecine est une activité professionnelle qui met en œuvre des moyens de télécommunications numériques permettant à des médecins et à d'autres membres du corps médical de réaliser à distance des actes médicaux, alors que la télésanté concerne l'utilisation des systèmes de communication pour protéger et promouvoir la santé.

La e-santé, ou santé électronique, décrit l'ensemble des moyens et services liés à la santé qui utilisent les nouvelles technologies de l'information et de la communication. La e-santé fait appel à Internet, aux applications pour smartphones et aux objets connectés (science, 2023).

C. LES RESSOURCES MEDICALES

Par ressource du système de santé, on désigne l'ensemble des facteurs qui concourent, en se recombinant entre autre la documentation, la finance, les différents appareils, à l'amélioration de la santé de la population, soit par le diagnostic et le traitement des pathologies, soit par leur prévention (Maes, 2018).

D. L'INTELLIGENCE ARTIFICIELLE

L'intelligence artificielle (IA) est un processus d'imitation de l'intelligence humaine qui repose sur la création et l'application d'algorithmes exécutés dans un environnement informatique dynamique. Son but est de permettre à des ordinateurs de penser et d'agir comme des êtres humains (netapp, 2023).

Pour y parvenir, trois composants sont nécessaires :

- Des systèmes informatiques ;
- Des données avec des systèmes de gestion ;
- Des algorithmes d'IA avancés (code)

Pour se rapprocher le plus possible du comportement humain, l'intelligence artificielle a besoin d'une quantité de données et d'une capacité de traitement élevées.

Selon Sage Advices, L'intelligence artificielle (IA) est une branche de l'informatique dédiée à la conception de machines capables d'imiter le cerveau humain dans des tâches telles que l'apprentissage ou le raisonnement. C'est une science interdisciplinaire qui implique un ensemble de logiciels, de logiques, de calculs et de disciplines philosophiques qui ont pour objectif d'amener les ordinateurs à exécuter des fonctions jusqu'alors perçues comme étant propres à l'homme (Sage advices, 2023).

E. LE CHATBOT

Au niveau le plus fondamental, un chatbot est un programme informatique qui simule et traite une conversation humaine (écrite ou parlée), permettant aux humains d'interagir avec des terminaux numériques comme s'ils communiquaient avec une personne réelle. Les chatbots peuvent être aussi simples que des programmes rudimentaires répondant à une requête simple avec une réponse sur une seule ligne, ou aussi sophistiqués que des assistants numériques qui apprennent et évoluent pour fournir des niveaux de personnalisation croissants à mesure qu'ils collectent et traitent des informations (Oracle, 2023).

Un chatbot est un terme anglais qui désigne un robot conversationnel capable de donner l'illusion de tenir une discussion et de répondre à des questions de la part des internautes. Les chatbots jouent un rôle grandissant dans le commerce conversationnel où les échanges ont lieu en langage naturel.

Le chatbot fait son entrée officielle dans le dictionnaire français Le Petit Robert en 2019 sous la définition « agent conversationnel ». Il s'agit d'un robot (terme contracté en « bot ») capable de simuler une conversation (un « chat » en anglais), plus ou moins simple, avec un être humain. Le chatbot interprète un message émis par l'utilisateur et lui répond, généralement par une action pré-déterminée ou acquise. Le chatbot est un nouveau canal de communication digitale, au même titre que les sites internet et les applications mobiles en leur temps (lab-e-santé, 2020).

F. LA REALITE VIRTUELLE

La réalité virtuelle (ou Virtual Reality en anglais) est une expression qui désigne les dispositifs permettant de simuler numériquement un environnement par la machine (ordinateur). Selon les technologies employées, elle permet à l'utilisateur de ressentir un univers virtuel par le biais de ses différents sens : la vue le plus souvent mais aussi le toucher, l'ouïe, l'odorat (Artefacto, 2023).

La réalité virtuelle permet donc à une personne de vivre une expérience d'immersion et de mener une activité senso-motrice dans un monde artificiel. Pour garantir une immersion totale, l'utilisateur se sert d'un casque de réalité virtuelle. Celui-ci utilise le principe d'affichage en 3D stéréoscopique pour placer le visualisateur dans un monde virtuel généré par une machine

"La réalité virtuelle est l'ensemble des sciences et technologies qui permettent à un utilisateur de se sentir présent dans un environnement artificiel. Ainsi la finalité de la réalité virtuelle est de permettre à un ou plusieurs utilisateurs une activité sensori-motrice et cognitive dans un monde artificiel, créé numériquement, qui peut être imaginaire, symbolique ou une simulation de certains aspects du monde réel."

Contrairement à la réalité augmentée, la réalité virtuelle est une technologie qui vise donc à "plonger" un utilisateur dans un environnement artificiel, tout en lui permettant d'interagir intuitivement et naturellement avec cet environnement virtuel (clarte, 2023).

Dans le cadre professionnel, les usages de la réalité virtuelle sont nombreux. Celle-ci peut intervenir à différentes étapes, en phase de conception ou revue de projet, dans un objectif de prévention ou formation, comme outil d'aide à la vente... Elle s'applique à de nombreux domaines comme ceux de l'industrie, de l'éducation, de la santé, de la défense, du sport...

G. LE BIG-DATA

Avant de définir le Big Data, ou les méga données, il est important de bien comprendre ce que sont les données. Ce terme définit les quantités, les caractères ou les symboles sur lesquels des opérations sont effectuées par un ordinateur. Les données peuvent être stockées ou transmises sous forme de signaux électriques et enregistrées sur un support mécanique, optique ou magnétique.

Le terme de Big Data désigne de vastes ensembles de données collectées par les entreprises, pouvant être explorées et analysées afin d'en dégager des informations exploitables ou utilisées pour des projets de Machine Learning (datascientest, 2023).

On définit souvent le Big Data par les « 3 V » qui le caractérisent : le volume et la variété des données, et la vitesse avec laquelle elles sont générées, collectées et traitées. C'est ce qui différencie les « méga données » des données traditionnelles.

La définition du Big Data est la suivante : des données plus variées, arrivant dans des volumes croissants et à une vitesse plus élevée. C'est ce que l'on appelle les trois « V ».

En d'autres termes, le Big Data est composé de jeux de données complexes, provenant essentiellement de nouvelles sources. Ces ensembles de données sont si volumineux qu'un logiciel de traitement de données traditionnel ne peut tout simplement pas les gérer. Mais ces énormes volumes de données peuvent être utilisés pour résoudre des problèmes que vous n'auriez jamais pu résoudre auparavant (oracle, 2023).

II.3. DEUXIEME SECTION : CHAMPS D'APPLICATION

Seuls ceux qui ont vécu dans une grotte ces dix dernières années ont pu ignorer l'incroyable révolution de l'apprentissage automatique, ou Machine Learning (ML) ou abusivement appellé intelligence artificielle (IA). Il ne se passe plus une semaine sans qu'il ne fasse parler de lui : cela a commencé par de formidables progrès en reconnaissance d'images, puis en analyse de la voix, le programme Watson d'IBM³ est ensuite devenu champion du jeu de Jeopardy, on a vu les premières voitures autonomes de Google silloner les routes, puis le programme Alpha GO⁴ de DeepMind a vaincu le champion du monde du jeu de go, le logiciel Libratus⁵ de l'université Carnegie Mellon a écrasé des champions de poker, des patients paralytiques ont pu contrôler le mouvement de leurs membres par la pensée, grâce à un programme qui avait appris à déchiffrer certaines de leurs ondes cérébrales... bref, les succès s'enchaînent et ne se ressemblent pas. Il y a dix ans, de telles intelligences artificielles n'existaient que dans les romans de science-fiction.

L'intelligence artificielle (IA) engendre beaucoup de fantasmes et de craintes, notamment parce qu'elle permet d'autonomiser certaines tâches de manière accrue, jouant ainsi un rôle majeur dans les mutations du travail. On parle d'une technologie qui va se substituer au travail humain et faire disparaître le travail ou, à tout le moins, le raréfier. Mais ce qui va jouer sur les mutations du travail, ce n'est pas tant l'existence de l'IA ou ses progrès, que la manière dont elle sera déployée dans les organisations et les gains de productivité attendus.

³ Le programme Watson d'IBM

⁴ Le programme Alpha GO de DeepMind

⁵ Le logiciel Libratus de l'université Carnegie Mellon

En matière de technologie et de travail en général, il n'y a pas de fatalisme, mais des choix et des orientations stratégiques portés par les dirigeants d'organisations, publiques comme privées, dans un contexte spécifique (économique, sociale, technologique, démographique...). C'est la combinaison de tous ces facteurs qui déterminera en grande partie les effets et les usages de l'IA sur le travail, que ce soit au niveau des pratiques, du contenu, des conditions ou encore des relations entre collègues, clients ou usagers.

II.3.1. AUTOMATISATION DE LA DETECTION

Ce n'est que lorsque nous essayons d'écrire un programme pour imiter le raisonnement des experts médicaux sur un ordinateur que nous pouvons commencer à apprécier pleinement la véritable complexité des tâches visuelles, de l'audition, de l'analyse du traitement et de la compréhension effectuée par leur cerveau. La simplicité de l'extraction des informations pertinentes des images est très trompeuse. L'approche de l'apprentissage automatique est très différente. Au lieu de coder manuellement les règles explicites qui imitent l'approche des cardiologues, nous spécifions un modèle d'apprentissage et laissons l'algorithme d'apprentissage déterminer automatiquement un ensemble de règles en examinant les données, c'est-à-dire en entraînant le modèle. Dans le cadre de l'apprentissage supervisé, un ensemble d'exemples ainsi que les résultats souhaités (par exemple, des images et leurs segmentations respectives) sont présentés à l'algorithme d'apprentissage. L'algorithme sélectionne ensuite les règles qui transforment le mieux les entrées en sorties souhaitées. Il est important que le modèle appris soit généralisable, c'est-à-dire qu'il puisse prédire de manière fiable les résultats pour des images non vues auparavant, tout en ignorant les différences d'acquisition non pertinentes.

II.3.1.A. INTELLIGENCE ARTIFICIELLE

L'intelligence artificielle (IA) est une discipline jeune d'une soixantaine d'années, qui est un ensemble de sciences, théories et techniques (notamment logique mathématique, statistiques, probabilités, neurobiologie computationnelle, informatique) qui ambitionne d'imiter les capacités cognitives d'un être humain (Stephan, 2018). Initiés dans le souffle de la seconde guerre mondiale, ses développements sont intimement liés à ceux de l'informatique et ont conduit les ordinateurs à réaliser des tâches de plus en plus complexes, qui ne pouvaient être auparavant que déléguées à un humain.

La paternité du terme « IA » pourrait être attribué à **John McCarthy** du MIT (Massachusetts Institute of Technology), terme que **Marvin Minsky** (université de Carnegie-Mellon) définit comme « *la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique* ». En 1956, le séminaire de Dartmouth a réuni pendant dix semaines des scientifiques de très haut rang pour définir l'intelligence artificielle. Parmi eux se trouvaient **John McCarthy**, **Marvin Minsky** (cofondateur du laboratoire d'intelligence artificielle au MIT, père des réseaux de neurones), **Herbert Simon** (prix Nobel d'économie 1978, inventeur de la rationalité limitée) et **Claude Shannon** (inventeur de la théorie de l'information) (McCarthy, 1995).

Le terme d'intelligence artificielle a été évoqué pour la première fois en 1950 par **Alan Turing**. Ce mathématicien a élaboré un test permettant de vérifier la capacité d'un système à faire preuve d'intelligence humaine. De nos jours, le test de Turing est toujours utilisé pour

mesurer l'intelligence de certaines machines. Depuis, les algorithmes d'intelligence artificielle contemporains s'attaquent principalement à deux types de problèmes : la première catégorie contient les problèmes qui sont facilement décrits à l'aide d'équations et de relations mathématiques sophistiquées, tandis que la seconde catégorie est constituée de problèmes difficiles à formuler.

Figure 1 : En 1956, le séminaire de Dartmouth



Un exemple de la première catégorie est "Deep Blue"⁶. Garry Kasparov, le champion du monde d'échecs, a été battu début 1997 par un algorithme informatique appelé Deep Blue. En effet, le jeu d'échecs est limité à 64 cases et 32 pièces et donc à un nombre maximum de coups. Cependant, pour le cerveau humain, l'exploration de tous les coups possibles semble un processus difficile et très lent comparé aux algorithmes informatiques effectuant la même tâche. Les problèmes de la deuxième catégorie sont assez différents. Bien qu'ils puissent sembler triviaux pour les humains, ils sont d'une difficulté prohibitive pour être formulés en équations mathématiques simples comme les problèmes de reconnaissance faciale ou d'estimation de l'âge par exemple.

La première approche représente tous systèmes dotés de mécanismes de raisonnement capables d'interpréter des données symboliques qui constituent une base de connaissance. Cette approche permet de traiter des questions de logique formelle pour prendre des décisions intelligentes en fonction de règles, faits et raisonnements établis par le savoir humain. La seconde approche, s'inspire du fonctionnement des neurones biologiques. En effet, cette méthode regroupe tous les systèmes composés de sous-élément interconnectés capables de traiter et de communiquer des informations entre eux. En outre, cette intelligence artificielle permet d'extraire des règles implicites contenues dans des bases de données volumineuses. Ces réseaux neuronaux artificiels sont généralement optimisés par des méthodes d'apprentissage afin d'apprendre de leurs erreurs par comparaison statistiques avec les résultats attendus.

Le développement général des volumes de données disponibles et le perfectionnement des nouvelles technologies, plus particulièrement des algorithmes et du matériel informatique, font de la santé un secteur de développement immense et extrêmement diversifié. Le croisement entre l'intelligence artificielle et la médecine permet de concevoir des systèmes

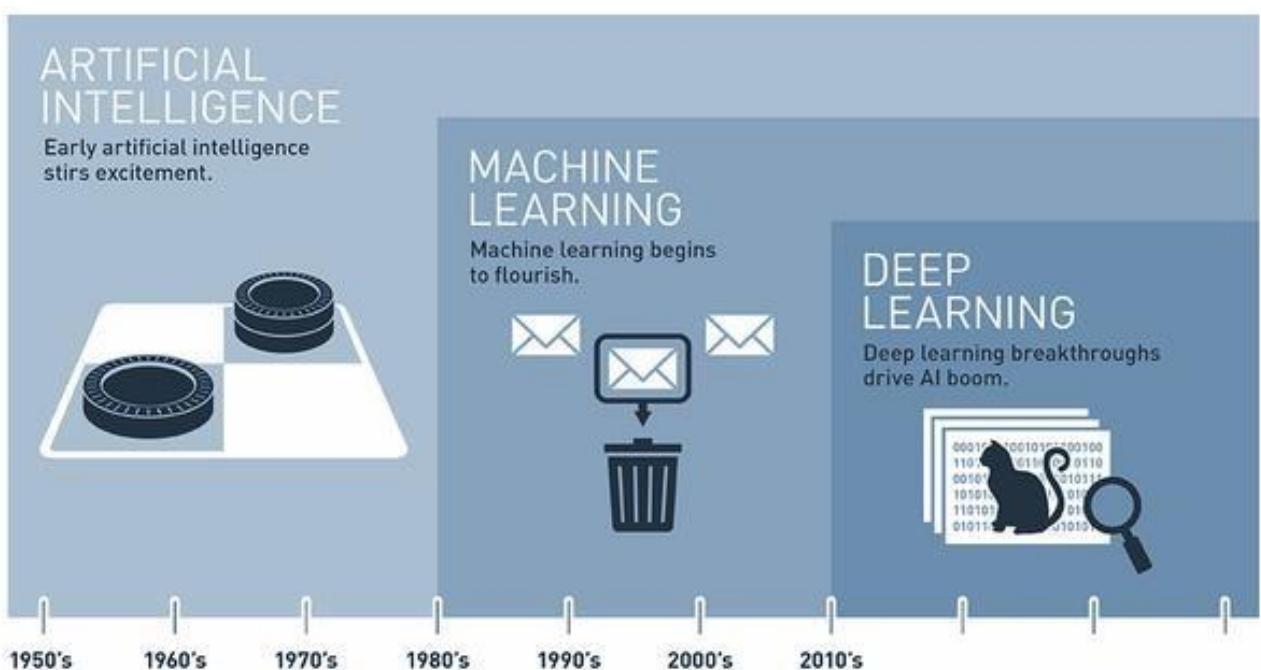
⁶ Deep Blue

plus préventifs et personnalisés, apportant une amélioration du suivi médical des patients et une assistance considérable aux professionnels de santé.

A.1. INTELLIGENCE ARTIFICIELLE POUR L'IMAGERIE MEDICALE

L'intelligence artificielle est devenue un acteur majeur de la médecine du futur en proposant des modèles toujours plus préventifs et personnalisés, offrant une amélioration constante de la qualité des soins et une assistance précieuse aux professionnels de santé. Ses applications ont été conçues et appliquées à des pratiques comme l'aide au diagnostic, les opérations assistées et les traitements personnalisés. L'imagerie médicale, et plus particulièrement l'aide au diagnostic, est le secteur de la santé où les recherches et les enjeux sont les plus importants. Le grand enthousiasme et le dynamisme du développement des systèmes d'IA en radiologie sont démontrés par l'augmentation des publications sur ce sujet. Il y a seulement 10 ans, le nombre total de publications sur l'IA en radiologie dépassait tout juste 100 par an. Par la suite, nous avons connu une augmentation considérable, avec plus de 700-800 publications par an en 2016-17. Au cours des deux dernières années, la tomographie assistée par ordinateur (CT) et l'imagerie par résonance magnétique (IRM) ont collectivement représenté plus de 50 % des articles, bien que la radiographie, la mammographie et l'échographie soient également représentées. La neuroradiologie (évaluée comme l'imagerie du système nerveux central) est la sous-spécialité la plus impliquée (représentant environ un tiers des articles), suivie de la radiologie musculo-squelettique, cardiovasculaire, mammaire, urogénitale, pulmonaire/thorax et abdominale, chacune représentant entre 6 et 9% du nombre total d'articles. L'IA a actuellement un impact sur le domaine de la radiologie, l'IRM et la neuroradiologie étant les principaux domaines d'innovation.

Figure 2 : une équivoque à levée : le Deep Learning, Le Machine Learning, L'intelligence Artificiel



Source : l'intelligence humaine associée à l'informatique cognitive (Cogito, 2019).

A.2. LE MACHINE LEARNING

L'apprentissage machine est une science qui consiste à faire en sorte qu'un ordinateur agisse sans le programmer. Le Deep Learning en est un sous-ensemble, qui peut être considéré comme l'automatisation de l'analyse prédictive. Il en existe trois types différents. Tout d'abord, l'apprentissage supervisé, où les ensembles de données sont étiquetés pour que des modèles soient détectés puis réutilisés. Puis, l'apprentissage non supervisé, où les ensembles de données ne sont pas étiquetés, mais sont triés en fonction des similarités ou des différences. Et enfin, l'apprentissage renforcé, où les ensembles de données ne sont pas étiquetés, mais l'IA reçoit un feedback de rétroaction après l'action (actualiteinformatique, 2023).

A.2.1. LES TYPES D'APPRENTISSAGE MACHINE

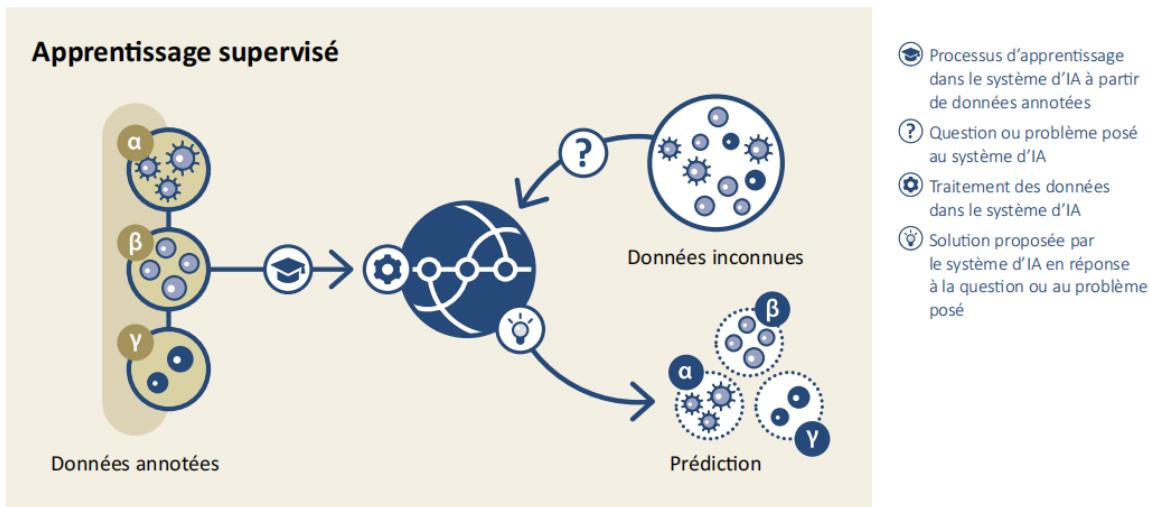
En apprentissage automatique, il existe deux types de données : les données étiquetées et les données non étiquetées. Si nous avons des données étiquetées, cela signifie que nos données sont marquées, ou annotées, pour montrer la cible, qui est la réponse souhaitée que notre modèle d'apprentissage automatique prédise. La nature de ces données identifie quatre catégories principales d'apprentissage utilisées aujourd'hui : l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage semi-supervisé et l'apprentissage par renforcement. Nous détaillerons ces catégories dans les paragraphes suivants.

APPRENTISSAGE SUPERVISE

L'apprentissage supervisé traite ou apprend avec des données étiquetées. Cela implique que les données doivent être déjà étiquetées avec les bonnes réponses. L'apprentissage supervisé peut être assimilé à un enseignant qui utilise ses connaissances pour enseigner et corriger les erreurs d'un apprenant. Cette analogie est utilisée par les algorithmes d'apprentissage supervisé pour apprendre à partir des données. Lorsque l'algorithme fait une prédiction sur un exemple, sa fiabilité peut être calculée en connaissant la réponse correcte qui correspond à l'exemple.

L'apprentissage supervisé englobe deux familles d'algorithmes : les algorithmes de classification lorsque la variable de sortie est une catégorie (chat ou chien par exemple) et les algorithmes de régression quand la variable de sortie est une valeur réelle (prix, poids, surface...).

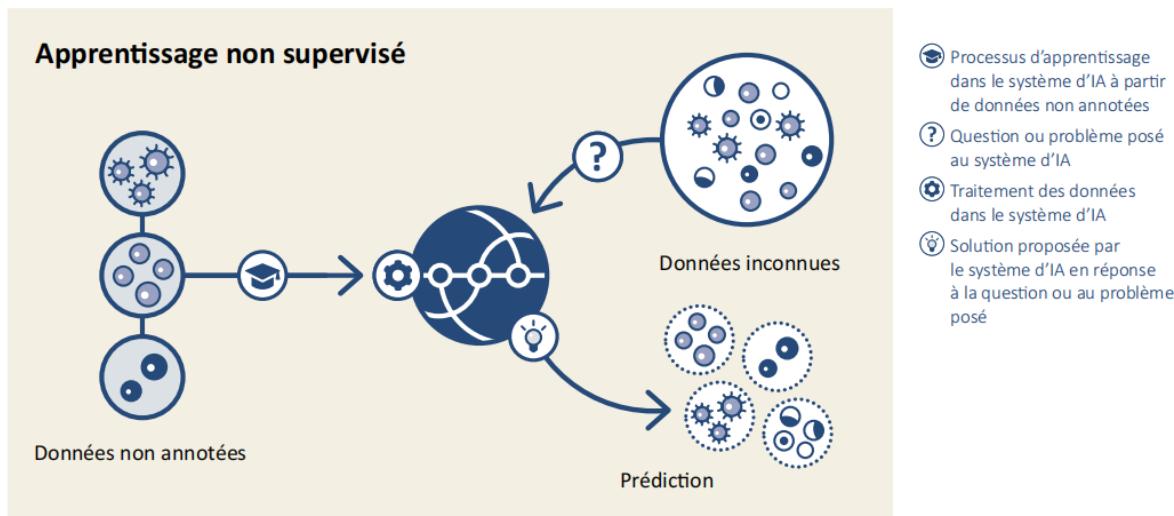
Figure 3 : modèle représentative : apprentissage superviser



APPRENTISSAGE NON SUPERVISE

L'apprentissage non supervisé consiste à entraîner un algorithme en utilisant des informations qui ne sont ni classées ni étiquetées et lui permettre d'agir sur ces informations sans aucun guidage. Dans ce cas, sa mission consiste à regrouper des informations non triées en fonction de similitudes, de caractéristiques et de différences. Contrairement à l'apprentissage supervisé, aucun enseignant n'est fourni, ce qui signifie qu'aucune instruction ne sera donnée. Par conséquent, l'algorithme doit trouver les relations implicites dans les données non étiquetées par lui-même. L'apprentissage non supervisé englobe deux familles d'algorithmes : les algorithmes de regroupement (clustering) et les algorithmes d'association. Les premiers consistent à découvrir les regroupements inhérents aux données (par exemple en regroupant les clients par leur comportement d'achat) alors que les derniers consistent à découvrir des règles qui décrivent une grande partie des données (par exemple, les personnes qui achètent X ont également tendance à acheter Y).

Figure 4 : modèle représentative : apprentissage non-supervisé



APPRENTISSAGE SEMI-SUPERVISE

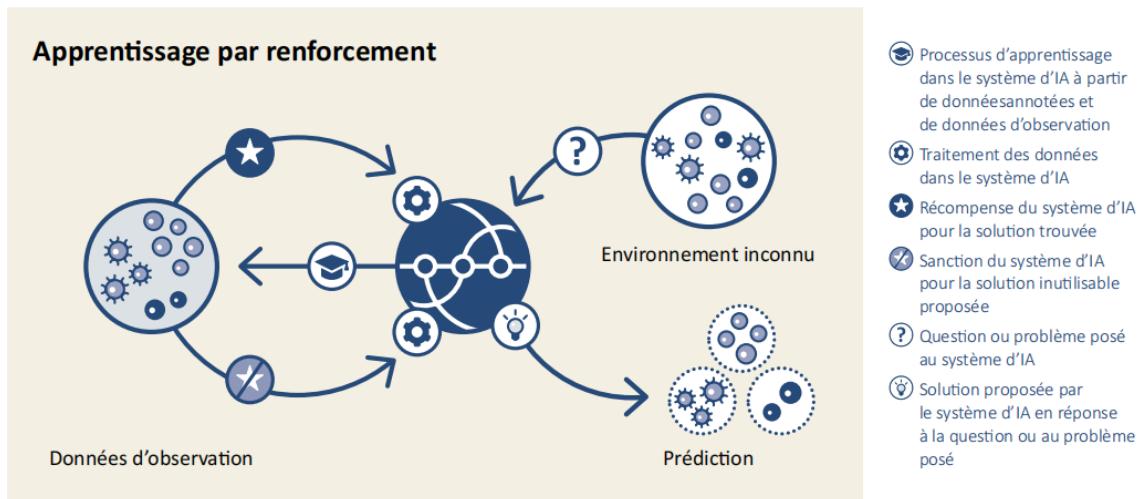
L'apprentissage semi-supervisé est un problème d'apprentissage qui implique un petit nombre d'exemples étiquetés et un grand nombre d'exemples non étiquetés. Les problèmes

d'apprentissage de ce type sont complexes car ni les algorithmes d'apprentissage supervisé ni les algorithmes d'apprentissage non supervisé ne sont capables d'utiliser efficacement les ensembles de données étiquetées et non étiquetées. L'apprentissage supervisé donne plus d'information, mais peut être très coûteux en quantité de données et temps de calcul alors que l'apprentissage non supervisé est plus automatisé, mais les résultats peuvent être beaucoup moins précis. L'apprentissage semi-supervisé combine quant à lui l'utilisation de données étiquetées et non étiquetées pour obtenir "le meilleur des deux" approches. Son principe est alors de modifier, ou de réorganiser, les hypothèses effectuées sur le modèle à partir des données d'apprentissage (ensemble en général petit) pour trouver la configuration la plus adaptée la fois aux données d'apprentissage et aux données de test. La plus grande utilité de cette approche est la capacité de labéliser les données non étiquetées en données étiquetées. Les données étiquetées sont comparativement plus coûteuses, difficiles et fastidieuses à acquérir que les données non étiquetées. Une autre utilité est de réduire le nombre de paramètres à apprendre pour prédire les données étiquetées à l'aide de tâches fictives.

APPRENTISSAGE PAR RENFORCEMENT

L'apprentissage par renforcement consiste à entraîner des modèles d'apprentissage automatique à prendre une séquence de décisions. Le modèle apprend à atteindre un objectif dans un environnement incertain et potentiellement complexe. Dans l'apprentissage par renforcement, il est confronté à une situation semblable à un jeu. Il procède par essais et erreurs pour trouver une solution au problème. Pour amener la machine à faire ce que le programmeur veut, le modèle reçoit des récompenses ou des pénalités pour les actions qu'il réalise. Son objectif est de maximiser la récompense totale. Bien que le développeur définisse la politique de récompense (c'est-à-dire les règles du jeu), il ne donne au modèle aucune indication ou suggestion sur la façon de résoudre le jeu. C'est au modèle de trouver comment accomplir la tâche pour maximiser la récompense, en commençant par des essais totalement aléatoires et en finissant par des tactiques sophistiquées et des compétences supérieures. L'apprentissage par renforcement a été appliqué avec succès à des problèmes variés, tels que le contrôle robotique, le pendule inversé, la planification de tâches, les télécommunications, le backgammon et les échecs.

Figure 5 : modèle représentatif : apprentissage par renforcement



Durant ce mémoire, nous avons fait face à des problématiques de nature segmentation et classification (segmentation des artères, segmentation d'anévrismes, détection des sténoses et classification des plaques). Ces tâches font partie de la discipline de l'apprentissage supervisé car nos modèles apprennent à partir d'un jeu de données annotées par des experts en neurologie et radiologie.

C'est dans ce cadre que nous concentrons les sous-sections suivantes sur l'étude de l'apprentissage supervisé. Nous détaillons sa formulation mathématique, ses principaux défis et les techniques utilisées pour assurer une bonne performance des modèles développés dans ce contexte.

A.2.2. LES MATHEMATIQUES AUTOUR DU MACHINE LEARNING

Pour formaliser ce problème, nous devons d'abord établir une modélisation mathématique de la tâche que nous voulons automatiser. Par conséquent, nous définissons un espace des hypothèses Φ dans lequel nous chercherons une solution. Nous devons également choisir une fonction de coût f qui évalue correctement les performances d'une solution $\varphi \in \Phi$ pour l'échantillon des données utilisé. Si le coût d'erreur de f est élevé alors les performances sont faibles, et vice versa. Soit D l'ensemble correspondante à toutes les données relatives à la tâche considérée, $d \in D$, alors le but ultime de l'apprentissage automatique est de trouver une fonction φ^* telle que :

Équation 1 : formalisation d'une fonction d'apprentissage

$$\varphi^* = \operatorname{argmin}_{\varphi} E_D [f(d, \varphi)]$$

En d'autres termes, nous cherchons à trouver une fonction φ^* qui minimise le coût de l'erreur sur l'ensemble des données. Ainsi, si nous avons accès à un ensemble infini de données, le problème d'apprentissage automatique se réduit au problème d'optimisation ci-dessus (en plus du problème de modélisation qui est le choix de Φ). En revanche, c'est rarement le cas et les données sont le plus souvent limitées à un ensemble fini D . Par conséquent, nous ne pouvons pas juger avec précision la fiabilité de la fonction choisie car cet ensemble D peut ne pas représenter la vraie distribution des données de la tâche en question. Dans ce cas, nous faisons face à un problème d'estimation où nous ne pouvons évaluer la performance du modèle que sur un ensemble limité et connu de données.

Cet apprentissage se fait alors dans un espace d'hypothèses Φ avec un ensemble de données fini D . En apprenant de cette manière, le système accumule deux erreurs :

- Une erreur d'approximation (biais) qui est l'erreur provenant de l'apprentissage sur un espace d'hypothèses différent de l'espace réel,
- Une erreur d'estimation (variance) qui est l'erreur résultant de la mauvaise représentation des données par l'hypothèse choisie. Idéalement, nous cherchons à minimiser ces deux erreurs ensemble (risque empirique). Malheureusement, il est généralement impossible de faire les deux en même temps. Ceci est appelé donc le dilemme biais-variance.

A.3. LE DEEP LEARNING

Le Deep Learning ou **apprentissage profond** consiste à ce qu'une intelligence artificielle parvienne à assimiler de nouvelles connaissances à travers un réseau de neurones artificiels. Il s'agit d'une pratique issue de l'apprentissage automatique, également connu

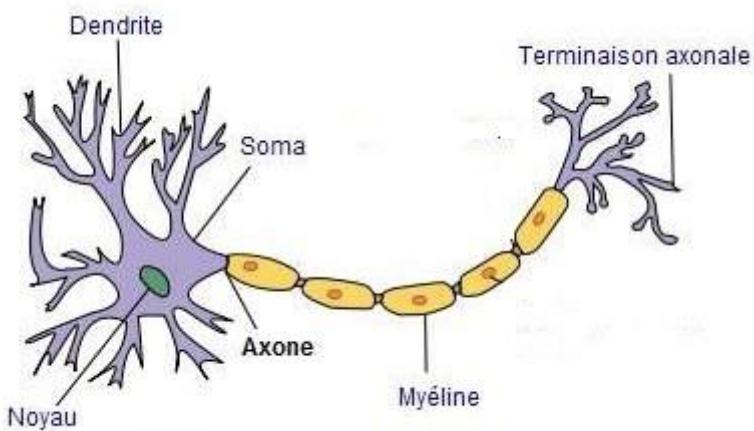
sous le terme de machine Learning. Cela permet à une intelligence artificielle (IA) de s'améliorer en intégrant de nouvelles règles. Leur ajout ne fait l'objet d'aucune intervention humaine. L'apprentissage profond utilise alors différentes couches neuronales qui forment un réseau artificiel.

L'apprentissage profond s'appuie sur les réseaux de neurones artificiels convolutifs, dits profonds. Originellement inspiré par des observations en neurosciences sur le fonctionnement du cerveau, ce type de modélisation existe depuis près de soixante ans, après l'invention du Perceptron (Rosenblatt, 1958). Dans un réseau biologique, les données d'entrée proviennent d'un autre neurone qui passe dans la dendrite, le corps cellulaire effectue les calculs et le résultat sort par l'axone (cf. figure 3). Le fonctionnement d'un neurone artificiel est modélisé sur celui d'un neurone biologique. Il peut réaliser des fonctions logiques, arithmétiques et symboliques complexes. Par contre, sa formulation initiale comportait d'importantes limitations (Marvin Minsky, 1969), notamment l'impossibilité de résoudre des problèmes non linéaires (surtout l'opération XOR). Il a donc été mis de côté jusqu'à l'invention des réseaux neuronaux multicouches artificiels et la rétro propagation (Williams, 1986).

A.3.1. LE NEURONE NATUREL

Le neurone biologique, qu'on peut nommer encore neurone naturel pour le distinguer du neurone artificiel des informaticiens, est une cellule spécialisée de l'organisme. C'est l'unité de base de notre cerveau, qui en contiendrait 86 milliards. Assemblés en réseau, les neurones donnent naissance à la pensée humaine et à ce que nous nommons intelligence. Mais ils ne sont pas les seules cellules du cerveau. D'autres types cellulaires jouent un rôle fondamental et participent au processus neuronal. Sans eux, les neurones ne pourraient exister. Mais seuls les neurones, ou plus exactement leurs propriétés électriques, ont inspiré les inventeurs du neurone artificiel. En effet, le neurone biologique a la particularité de produire, sous certaines conditions, un signal électrique et de le transmettre. C'est par ce signal électrique que se fait la diffusion du message nerveux (David, 2021).

Figure 6 : représentation d'un neurone naturel



A.3.2. LE NEURONE ARTIFICIEL

I. LE PERCEPTRONS

Le perceptron est une modélisation grossière du fonctionnement des neurones naturels. Cette modélisation est assez intéressante car ainsi le neurone se prête mieux au calcul (au traitement informatique). C'est en effet ce modèle qui définit un neurone artificiel (celui qui est implémenté sur les plateformes informatiques).

Figure 7 : Neurone Artificiel complet

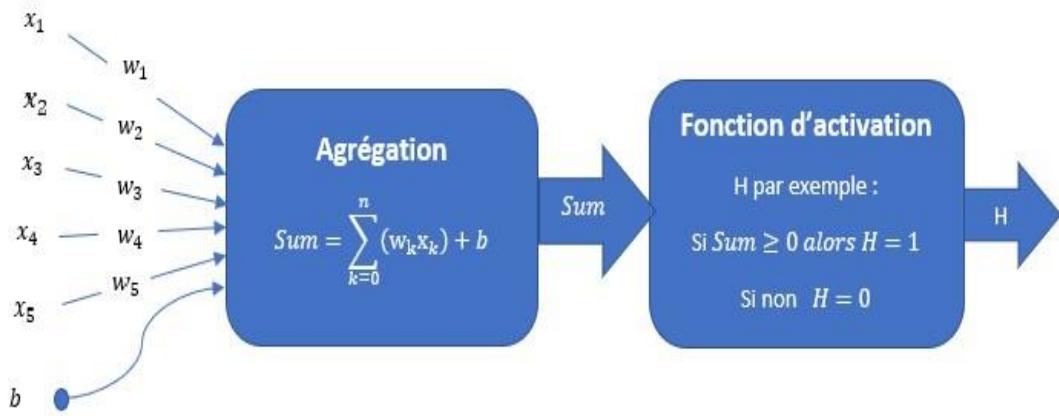
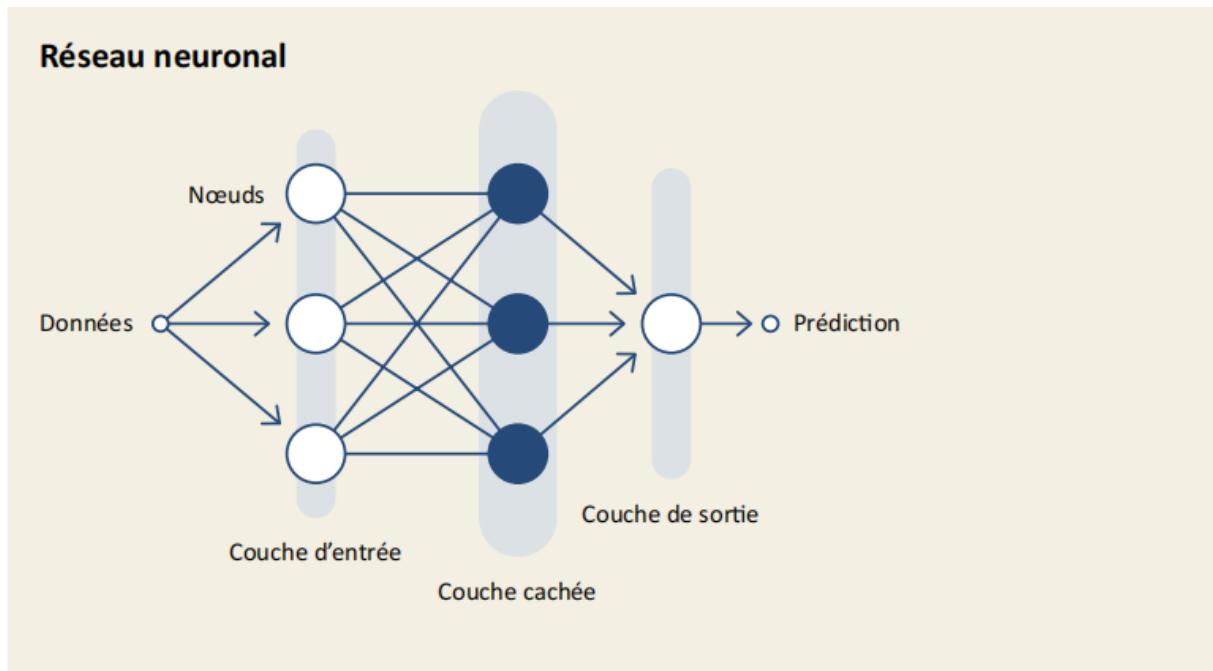


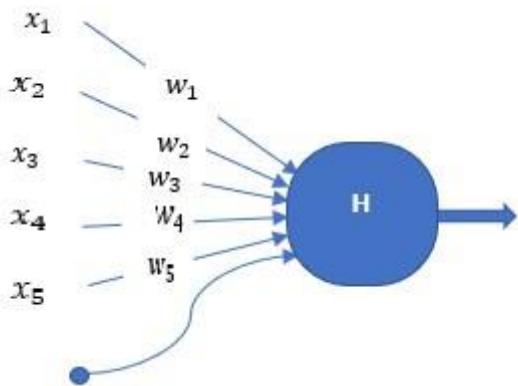
Figure 8 : représentation schématique d'un réseau de neurones



II. INFLUENCE DU BIAIS

Le biais du neurone est fondamentalement nécessaire. Pour le comprendre, essayons d'explorer plus profondément le fonctionnement du neurone présenté à la première figure 4. Convenons-nous de représenter un neurone et sa fonction d'activation plus simplement par un seul bloc dans lequel sera inscrit uniquement la fonction d'activation. Donc, désormais, un neurone sera de la forme:

Figure 9 : neurone complet

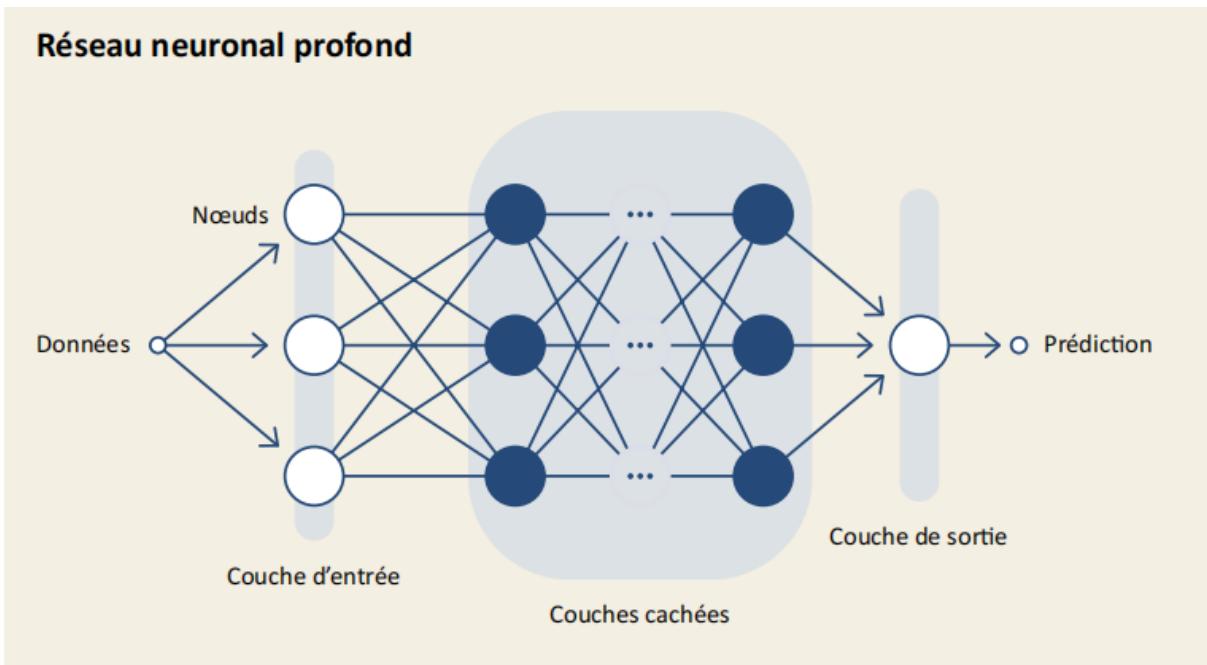


A.3.3. LE RESEAU DE NEURONES ARTIFICIEL : ANN

Le réseau de neurones artificiels (ANN pour Artificial neural network) repose sur la notion perceptrons multicouches (MLP pour Multilayer Perceptrons). Un réseau de perceptrons multicouche peut être vu comme un ensemble d'unités de traitement (neurones), reliées entre elles par des connexions pondérées. Les poids de ces connexions sont les paramètres du modèle. Ces neurones et ces connexions sont organisés par couches : La première couche est appelée couche d'entrée, la dernière couche est appelée couche de sortie et les couches intermédiaires sont appelées couches cachées. Les neurones de ces couches cachées, ainsi que ceux de la couche de sortie, appliquent deux traitements :

- Une combinaison linéaire de leurs entrées (dont les poids sont des paramètres du réseau),
- Une fonction linéaire ou non linéaire appelée fonction d'activation.

Figure 10 : réseau de neurones profond



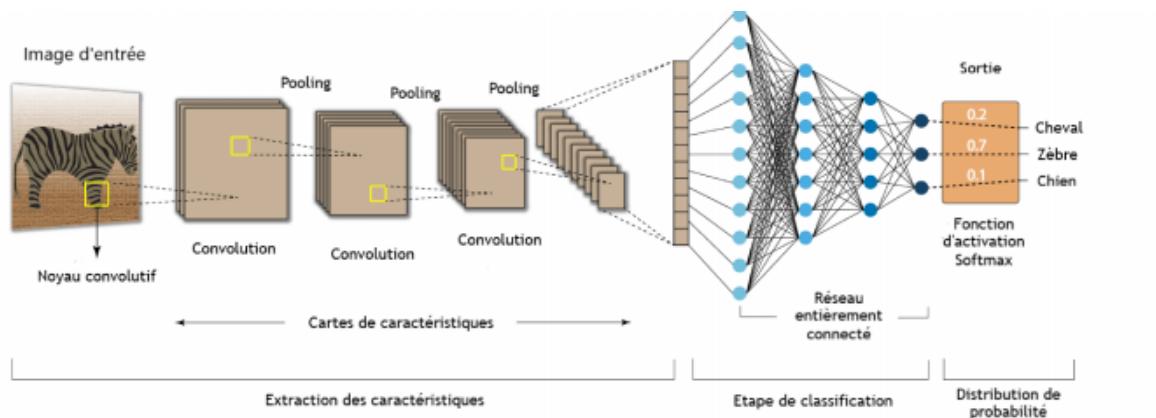
A.3.4. RESEAU DE NEURONES CONVOLUTIFS : CNN

Ce type de réseau s'inspire du fonctionnement biologique du cortex visuel. Les premières tentatives remontent à 1980 avec le néocognitron de **Fukushima** (Kunihiro, 1980), ancêtre des réseaux de neurones convolutifs. **Lecun** et al (Yann Lecun, 1998) se sont inspirés de ce dernier pour proposer le premier exemple de réseau de neurones convolutifs (appelés aussi CNN ou ConvNet) conçu dans le but de la reconnaissance de l'écriture manuscrite. Ce travail innovant a été appelé LeNet-5 et comprenait sept couches différentes avec deux couches de convolution.

Un réseau de neurones convolutifs est un modèle proche d'un réseau de perceptron multicouche. Il repose sur trois principes fondamentaux :

- Le champs réceptifs locaux associés à des convolutions permettent de détecter des caractéristiques élémentaires de l'image, formant ainsi une carte de caractéristiques ;
- Le partage de poids, qui consiste à apprendre les mêmes paramètre (poids) d'une convolution (et par conséquent à extraire les mêmes caractéristiques) pour toutes les positions sur l'image ;
- Les opérations de pooling réduisent la sensibilité aux translation et le coût du traitement de données.

Figure 11 : schéma complet d'un réseau de neurone convolutifs



II.3.1.B. LE BIG DATA ET L'INTELLIGENCE ARTIFICIEL

Le Big Data et l'intelligence artificielle sont deux technologies inextricablement liées, au point que l'on peut parler d'une Big Data Intelligence. L'IA est devenue omniprésente dans les entreprises de toutes les industries au sein desquelles la prise de décision est transformée par les machines intelligentes. Le besoin en matière de décisions plus intelligentes et de gestion du Big Data sont les critères qui dirigent cette tendance. La convergence entre le Big Data et l'IA semble inévitable à l'heure où l'automatisation des prises de décisions intelligentes se présente comme la prochaine évolution du Big Data. Une agilité en hausse, des processus business plus intelligentes et une meilleure productivité sont les bénéfices les plus probables de cette convergence.

L'évolution de la gestion des données ne s'est pas déroulée sans embûches. Beaucoup de données sont aujourd'hui stockées sur ordinateur, mais il reste encore de nombreuses informations sur papier, malgré la possibilité de numériser les informations sur papier et de les stocker sur des disques ou dans des bases de données (actualiteinformatique, 2023).

Aujourd'hui, l'usage du machine Learning, des systèmes experts et des technologies analytiques en combinaison avec le Big Data se présente comme l'évolution naturelle de ces deux disciplines. La convergence est inéluctable. L'Internet des Objets représente également une convergence entre Big Data et Intelligence artificielle. Sans un cerveau humain numérisé suffisamment intelligent pour permettre aux humains d'utiliser un réseau IoT capable de traiter, de distribuer et de collecter le Big Data, il ne sera pas possible de mettre en place un tel réseau. Même les capteurs, les puces, les nœuds de réseaux et les logiciels qui permettent de faire fonctionner les réseaux IoT sur le Cloud seront liés à l'intelligence artificielle. Ce phénomène est déjà en place dans le domaine des communications Machine to Machine. La capture de données pour identifier des tendances ou des patterns dans le comportement des clients ou des employés peut être très utile. Toutefois, l'extraction d'un sens, et son automatisation, pour découvrir des méthodes optimales d'améliorer la productivité ou la résolution de problèmes pourrait être encore plus utile. L'intelligence artificielle va être utilisée pour extraire du sens, déterminer de meilleurs résultats, et permettre des prises de décisions plus rapides à partir de sources Big Data massives (datascientest, 2023).

II.3.1.C. LA PRISE DE DÉCISION ET L'INTELLIGENCE ARTIFICIEL

Quant à la prise de décision, de quoi s'agit-il ? Ce concept est défini comme ceci : "La prise de décision est un processus cognitif complexe, différent de la réaction instinctive et immédiate, visant à la sélection d'un type d'action parmi différentes alternatives. Ce processus est théoriquement basé sur des critères de choix, et sur une analyse des enjeux et des options et conduit à un choix final." Dans cette définition, nous apprenons donc qu'une prise de décision est un processus qui met en place à la fois l'utilisation des connaissances apprises dans certains environnements pour choisir une action parmi tout un panel, et également l'apprentissage de part ces mêmes environnements qui permettra lors de prochaines prises de décision d'avoir une expérience plus grande afin de choisir l'action qui apportera le meilleur ratio entre le gain et les conséquences.

On peut également définir la prise de décision comme une réaction normale d'un organisme à rechercher comment réagir face à une situation, ou lorsque plusieurs choix s'offrent à ce dernier. De plus il n'est pas nécessaire d'apporter une réponse de vive voix, mais il est très important que la prise de décision se fasse dans un état conscient. Cependant ce n'est pas toujours par la réflexion et le passage vers notre soi conscient que la décision est prise. Certaines décisions sont prises par notre inconscient, les exemples donnés le plus fréquemment sont l'arrêt à un feu rouge lorsque nous sommes au volant d'un véhicule, ou le fait de s'éloigner d'un feu lorsqu'il commence à heurter notre corps. Nous ne réfléchissons pas dans ces cas-là à la décision que nous devons prendre, ainsi nous prenons des décisions sur le coup, pourtant les processus ont bien été mis en place et se sont déroulés de manière inconsciente. On peut donc distinguer deux types de prises de décision :

- Les prises de décision conscientes, où nous réfléchissons et où nous passons consciemment par toutes les étapes de la prise de décision ;
- Les prises de décision inconscientes, où nous ne nous rendons pas compte de ce processus mais qui se produit tout de même, et qui apportent une réponse ;

Si nous avions à imaginer ce que serait une prise de décision visuellement, le plus simple serait de s'imaginer un arbre de décision. Il s'agit là d'un outil qui aide à la prise de décision grâce à la représentation des différents scénarios qui sont proposés et des possibles

conséquences de chaque "branche" (chemin qui mène à une action et donc à une récompense). Ainsi, une prise de décision serait constituée essentiellement d'un point de départ, comme un problème à résoudre ou une situation quelconque, d'un ensemble d'actions possibles à chaque étape de la réaction ou de la résolution du problème, et à la fin de l'action entreprise une conséquence et potentiellement l'évaluation de la décision prise afin d'enrichir sa propre connaissance grâce aux processus cognitifs mis en place (Contribution du système d'information à la prise de décision :Cas des entreprises de Bejaia. YAHIAOUI Brahim).

Les résultats de ces définitions nous mènent à dire que la décision est caractérisée par :

- Processus composé d'un ensemble d'étapes.
- L'existence d'un problème qui nécessite une résolution.
- L'existence de plusieurs solutions possibles pour répondre à ce problème.
- Le choix de la solution doit être conscient et pris par un certain nombre de critères bien précis.
- La solution doit être satisfaisante par rapport au but fixé et La décision nécessite un temps limité.

II.3.2. LA SANTE NUMERIQUE ET LE CHATBOT

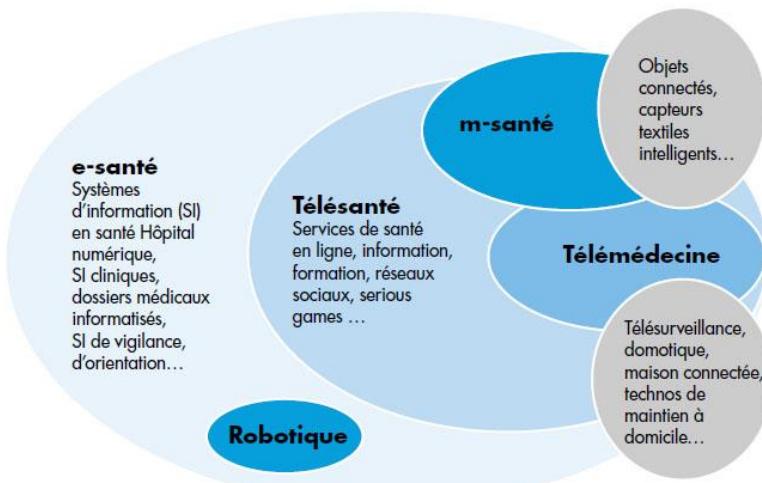
L'IA en santé pourrait être un outil performant et révolutionnaire. En effet, elle pourrait permettre de faire des diagnostics. Le but : gagner du temps et désengorger les cabinets de santé, les hôpitaux et tout autre centres médicaux. Pour cela, on imaginera un patient se trouvant face à une sorte de robot (Laura DI ROLLO, 2019). De chez lui ou dans un centre, il répondrait assez rapidement à un questionnaire lui permettant d'apporter une solution personnalisée en fonction des réponses qu'il aura cochées.

Selon de nouvelles orientations publiées récemment par l'OMS, l'intelligence artificielle (IA) constitue un grand espoir pour améliorer la prestation des soins et la médecine dans le monde entier, mais à condition de placer l'éthique et les droits humains au cœur de sa conception, de son déploiement et de son utilisation. « Comme toute nouvelle technologie, l'intelligence artificielle présente un immense potentiel pour améliorer la santé de millions de personnes dans le monde, mais comme pour toute technologie, il peut aussi en être fait mauvais usage et elle peut entraîner des effets préjudiciables, a déclaré le Dr Tedros Adhanom Ghebreyesus, Directeur général de l'OMS. Ce nouveau rapport, d'une importance capitale, constitue un guide précieux pour les pays qui souhaitent maximiser les avantages de l'IA, tout en minimisant les risques et en évitant les pièges. » (dataanalytics, 2023).

II.3.2.1. LA E-SANTE

L'Intelligence artificielle est utilisée dans plusieurs solutions proposées dans le domaine de la santé. Elle est utilisée pour accélérer le processus de diagnostic, avec un taux de précision qui pourrait dépasser celui de l'homme. De même, il devient plus facile de traiter des pathologies graves comme le cancer. L'IA peut aussi favoriser la conception de nouveaux médicaments et réduire le temps qui sépare la découverte d'une molécule de sa mise sur le marché. Par ailleurs, de nombreuses entreprises mettent sur le marché des applications mobiles destinées à suivre le traitement de certaines pathologies chez les patients (techno-science, 2023). Dans les pays asiatiques, des robots sont déjà expérimentés pour venir en aide aux personnes en difficulté dans leurs activités quotidiennes. Ce sera un véritable plus pour les personnes âgées par exemple.

Figure 12 : composante et structure de la e-santé



L'e-Health est un terme générique qui répond à des définitions à géométrie très variable, toutes en lien avec l'utilisation des technologies de l'information et de la communication (TIC) dans le domaine de la santé et/ou des soins de santé. Cela peut recouvrir des domaines tels que le télé monitoring, les téléconsultations, la télé-expertise ou la téléassistance, que l'on peut aussi regrouper sous l'appellation télémédecine. On inclut également ces services dans le domaine de la mHealth lorsqu'ils ont recours au support d'applications mobiles. La mHealth comprend également des applications de bien-être ou de style de vie, parfois connectées à des dispositifs médicaux ou à des capteurs (bracelets ou montres).

II.3.2. LES AGENTS CONVERSATIONNELS : CHATBOT

Historiquement, c'est le domaine de la psychologie qui a donné naissance au tout premier chatbot : ELIZA⁷, créé par le professeur **Joseph Weizenbaum** du MIT en 1966. Il s'agissait d'un programme informatique d'approche rogérienne (du nom du psychologue nord-américain Carl Rogers [1902-1987]) destiné à reformuler les propos de l'utilisateur sous forme de questions et ainsi simuler un entretien avec un thérapeute. Ensuite il a fallu attendre que la technologie se développe et devienne accessible au grand public pour se diffuser plus largement. Les chatbots de type assistant personnel de santé ou symptom checker ont commencé à voir le jour en Amérique du nord au début des années 2010. Mais **c'est surtout à partir de 2016 que les cas d'usage se sont multipliés** et que la technologie s'est diffusée dans plusieurs régions du monde, en particulier celles où le taux d'équipement en smartphone est le plus élevé.

La cinquantaine de solutions étudiées dans le cadre de notre étude permet aussi de dresser une typologie du phénomène. **Les chatbots visant à automatiser certaines tâches répétitives** telles que la prise de rendez-vous, la collecte d'informations auprès du patient ou l'établissement d'un pré-diagnostic ont pour vocation d'améliorer la performance médico-économique du système de santé. **D'autres solutions s'inscrivent davantage dans une logique de care.** Soit qu'il s'agisse de favoriser l'empowerment et contribuer à une meilleure expérience patient : accès facilité à une information fiable et pertinente, aide à l'éducation

⁷Eliza: le tout premier chatbot, a été créé par le professeur Joseph Weizenbaum du MIT en 1966,

thérapeutique, facilitation des démarches. Soit qu'elles permettent de rompre l'isolement dans le cas de maladies rares ou dans celui de troubles pour lesquels le regard de l'autre et la honte ressentie peuvent constituer un frein à la démarche de prise en charge.

Le marché des chatbots, et plus généralement de la e-santé, se développe à un rythme rapide et il faut s'attendre à une concurrence forte dans les mois et les années à venir. C'est pourquoi la question de l'usage doit plus que jamais demeurer au centre des préoccupations des porteurs de projet comme des concepteurs. Pour aboutir, le projet devra être réaliste, utile (c'est-à-dire satisfaire un besoin non adressé ou insuffisamment comblé), et apporter une valeur ajoutée par rapport à l'existant. Dans cette optique, **il est nécessaire de considérer la technologie non pas comme une fin en soi, mais comme un adjuvant au service de l'objectif à atteindre.**

II.3.2.A. LA TECHNOLOGIE SOUS-JACENTE

Dans la démarche de construction d'un chatbot, le rôle de la technologie est de mimer les mécanismes d'une conversation. Lorsque l'on décrypte les besoins conversationnels, il s'agit de comprendre une intention, et de proposer une réponse adaptée, et ce en temps réel. Enfin pour s'approcher davantage du mode conversationnel humain, il faut prendre en compte d'autres éléments : en particulier les paramètres de contexte, de mémorisation des précédents échanges et d'adaptation à l'interlocuteur.

Ces concepts peuvent être mimés de manière plus ou moins fidèle : pour les plus simples grâce à des outils développés spécifiquement, et pour les plus intuitifs grâce à de l'intelligence artificielle. Aujourd'hui, la maturité des solutions disponibles est variable et dépend de la complexité des performances attendues.

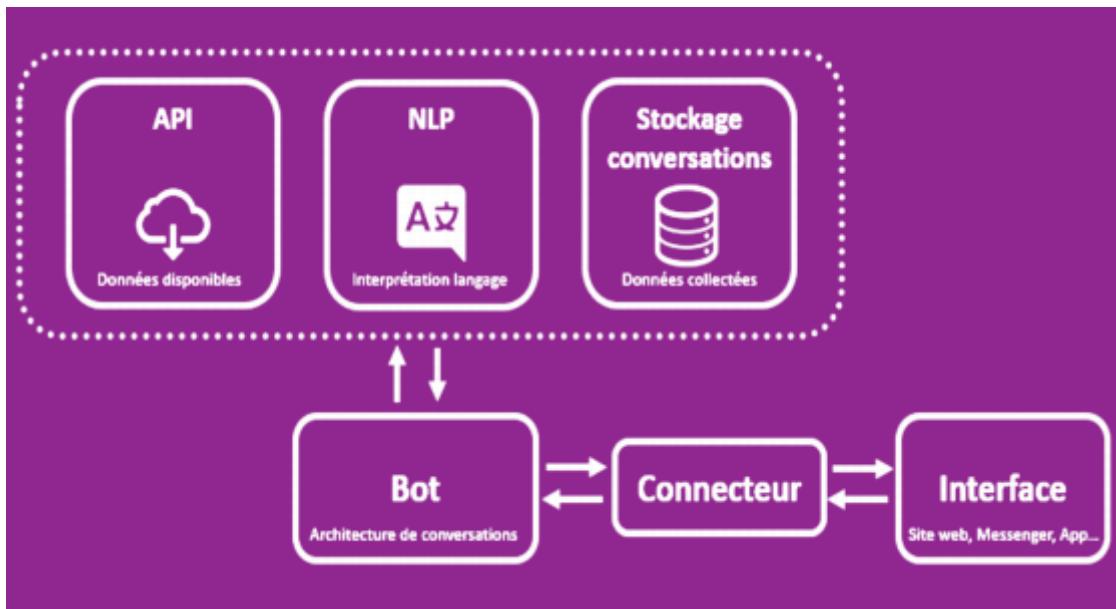
Des plateformes de construction de chatbot (par exemple Chatfuel ou ManyChat) permettent de répondre aux besoins conversationnels. Ce type de solution « sur l'étagère » est acceptable pour construire **une solution relativement simple**. Il est aussi possible de concevoir une solution propriétaire pour disposer d'un outil sur mesure, ou adapter une solution existante en la personnalisant.

L'intelligence artificielle, quant à elle, présente encore de nombreux champs de recherche. L'intérêt de ces technologies est de rendre **la conversation toujours plus fluide, ouverte et intuitive**. Les avancées sont rapides ce qui explique les améliorations des performances des chatbots. Les solutions développées par Google ou IBM sont parmi les plus puissantes du marché, mais des solutions propriétaires se développent en ciblant les conversations pour permettre des performances convenables sur un périmètre limité. En fonction des situations et des besoins, plusieurs approches peuvent être utilisées :

- **Le NLP** (Natural Language Processing) ou **NLU** (Natural Language Understanding), pour élaborer sa propre syntaxe autour d'un champ de langage dédié ;
- **Le deep-learning**, pour apprendre et automatiser la sélection des réponses à partir d'une large banque de données.

II.3.2.B. ORGANISATION D'UN OUTIL DE CHATBOT

Figure 13 : L'organisation technique d'un chatbot peut être représentée de la manière suivante :



- **Les API** servent à mettre à disposition des données ou du contenu validé auquel pourra avoir accès le bot. Les conversations seront ensuite stockées dans des bases de données. L'ajout d'une couche de NLP (Natural Language Processing) permet d'optimiser la compréhension du langage naturel.
- **Le bot** permet d'architecturer les conversations en fonction des questions reçues, en proposant des réponses adaptées.
- **Le connecteur** permet de faire la jonction avec l'interface choisie, qui peut être un site internet ou une application de messagerie.

Dans un contexte de santé, chaque brique technologique doit être considérée par rapport aux bénéfices et aux risques qu'elle peut faire courir. Par exemple : un chatbot avec de l'IA ouvre la conversation mais autorise des questions non contrôlées ; l'utilisation d'une messagerie tierce permet une augmentation de la visibilité mais expose à un partage des données avec un tiers ; préfère-t-on bénéficier des puissants moteurs des GAFAM ou a-t-on besoin de conserver les données dans un environnement sécurisé ? À chaque étape, une réflexion doit être menée en amont afin **d'identifier les options d'architecture utilisables**.

À retenir

Les barrières technologiques à l'adoption du chatbot ont en grande partie été levées. Certaines technologies matures sont disponibles, d'autres sont en évolution constante et rapide. Un accompagnement autour de la conception d'un chatbot permet de trouver des solutions techniques adaptées aux besoins de chaque projet.

Dans un contexte de chatbot dédié au domaine de la santé, le choix de l'approche technologique nécessite une attention particulière, d'une part **pour garantir la qualité du message transmis** à l'utilisateur, mais aussi pour assurer **la protection des données échangées**. L'objectif étant d'apporter des conversations à la fois **fluides, bienveillantes et en conformité avec les pratiques recommandées et la réglementation**.

II.3. CONCLUSION

Regardons les choses en face : avec la montée en puissance du numérique et le bouleversement du traitement de l'information (mode d'écriture, archivage, partage, communication, réseaux sociaux et, désormais, intelligence artificielle), nous sommes en train de vivre une véritable révolution technologique. Contrairement aux grandes transformations d'hier qui venaient modifier notre accès à la nourriture ou à l'énergie (ce fut le cas de la révolution agricole du néolithique, il y a dix mille ans, ou de la révolution industrielle, il y a cent cinquante ans), il s'agit aujourd'hui d'une révolution « intellectuelle », qui touche à l'immatériel et qui va modifier la manière dont nous utilisons notre cerveau, tout en bouleversant l'organisation de notre société dont les économies développées sont précisément centrées sur le tertiaire et donc sur le traitement (plus ou moins sophistiqué) d'informations.

On peut donc s'attendre à ce que cette nouvelle révolution transforme l'être humain, non pas sur le plan biologique, mais dans sa manière d'utiliser ses propres ressources.

Comme nous l'avons signifié tout au début que ce présent chapitre qui comprenait deux sections dont la première constituait dès la présentation des approches conceptuelles où nous avions présenté certaines définitions des différents mots clés et concepts de base liés de près ou de loin à notre sujet de recherche. À sa lecture certaines figures vous expliciteront mieux les aspects vaporeux de concepts technologiques.

CHAPITRE III. ETAT DE L'ART

III.1. INTRODUCTION

Certains problèmes connaissent une croissance exponentielle dans leur temps de traitement en fonction de la taille des exemples qui sont donnés. Il est possible qu'un problème ne puisse être résolu dans le temps à l'échelle humaine, et c'est pour cela que des techniques de traitement sont apparues pour palier à ce souci, comme le fait de pouvoir subdiviser un problème en sous-problèmes qui sont quant à eux résolubles.

De nos jours l'état de l'art dans l'univers de l'intelligence artificielle et celui des agents conversationnels comporte **le Transfer Learning**. Il s'agit d'une méthode d'apprentissage dans laquelle on suppose que la connaissance acquise par un modèle entraîné de Machine Learning peut être "transférée" pendant le processus d'entraînement de ce dernier. Cela permet de réduire la quantité de données nécessaires à un modèle pour apprendre une nouvelle tâche.

Dans le contexte actuel où l'on cherche à généraliser le plus possible les modèles afin qu'ils puissent traiter plusieurs types de problèmes différents, cela consiste en une étape clé vers ce but. Les évolutions dans le domaine de l'intelligence artificielle s'accompagnent également des évolutions dans le monde des hardwares : des composants plus puissants, qui permettent des traitements de plus en plus lourds en calculs et de réduire les temps de ce des derniers de façon optimale grâce à la parallélisations, sont à la source même de ces évolutions. Une plus grande parallélisations des traitements signifie un entraînement et une itération des modèles beaucoup plus rapides. C'est une frontière que l'on cherche constamment à repousser et qui chaque année nous offre son lot de surprises.

L'Intelligence Artificielle (IA) recouvre des communautés de recherche distinctes qui contribuent à divers titres aux succès récents du domaine. L'essor de l'IA moderne repose sur la disponibilité de grandes masses de données ainsi que la mise au point d'architectures de calcul et d'environnements logiciels spécialisés, et a bénéficié des avancées récentes en recherche fondamentale, notamment dans le domaine de l'apprentissage automatique. Au-delà des thématiques du numérique (vision par ordinateur, traitement du langage naturel, localisation et pilotage de systèmes autonomes, etc.), l'IA a des répercussions dans de nombreux autres secteurs (santé, humanités numériques, environnement, biologie, observation de l'univers et de la planète, etc....) (actualiteinformatique, 2023).

Ce dans cette marge de réflexion que nous avons inscrit la philosophie de cette recherche pas pour nier ni même pour bafouer les efforts des recherches déjà accomplis mais avec la vision de les étudier minutieusement, le prendre pour appuis et trouver notre apport scientifique.

Ce chapitre comme son intitulé l'insinue va présenter dans sa première section les travaux de certains chercheurs qui par leur dur labeur ont marqués et captés notre attention en travaillant sur une problématique similaire au notre. La deuxième section mettra les projecteurs sur nos objectifs de recherche en présentant la ligne de démarcation de ce présent travail de recherche pour donner à cette dernière sa scientificité.

III.2. SECTION PREMIERE : DE QUOI IL ETAIT QUESTION ?

Certes, le domaine d'intelligence artificielle (IA), l'e-santé, la réalité virtuelle et celui du Bigdata sont maintenant un aimant d'attraction pour plusieurs chercheurs vues premièrement les ouvertures qu'apportes ces nouvelles technologies et les budgets y allouent. La lecture de travaux antérieurs nous a permis d'avoir de larges connaissance sur le thème traité et aussi dans les domaines de recherche. Voici quelques travaux de recherche scientifique qui ont marqués nos esprits :

- **KAHUSI JOSEPH Jamal** a parlé sur « **the setting in place of a patient control and monitoring system in the room of a hospital** ». (Jamal, 2015) Ce chercheur a traité sur la notion du monitoring aux seins d'un hôpital en se focalisant sur la suivie de l'état vital du patient et le contrôle de ses mouvements pendant que le patient est interné. Il a aussi souligné que pour améliorer la vie économique de la population africaine, nous devrions faire de la santé et de son amélioration une priorité. Associant les connaissances informatique et électronique, il a pris comme objectif principal de concevoir un prototype grâce à Arduino et quelques capteurs pour la récolte les données de signe vital du patient. Une application desktop implémenté, conçu en java pour l'affichage graphique de ces données. Ce dans ce contexte de résultat de la recherche que le chercheur a fait des recommandations au ministère de la santé, aux soignants pour qu'il cherche comment améliorer la qualité des services à donner aux patients fréquentant les infrastructures sanitaires.
- **RONALD NACUA** qui a mené une étude sur : « **Conception et développement d'un système ambulatoire pour la mesure de l'activité du Système Nerveux Autonome pour la surveillance de personnes âgées.** ». (NACUA, 2009) Dans ce cadre, il montre qu'il est donc nécessaire de surveiller l'état de santé de personnes âgées à domicile. Une première solution consistait à équiper l'habitat de la personne avec des capteurs non-invasifs pour surveiller son activité quotidienne afin de détecter, par exemple, une éventuelle perte d'autonomie. Cependant, pour le suivi en continu de paramètres physiologiques ou la reconnaissance de motifs cinématiques, la portabilité des dispositifs devient une nécessité incontournable.

L'objectif principal de sa première expérimentation était de caractériser les réactions physiologiques lors de chutes simulées et de transitions debout-couché. A partir, des réactions physiologiques pour chaque situation, il avait implanté deux principales méthodes de classification (Réseaux de Neurones et SVM⁸) afin de classifier ces deux situations et fournir ainsi une information additionnelle pour la détection de chute. La deuxième expérimentation est consacrée à l'étude de l'activité du SNA⁹ chez des sujets âgés atteints de troubles orthostatique afin de détecter un dysfonctionnement à ce niveau. La population de l'étude est composée par 19 patients âgés ($85,5 \pm 6,9$). Sur l'ensemble des patients, 4 patients sur 19 présentent des problèmes d'hypotension orthostatique, vérifié pendant l'expérimentation à partir de la mesure de la pression artérielle en positions couché et debout. Au niveau de la fréquence cardiaque, il avait aussi remarqué qu'une quantité importante de sujets inclus dans son étude, présentaient divers problèmes de type cardiaque. Ainsi, à partir de la quantification de la variabilité de la fréquence cardiaque, nous n'avons pas pu conclure un dysfonctionnement du système sympathique. Au niveau de

⁸ SVM :

⁹ SNA :

la mesure de la résistance cutanée, il avait remarqué une augmentation significative de l'activité électrodermale chez les patients atteints de troubles orthostatique.

- **MOURAD AHMED Dendane** qui a consacré ces efforts sur : « **Conception et réalisation d'un capteur passif biomédical implantable et d'un lecteur assurant l'activation et la communication sans fil** » (Dendane, 2017). L'objectif de son projet consistait à concevoir un système biomédical inspiré par la RFID¹⁰, donc composé d'un lecteur actif et d'une plateforme passive sous-cutanée capable de mesurer un certain paramètre corporel, en l'occurrence la température interne. Puis, son système devait être apte à transférer les résultats sans fil à un appareil permettant l'analyse et l'affichage des données récoltées. L'implant serait injecté au niveau du poignet et le lecteur prendrait la forme d'un bracelet externe. Le premier objectif est donc l'alimentation par champ électromagnétique de l'implant passif.

Cette étape était nécessaire afin que le capteur puisse éventuellement transiger de l'information biomédicale sans fil avec un lecteur externe. L'alimentation d'un dispositif passif par induction électromagnétique soulève cependant plusieurs contraintes. D'abord, la géométrie et l'allure physique du lecteur RFID était des facteurs importants à considérer, puis une étude des implications en matière de puissance et l'énergie minimale pour l'activation d'un tag est aussi nécessaire.

Enfin, le principe de Faraday, le couplage électromagnétique et l'induction mutuelle étaient des aspects clés de l'accomplissement de cet objectif. Son second objectif était la conception d'une plateforme implantable d'acquisition de données capable de communiquer des mesures corporelles avec le bracelet externe. Le paramètre mesuré dans le cadre de son projet était la température interne du corps humain. À sa dernière vison le chercheur voulais concevoir une interface sans fil entre le bracelet et l'appareil mobile tel qu'un téléphone intelligent, afin de recueillir, d'afficher et d'analyser les données acquises. Le but est de mettre au point une application mobile intuitive permettant de relayer l'information biomédicale.

À son terme, le chercheur à recommander au futur chercheur de terminer la miniaturisation de l'implant et de développer une gamme de capteurs pouvant s'y intégrer de manière à récolter des données biomédicales auparavant inaccessibles. Pour ce faire, la mise en œuvre de technologies telles que les micro-électro-mécanisa (MEMS) ou le micro fluidique débloquerait indéniablement la possibilité de récolter plus facilement les données nécessaires à établir un diagnostic d'état de santé de base.

- **BOUAMRANE SOUAD FATIMA ZOHRA** jugea mieux de parler du « **Système d'Information Hospitalier : Admission et Planification des blocs opératoires** » (Zorha, 2019) dans sa recherche il précise que la planification doit permettre aussi bien une optimisation de l'utilisation des ressources, que la satisfaction d'une demande de soins de plus en plus exigeante. Il s'agit donc d'apporter aux systèmes de soins des gains significatifs en termes d'efficacité et de productivité tout en assurant qualité des soins et satisfaction des patients. Il démontre que le problème essentiel de toute planification des admissions réside dans les aspects aléatoires liés au milieu hospitalier avec occurrences de situations complètement imprévisibles. Disponibilité de toutes les ressources critiques intervenant dans ce processus.

¹⁰ RFID : Radio Frequen

L'objectif principal de son mémoire reposait sur la planification des admissions dans un système hospitalier et de répondre autant que possible à la demande de soins étant donnée la capacité finie des ressources tant humaines que matérielles.

L'approche proposée dans son travail était illustrée par la planification des admissions dans les cas de processus de soins passant par un service de chirurgie. En effet, ce type de processus représente une bonne partie des cas d'hospitalisation. Le coût que représentent le bloc opératoire et les attentes importantes qu'il génère justifie le choix de l'étude pour le chercheur. À cet effet, il présentait une méthode pour aider au suivi des admissions et à la construction d'un planning opératoire, et en comparant son application dans le cadre de différentes politiques de programmation opératoire. Les blocs opératoires représentent un univers composite dans lequel s'exprime une multitude de professions et de cultures différentes dont la finalité devrait être identique : les soins aux patients. La planification, l'ordonnancement et l'optimisation du fonctionnement des blocs opératoires est un problème vaste, complexe (caractère aléatoire du problème) et se trouve au conflit de plusieurs disciplines et techniques ayant comme objectif l'obtention d'un programme opératoire réalisable et efficace. De multiples contraintes telles que : L'emploi du temps des chirurgiens, leurs compétences spécifiques, le matériel médical spécialisé, la disponibilité des lits d'hospitalisation et les lits post opératoires, doivent être pris en compte dans les méthodologies de résolutions des gestionnaires pour assurer les meilleurs compromis entre une offre de soins compatible aux besoins des patients, et afin de lui garantir une prise en charge optimale.

III.3. SECTION DEUXIEME : LIGNE DE DEMARCACTION

En effet, nous ne sommes pas le premier à mener des recherches dans le domaine de l'intelligence artificielle ni même sur le système de contrôle et de surveillance des données des patients. Ce thème est orienté dans différents domaines d'application. Dans la plupart des ouvrages publiés, l'étude est orientée vers le contrôle et la surveillance des certains phénomènes liés aux biens et aux personnes comme la prévention des intrusions, des tentatives de vol, des incendies, protection des maisons de personnes vivant avec handicap, des attaques, la destruction de patrimoine international, etc. Néanmoins certains de ces études se sont également focalisés sur la surveillance et le contrôle l'évolution de l'état de santé, ces études portent principalement sur le contrôle et le suivi de l'état de santé des personnes âgées à domicile ou en maison de retraite ou encore sur le perfectionnement de méthode de prise en charge de malade.

Pour montrer l'originalité de notre travail de recherche nous devons trouver une ligne de démarcation pour chaque travail cité dans la première section de ce présent chapitre. Sur ce nous citerons de nouveau respectivement les auteurs, une synthèse de sa philosophie de travail et en suite nous préciserons ce dont il est question pour notre cas.

Pour **KAHUSI JOSEPH Jamal** (Jamal, 2015) ayant parlé d'une mise en place d'un système de télésurveillance de l'état vital du patient lorsque ce dernier est dans la chambre d'hôpital. Il a associé à son système de capteurs électronique pour la récupération de données. Ayant couvert un grand secteur vu les nombreuses solutions que son projet à apporter. Dans cette brillante recherche, effectuer minutieusement le chercheur a omis **de prendre en compte l'état du patient une fois dans son milieu habituel et la gestion des données des patients n'a pas intégré un aspect prospectif de son insertion dans la recherche clinique.**

En 2009, **RONALD NACUA** (NACUA, 2009) traitera sur la Conception et le développement d'un système ambulatoire pour la mesure de l'activité du Système Nerveux Autonome pour la surveillance de personnes âgées. C'est dans cette perspective qu'il a pris pour population d'étude les personnes âgées de plus de 55 ans. Il nous a montré que le nombre d'infrastructures dédiées pour la surveillance des personnes âgées ne sera donc pas suffisant et une de solutions qui permettra de soulager ces institutions est celle de surveiller ces personnes dans leur maison de retraite. Cependant, pour assurer un suivi plus approfondi de l'état de la personne (suivi de l'activité et de l'état de santé) il est indispensable de réaliser ce suivi avec des dispositifs portables sur la personne. Les travaux de cette thèse consistent au développement d'un dispositif portable pour le suivi de l'activité du système nerveux autonome de façon ambulatoire. La notion de l'intelligence artificielle fut évoquée dans cette thèse pour le développement d'un réseau de neurones pour l'interprétation et la prédiction de temps de chute ou de rechute de vieillard. **La prise en compte des données des personnes dont l'âge est inférieur à 55 ans dans son algorithme serait un atout majeur dans cette thèse car la perte de l'autonomie est devenue aujourd'hui une affaire de tout le monde.**

L'Algérien **MOURAD AHMED Dendane** (Dendane, 2017) qui a consacré ces efforts sur la Conception et la réalisation d'un capteur passif biomédical qu'on peut implanter et d'un lecteur assurant l'activation et la communication sans fil nous a démontré qu'en traitant sur la mise en application de principes d'ingénierie pour établir un pont entre la médecine et la biologie il a montré comment on peut accroître les possibilités dans le domaine de la santé. La prémissse de son mémoire est qu'il a prouvé la possible d'adapter la technologie d'identification par radiofréquence (RFID) pour activer et pour communiquer avec une plateforme implantable servant d'interface à une large variété de capteurs miniaturisés. Il a aussi souligné que l'implantation d'un capteur sous-cutané a pour avantage de fournir une variété de données autrement hors d'atteinte en empruntant une approche non-invasive. Avec trois objectifs cette recherche **montre un manque considérable dans la gestion des dossiers et différent document lier de près ou de loin au patient.**

BOUAMRANE SOUAD FATIMA ZOHRA (Zorha, 2019) développa un Système d'Information Hospitalier qui prenait en charge les Admissions et la Planification des blocs opératoires. Pour elle implanté un SIH reposant sur la planification des admissions et répondant autant que possible à la demande de soins étant donnée la capacité finie des ressources tant humaines que matérielles était la meilleure approche pour la numérisation de tache. Cette enrichissante recherche nous ayant aidé à élargir nos connaissances sur plusieurs notions liées à la planification et à la gestion des tâches dans un système hospitalière nous y avons néanmoins descelle quelque manquement. **Le fait de prendre uniquement le bloc opératoire comme milieu d'étude a largement handicapé ce travail car dans une machine toute les pièces travaillent en harmonie pour que chacun aide à la production d'un meilleur résultat. Numériser uniquement le bloc opératoire et les admissions cela sera à la base d'une perte d'équilibre considérable dans les autres blocs.**

Ayant pris l'audace de critiquer et d'apporter une vision actuelle aux travaux de nos prédécesseurs nous sommes assigne l'obligation de définir une ligne de démarcation qui pourra prouver l'originalité et la scientificité de ce présent travail. Aujourd'hui, nous avec **AMEN (Assistante Médicale Extra Numérique)** nous avons après la définition de notre problématique nous nous sommes assigne comme objectif et vison de nous appuyer sur dont il était question pour définir ce dont il est question.

L'association de capteurs portés et ambients permet de consolider le suivi en intérieur en proposant une analyse plus fiable de la détection de danger, mais également d'étendre ce suivi à l'extérieur. Ce dispositif permet de réaliser des mesures en ambulatoire, et donc en continu au cours d'une journée, pour une analyse plus complète et précise de l'acétimétrie. Le personnel soignant dispose ainsi de données plus complètes et plus fines pour le suivi et l'aide à la prise de décisions. Le bien être recommandent acétimétrie comme indicateur de base de l'état de santé mais demain des mesures physiologiques pourront s'inscrire dans la même logique de suivi ambulatoire. Notre action s'inscrit dans une dimension plus large qui touche au système de santé, au domaine de la télémédecine et de la télésurveillance qui tend à rapprocher l'hôpital et le domicile. Elle cherche à proposer un suivi médical à distance et des systèmes de détection de danger permettant aux personnes tant que distant du personnelle soignant de rester dans leur environnement de vie en toute sécurité.

Le bon fonctionnement et une harmonieuse coordination dans l'exécution des tâches dans les services journaliers des hôpitaux restent encore un défi majeur. Pour bien préciser objectivité de notre travail en nous basant sur les faits de notre problématique, notre contribution se situe dans cette logique de conception d'un écosystème intelligent qui permettra de résoudre les problèmes relative aux antécédent médicale, un système permettant aux médecin et infirmier de surveiller en temps réel certain signes vitaux des patients postopératoire ou ceux avec des pathologie à risque de perte d'autonomie et fonctions corporelles des patients et d'être alerte en cas de problèmes, un système pouvant faire sortir de statistique pour la compréhension de l'évolution de l'état du patient et utilisable pour la recherche médicale.

Toutefois signalons que les études précédemment cité ont un trait avec la nôtre. Elles ont traité d'une part de l'intelligence artificielle bien qu'à partie, du monitoring, du système de surveillance, de la télé-expertise, de la télésurveillance A l'heure actuelle, nous pouvons confirmer que notre travail se différencie des autres du fait nous nous basons sur l'amélioration de qualité de services d'urgences, celui du stockage des documents liés aux patients et du renforcement de l'interaction patient-médecin même si ce dernier et dans son milieu habituelle donc loin de l'hôpital.

III.4. CONCLUSION

Nous avons abordé dans ce chapitre, un état de l'art dans lequel nous nous sommes intéressés à donner un aperçu sur les systèmes d'informations, les systèmes d'information hospitaliers, l'application de télésurveillance et autres travaux scientifiques qui nous ont permis de nous imprégner dans le domaine hébergeant notre problématique.

Dans la littérature, beaucoup de chercheurs se concentrent sur les études concernant le problème de planification des interventions, seulement une petite partie de travaux a été effectuée sur le problème de l'ordonnancement des interventions dans les salles opératoires. De plus, un nombre encore plus restreint s'est intéressé à la planification suivie de l'ordonnancement dans le bloc opératoire, Néanmoins certains de ces études se sont également focalisés sur la surveillance et le contrôle l'évolution de l'état de santé, ces études portent principalement sur le contrôle et le suivi de l'état de santé des personnes âgées à domicile ou en maison de retraite ou encore sur le perfectionnement de méthode de prise en charge de malade.

Sur sa deuxième section nous avons explicité nos objectifs de recherche en nous basant sur les manquements de précédent travail. Toutes ces études précédentes que nous avons considérablement lu et étudie avec intérêt ont était pour nous une manière efficace de nous démarque sachant à l' avance de quoi il était question.

CONCLUSION DU PREMIER VOLET

La pluridisciplinarité de la recherche en intelligence artificielle et en système hospitalier se retrouve non seulement dans la diversité des applications présentées, qui rassemblent naturellement des chercheurs et de spécialistes des secteurs concernés, mais aussi dans les thématiques abordées, par exemple : robotique bio-inspirée, pour laquelle automaticiens et informaticiens collaborent avec mécaniciens, biologistes, chercheurs en sciences du comportement ; traitement automatique de la langue, qui mobilise linguistes, spécialistes de l'interaction, et chercheurs en algorithmique probabiliste ; Web sémantique, pour lequel les chercheurs en représentation des connaissances travaillent avec des sociologues pour les aspects réseaux sociaux. Cette pluridisciplinarité est une grande richesse, porteuse de ruptures scientifiques et technologiques, et donc de valeur ajoutée pour la société et l'économie.

Dans le second chapitre de ce volet nous avons présenté les concepts clés de cette recherche. Nous avons élargi ces contretypes dans sa deuxième section en parlant de : monitoring, de l'e-santé, du healthCare, du système informatique hospitalière, de l'intelligence artificiel, du deep Learning, du machines Learning, etc.

Le troisième chapitre s'est occupé de la présentation des travaux précédent qui ont d'un côté inspiré notre recherche. Après avoir présenté la synthèse philosophique ce recherche scientifique et leurs objectivités nous avons fait une critique constructive. Une ligne de démarcation a mis un point à ce chapitre.

Cette large littérature est très primordiale car il vient de poser une large et dure fondation pour ce qui suivra. Nous pensons avoir largement éclairé les domaines hébergeant notre sujet de recherche, si ce n'est pas le cas la bibliographie regorgent une liste détaillé de notre documentation pour vous enrichir.

Le volet suivant traitera maintenant de l'analyse de l'objet d'étude pour montrer concrètement les manquements des systèmes existant.

VOLET D'ANALYSE ET DE TECHNICITE

Comme décrit auparavant, la croissance des données dans le domaine de santé explose, et les cliniciens ont accès à une grande quantité de données. En effet, 153 exaoctets de données de santé ont été produits en 2013, et 2314 exaoctets produits en 2020 (Statista, 2023). La gestion et le traitement des données générées sont difficiles à réaliser de manière efficace et exhaustive.

L'augmentation constante du volume de données sur les soins de santé, créées par diverses sources et technologies (Shilo, 2020), fait qu'il y a plus de données collectées qu'il n'est possible d'analyser et d'utiliser dans les soins de santé. Une solution pour exploiter la diversité des données sur la santé consiste à utiliser le potentiel de l'IA. En bref, l'IA est un ensemble de concepts et de technologies qui permettent aux machines et aux ordinateurs d'imiter l'intelligence humaine. Selon la définition d'Oracle (oracle, 2023), c'est la capacité d'un système à interpréter correctement des données externes, à apprendre à partir de ces données et à utiliser ces apprentissages pour atteindre des objectifs et des tâches spécifiques grâce à une adaptation flexible. L'IA recouvre un large éventail de technologies, de cadres et de sous-ensembles présentant des possibilités d'application et des avantages différents.

Dans ce volet, nous allons nous atteler à la présentation des approches algorithmiques et modèle déjà proposé comme apport contributif dans le domaine de agents conversationnels médicaux qui est devenu un facteur clé de l'automédication, dans le domaine de l'authentification biométrique par l'iris (l'une des modalités les plus précises et difficiles à pirater), le visage (l'une des modalités les moins intrusives et les moins coûteuses) et celui d'assurer une prise en charge appropriée des patients victimes d'un accident vasculaire cérébral (AVC) et ensuite l'objet d'étude en montrant le manque de nos SIH.

CHAPITRE IV. APPROCHES, ANALYSES ET TECHNICITES

IV.1. INTRODUCTION

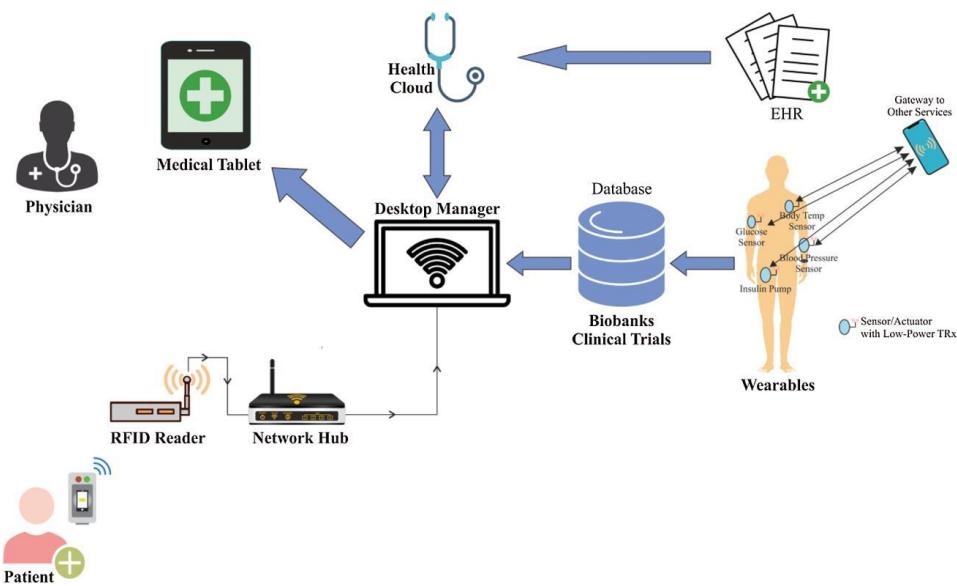
Au niveau de la littérature ou nous nous sommes fixé comme objectif de présenter les définitions des concepts clés de notre recherche et une approche thématique chose faite, nous avons ensuite attaqué l'état de l'art ou nous avons pu trouver la ligne de démarcation pour ce présent travail. En espérant que nous avons pu étancher la soif de curiosité de nos lecteurs nous voici à présent au point où nous allons analyser et étudier l'objet de notre recherche scientifique tout en présentant en avale les approches algorithmiques associer à notre problématique.

Récemment, l'IoT¹¹ est devenu plus productif dans les systèmes de santé (Shilo, 2020). Plus précisément, l'IoT dans le domaine de santé associe des capteurs, des microcontrôleurs, etc., afin d'analyser et d'envoyer les données des capteurs vers le cloud, puis vers le personnel soignant (médecins). L'intégration des fonctionnalités de l'IoT dans les dispositifs médicaux améliore la qualité de service des soins pour les patients âgés et pour les enfants. L'IoT dans les soins de santé pourrait conserver des milliers de données sur les patients qui sont informatisées et aider les patients à saisir leurs données à tout moment. De nombreux composants de détection de la santé ont été développés, la plupart d'entre eux sont portables, ce qui permet aux patients de les porter pour les surveiller. Le dispositif de surveillance de la santé est connecté au patient de sorte que le médecin puisse observer et suivre l'état du patient à tout moment.

Comme les patients assistés par des objets connectés sont accessibles par la technologie sans fil, l'état de santé du patient peut être détecté au bon moment afin de prendre les mesures appropriées (Statista, 2023). D'autre part, la plupart des pays en développement disposent d'infrastructures de soins de santé très médiocres. Si le dispositif de détection de la santé est conçu pour communiquer avec des appareils portables tels que des téléphones intelligents, des tablettes, etc., la communication avec le Cloud est possible. Les gens ont accès à ces appareils de communication portables qui deviennent maintenant la bonne solution. L'industrie de santé a rendu les soins aux patients plus fiables. Les informations en temps réel sur les données du patient sont analysées et enregistrées, et les médecins/soignants peuvent les surveiller à l'aide d'un smartphone, comme le montre la figure ci-dessous.

Figure 14 : internet des objets connectée lier à la e-santé

¹¹ IoT : internet of Things



Les données jouent un rôle important dans le diagnostic, le traitement et la prise en charge des patients. Les cliniciens se fient à leur capacité d'évaluer l'état de santé d'un patient sur la base des données et des informations disponibles. La disponibilité et la production de données dans le domaine de santé ont évolué de manière significative en raison de la numérisation de ce secteur voilà pourquoi dans la partie suivante nous allons directement voir les approches algorithmiques d'intelligence artificielle usuelle, des modèles d'analyse et traitement de données associé à la e-santé et une liste de réalisations des agents conversationnel existant dans le domaine clinique.

IV.2. SECTION PREMIERE : APPROCHES TECHNIQUES

La numérisation et les progrès technologiques actuels du secteur de santé ont créé une grande quantité de données cliniques. La création de données ne s'applique pas seulement aux hôpitaux et aux autres prestataires de soins de santé, mais des données exogènes sont également générées à partir de sources telles que les systèmes de surveillance à domicile, les capteurs personnalisés des smartphones et des technologies portables, et d'autres sources de données fiables externes (Blonde, 2018). La disponibilité accrue des données sur la santé permet, dans une plus large mesure, d'adopter surtout une approche holistique pour comprendre la santé des patients. L'évaluation médicale de l'état de santé d'un patient est de préférence considérée comme un processus multifactoriel, plutôt que comme une détermination de la maladie sur la base d'un symptôme spécifique qui, en d'autres termes, renvoie à un processus complexe très perturbé. A titre d'exemple, les données multi-ohmiques désignent le processus d'utilisation de types de données biologiques tels que la génomique, l'épigénétique et la protéinique dans l'analyse de maladies complexes (Subramanian, 2020).

IV.2.1. MAPPING DES CHATBOTS EXISTANT

L'une des fonctions premières du chatbot est **de répondre aux questions les plus fréquentes sur un sujet donné**, c'est pourquoi il suscite un fort intérêt dans le milieu médical. Il peut prendre en charge les interrogations des patients qui aspirent de plus en plus à devenir acteurs et responsables de leur santé. D'autant que les traitements n'ont de cesse de se

complexifier, et que les professionnels de santé ne trouvent plus le temps d'expliquer et de dialoguer avec les patients.

L'assistant conversationnel peut permettre de renforcer le lien patient-professionnels de santé, suscitant ainsi de grands espoirs pour l'efficience du suivi. En effet, la continuité et la qualité de la relation entre le patient et le professionnel de santé, même à distance, constituent un élément majeur de la réussite du traitement, en particulier pour les pathologies chroniques. En amont, ils sont même utilisés dans le cadre d'études cliniques pour gagner du temps sur le recrutement des patients et sur la compréhension de la partie législative du consentement.

Ainsi, c'est tout le parcours de soin du patient qui pourrait, à terme, être impacté par cette interaction entre l'homme et le robot : non seulement pour le bien-être et la prévention, mais aussi pour le curatif et l'accompagnement dans le traitement. Le chatbot, assistant conversationnel doté **d'une compréhension du langage naturel**, apprenant, capable à terme de détecter des émotions, et **disponible 24h/24, 7j/7**, devrait par nature répondre à l'ensemble de ces besoins, bien que le caractère récent de cette technologie ne permette ni encore de l'affirmer avec certitude, ni encore moins de le prouver scientifiquement.

IV.2.1.A. CHATBOTS DELIVRANT UNE AIDE PONCTUELLE

Parce qu'il s'agit de questions intimes que l'on peut hésiter à partager avec un professionnel de santé, il existe des chatbots traitant de sexualité (voir **Roo**¹², uniquement accessible sur smartphone et dédié aux adolescents et jeunes adultes), de gynécologie, de contraception ou de cycle menstruel (voir **Léa**¹³, le chatbot sans tabou de la marque de tampons Nett).

Dans le cadre pré ou post-opératoire, des solutions proposent d'automatiser le suivi des patients à domicile afin de dégager du temps soignant auprès des malades. Citons deux cas d'usage français : **Citizen Doc**¹⁴ et **MemoQuest**.

Le second, développé en 2017 par l'AP-HP en partenariat avec **Calmedica**¹⁵, propose le suivi patient par intelligence artificielle. Les chatbots offrent de nouvelles perspectives dans le domaine des essais cliniques. La recherche d'essais cliniques en cours est par exemple l'une des fonctionnalités proposées par **Vik Sein**¹⁶. À plus grande échelle, Microsoft a annoncé en mars 2019 avoir développé un Clinical Trials Bot à destination des laboratoires pharmaceutiques afin de faciliter le recrutement de patients sous forme de chat.

IV.2.1.B. CHATBOTS D'ACCOMPAGNEMENT DANS LA DUREE

¹² <https://roo.plannedparenthood.org/onboarding/intro>

¹³ <https://www.facebook.com/nett/>

¹⁴ <https://www.citizendoc.fr/>

¹⁵ <https://www.memoquest.com/>

¹⁶ <https://www.facebook.com/HelloVikSein/>

Le plus souvent accessible via son smartphone, et donc à portée de main 24/24h, le chatbot est assez naturellement susceptible de devenir un outil d'accompagnement au long cours.

Chat Yourself¹⁷ fournit une aide aux personnes qui souffrent de troubles cognitifs, tels que la maladie d'Alzheimer. L'outil est capable de mémoriser et envoyer sur demande une multitude de détails sur la vie quotidienne et pratique d'une personne telles que ses coordonnées ou les allergies dont il soufre.

Comme évoqué plus haut, les chatbots adressent également des maladies chroniques telles que le diabète, les maladies cardiaques ou encore l'asthme, auquel **WeFight¹⁸** vient de consacrer une nouvelle déclinaison de son chatbot Vik : **Vik Asthme¹⁹**

Enfin, nombreuses sont les solutions qui proposent de coacher notre activité physique ou notre alimentation dans une optique de bien-être et de prévention primaire. La Fondation Ramsay Générale de Santé propose par exemple **Ramsay²⁰**, le chatbot prévention santé spécialisé dans trois thématiques : le tabac, la nutrition et le stress. Les mutuelles de santé développent des solutions similaires à l'image de la **MGEN²¹** et son chatbot « JAM » dédié aux jeunes assurés.

IV.2.1.C. CHATBOTS GENERALISTES

Au nombre des chatbots généralistes, on peut citer :

Les outils vérificateurs de symptômes (ou symptom checker) tels que **Buoy²²**, **Your.MD²³**, **Gyant²⁴**, **SENSELY²⁵** ou **INFERMEDICA²⁶** sont très nombreux et principalement proposés en Amérique du Nord et au Royaume-Uni. Ce type de chatbot pose au patient une série de questions qui, au fur et à mesure qu'elles se précisent, permettent de dresser une liste de causes ou d'affections possibles. Cette étape de pré-consultation peut précéder la localisation d'un praticien, voire une téléconsultation ou, dans le cas d'une pathologie bénigne, proposer des options de self-care.

Les assistants personnels de santé comme **Florence Chat²⁷** aident les patients à gérer leur traitement sous forme de rappels et d'alarmes et les encouragent à respecter leur protocole thérapeutique. **Mabu²⁸**, le compagnon de santé pour les seniors développés par IDEO et **Catalia Health²⁹**, recueille également des données sur les progrès du patient et peut alerter son équipe médicale en cas de besoin.

Le chatbot de triage est programmé pour délivrer des conseils médicaux dans le but d'optimiser la charge d'un service d'admission aux urgences ou d'assistance téléphonique.

¹⁷ <https://www.facebook.com/chatyourself/>

¹⁸ <https://www.wefight.co/>

¹⁹ <https://www.facebook.com/VikAsthme/>

²⁰ <https://ramsaygds.fr/ramsay-chatbot-prevention-sante>

²¹ <https://www.facebook.com/MGEN.et.vous/>

²² <https://www.buoyhealth.com/>

²³ <https://www.your.md/>

²⁴ <https://gyant.com/>

²⁵ <http://www.sensely.com/>

²⁶ <https://infermedica.com/>

²⁷ <https://florence.chat/>

²⁸ <http://www.cataliahealth.com/>

²⁹ <https://www.babylonhealth.com/>

C'est le cas d'un test mené par le National Health Service (système de santé publique du Royaume Uni) dans le nord de Londres sur un bassin de plus d'un million de personnes. En partenariat avec **Babylon Health**³⁰, ce pilote a pour objectif affiché de réduire la charge du service d'assistance téléphonique du NHS, le 111. Censé durer six mois à compter de la fin janvier 2018, le retour d'expérience de ce test n'est, à notre connaissance, toujours pas divulgué.

IV.2.1.D CHATBOTS DE SPECIALITE

Contrairement aux chatbots généralistes, les chatbots de spécialité s'intéressent à un type de pathologie ou une situation particulière. Les chatbots relatifs à la santé mentale sont assez nombreux. Selon l'OMS, le coût induit par la dépression se chiffre à des centaines de milliards de dollars avec plus de 300 millions de personnes concernées chaque année (OMS, 2023).

De plus, nombre de personnes atteintes de problèmes de santé mentale n'ont pas accès aux traitements ou quand elles peuvent y accéder, hésitent à consulter en raison de la stigmatisation entourant la maladie. C'est pourquoi se développent les outils visant à offrir aux patients souffrant de troubles mentaux un soutien de première ligne, bienveillant et confidentiel.

De ce que nous avons pu observer, il apparaît que la plupart des solutions rencontrées telles **Woebot**³¹, **Wysa**³², **Elizzbot**³³ ou la française **Owlie**³⁴, sont basées sur des techniques de thérapie cognitive et comportementale ou sur du « coaching psychologique ». Les chatbots de suivi de maladies chroniques sont également assez fréquents. Ils s'inscrivent dans une double logique d'observance et de monitoring. **Diabetio**³⁵ est un outil de surveillance de la glycémie à destination des patients diabétiques disponible en deux versions, adultes et enfants. **Cardiocube**³⁶ est un assistant vocal qui aide le patient à gérer sa maladie cardiaque chronique, et communique en temps réel avec l'hôpital ou la clinique dont il dépend afin de faciliter et accélérer la prise de décisions en cas de besoin.

L'oncologie est un domaine dans lequel les chatbots offrent des perspectives intéressantes en raison notamment de l'accumulation de données de vie réelle et de la nécessité de soutenir les patients comme les aidants tout au long du traitement. Le nord-américain **Cancer Chatbot**³⁷ traite le cancer de manière globale quand la startup française **Wefifight**³⁸ et sa solution **Vik Sein**³⁹ se consacre à la seule thématique du cancer du sein

COMPORTEMENTS DE SANTE ADRESSES PAR LES CHATBOTS

Utilisés dans le cadre de dépistage, de prévention, d'observance, voire de cessation (de pratique à risque ou addictive par exemple), les chatbots sont susceptibles de traiter tout type de comportements de santé. Pour déterminer le bon diagnostic le plus tôt possible, permettre une prise en charge anticipée et identifier le meilleur traitement, l'**observation des**

³⁰ <https://www.babylonhealth.com/>

³¹ <https://woebot.io/>

³² <https://www.wysa.io/>

³³ <https://www.facebook.com/Elizzbot/>

³⁴ <https://www.facebook.com/owliechatbot/>

³⁵ <https://www.diabetio.com/>

³⁶ <https://www.cardiocube.com/>

³⁷ <https://masterofcode.com/portfolio/cancerchatbot>

³⁸ <https://www.wefifight.co/>

³⁹ <https://www.facebook.com/HelloVikSein/>

symptômes à des fins de détection de la maladie est cruciale. Favoriser cette observation et ainsi éviter au patient une « perte de chance » s'avère particulièrement utile dans les cas de pathologies graves, de troubles mentaux ou encore de risque de suicide.

La presse a par exemple fait état de la mise en service d'un **chatbot de dépistage des troubles de l'autisme** (esechos.fr, 2023). Utilisé par le personnel des crèches pour détecter les symptômes de l'autisme le plus tôt possible, soit dès 18 mois, il doit permettre une prise en charge plus précoce, favorable à un meilleur développement de l'enfant. Des chercheurs ont également commencé à développer des chatbots capables de réaliser un entretien clinique avec des interactions empathiques afin **de dépister des troubles dépressifs chez un patient**. Des expériences similaires sont en cours dans le but de déceler des problèmes d'addiction à l'alcool ou au tabac chez les patients.

En phase avec la Mission interministérielle de lutte contre les drogues et les conduites addictives qui dévoilait mardi 14 mai 2019 un rapport préconisant l'utilisation de la e-santé pour soigner les addictions, le chatbot peut être utilisé dans la prévention, le repérage, le diagnostic et la prise en charge des addictions. Dans le cadre **de cessation de comportements à risque et d'addictions** par exemple, il peut être utile à envoyer des messages d'encouragement dans le cadre de thérapies comportementales et motivationnelles, ou encore délivrer des informations d'orientation et de conseils.

Le chatbot peut également servir des objectifs de **prévention primaire et secondaire** : rappels de vaccins, alertes pour des examens de dépistage (cancer du sein, cancer colorectal), prophylaxie de l'infarctus, informations sur les substances addictives, etc.

Dans le cadre de l'observance, les chatbots peuvent aider à une meilleure prise en charge et une meilleure acceptation de la maladie : envoi de rappels, demande de confirmation de la prise effective du traitement, coaching, suivi dans le temps. De par la création d'un lien continu et durable, ils sont susceptibles de favoriser **un accompagnement au changement de comportement durable**. Même s'il convient de souligner que la technologie, aussi performante soit-elle, ne sera jamais un remède miracle contre une problématique éminemment humaine et complexe, dans laquelle entrent en jeu de nombreux facteurs situationnels et psychosociaux.

Enfin les laboratoires pharmaceutiques auraient un intérêt évident à concevoir des chatbots liés à un médicament précis afin d'en améliorer l'utilisation et de veiller à la bonne observance du patient. Ce domaine est certes peu exploité aujourd'hui, mais des solutions existent et peuvent être envisagées pour communiquer, en ligne avec la réglementation.

IV.2.2. MATHEMATIQUE DES RESEAUX DE NEURONES

Le but de cette partie est de comprendre les mathématiques liées aux réseaux de neurones et le calcul des poids par rétro propagation, est de se consacrer à la convolution qui est une opération mathématique simple pour extraire des caractéristiques d'une image et permet d'obtenir des réseaux de neurones performants.

LES DERIVEES

La notion de dérivée joue un rôle clé dans l'étude des fonctions. Elle permet de déterminer les variations d'une fonction et de trouver ses extrêmes. Une formule fondamentale pour la suite sera la formule de la dérivée d'une fonction composée.

Soit $f : I \rightarrow \mathbb{R}$ une fonction, où I est un intervalle ouvert de \mathbb{R} (par exemple du type $]a, b[$). Soit $x_0 \in I$.

La dérivée de f en x_0 , si elle existe, est le nombre

Équation 2 : définition de la dérivée

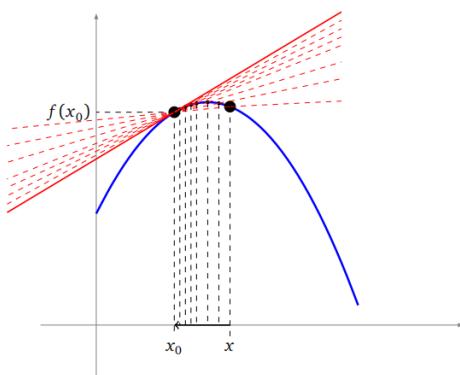
$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

L'interprétation géométrique de la dérivée est essentielle ! Le coefficient directeur de la tangente au graphe de f en x_0 est $f'(x_0)$. Une équation de la tangente au point $(x_0, f(x_0))$ est donc :

Équation 3 : équation tangentielle d'une dérivée

$$y = (x - x_0)f'(x_0) + f(x_0)$$

Figure 15 : représentation graphique tangentielle d'une dérivée



LE GRADIENT

Le gradient est un vecteur qui remplace la notion de dérivée pour les fonctions de plusieurs variables. On sait que la dérivée permet de décider si une fonction est croissante ou décroissante. De même, le vecteur gradient indique la direction dans laquelle la fonction croît ou décroît le plus vite. Nous allons voir comment calculer de façon algorithmique le gradient grâce à la « différentiation automatique ».

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction admettant des dérivées partielles. Le **gradient** de f en $(x_0, y_0) \in \mathbb{R}^2$, noté $\text{grad } f(x_0, y_0)$, est le vecteur :

Équation 4 : le vecteur gradient

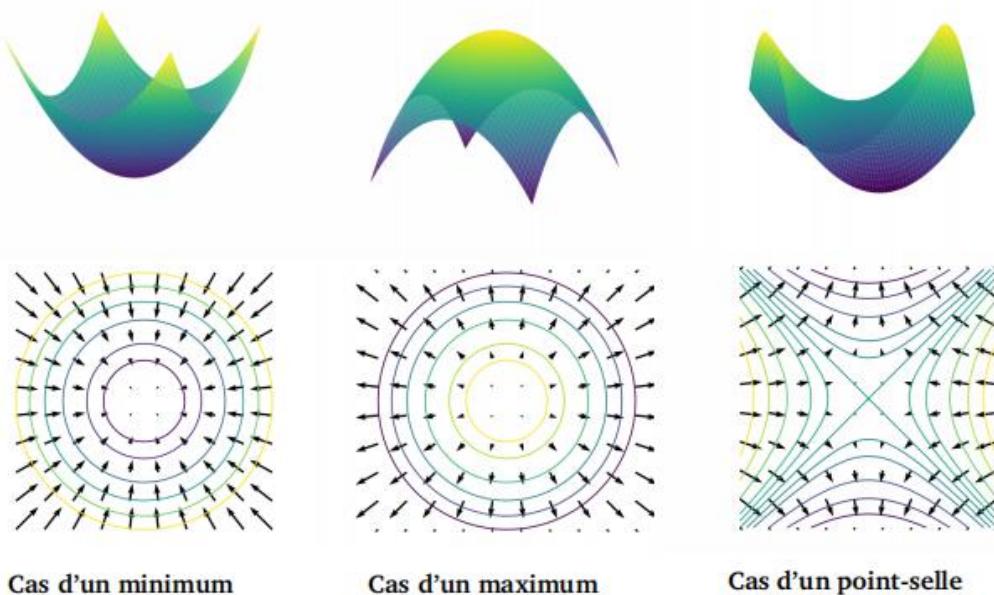
$$\text{grad } f(x_0, y_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}.$$

Le gradient est un élément de \mathbb{R}^n écrit comme un vecteur colonne. Parfois, pour alléger l'écriture, on peut aussi l'écrire sous la forme d'un vecteur ligne.

Comme il peut être difficile de calculer les points critiques de façon exacte, nous allons utiliser des méthodes numériques. L'idée qui sera détaillée dans le prochain chapitre est la suivante : comme le gradient indique la direction dans laquelle la fonction f croît le plus rapidement, nous allons suivre la direction opposée au gradient, pour laquelle f décroît le plus rapidement. Ainsi, partant d'un point (x_0, y_0) au hasard, on sait dans quelle direction se déplacer pour obtenir un nouveau point (x_1, y_1) en lequel f est plus petite. Et on recommence.

Sur les trois dessins ci-dessous, on a dessiné les lignes de niveau d'une fonction f ainsi que les vecteurs $\text{grad } f(x, y)$. On voit que ces vecteurs pointent bien vers le minimum (figure de gauche), s'éloignent d'un maximum (figure centrale), le cas d'un point-selle est spécial (figure de droite). Dans tous les cas, la longueur des vecteurs gradients diminue à l'approche du point critique.

Figure 16 : méthode numérique pour les calcul du gradient



LA DESCENTE DE GRADIENT

L'objectif de la méthode de **descente de gradient** est de trouver un minimum d'une fonction de plusieurs variables le plus rapidement possible. L'idée est très simple, on sait que le vecteur opposé au gradient indique une direction vers des plus petites valeurs de la fonction, il suffit donc de suivre d'un pas cette direction et de recommencer. Cependant, afin d'être encore plus rapide, il est possible d'ajouter plusieurs paramètres qui demandent pas mal d'ingénierie pour être bien choisis.

On nous donne une fonction f de deux variables (a, b) et nous cherchons un point (a_{\min}, b_{\min}) en lequel f atteint un minimum. Voici la méthode expliquée par des dessins sur lesquels ont été tracées des lignes de niveau:

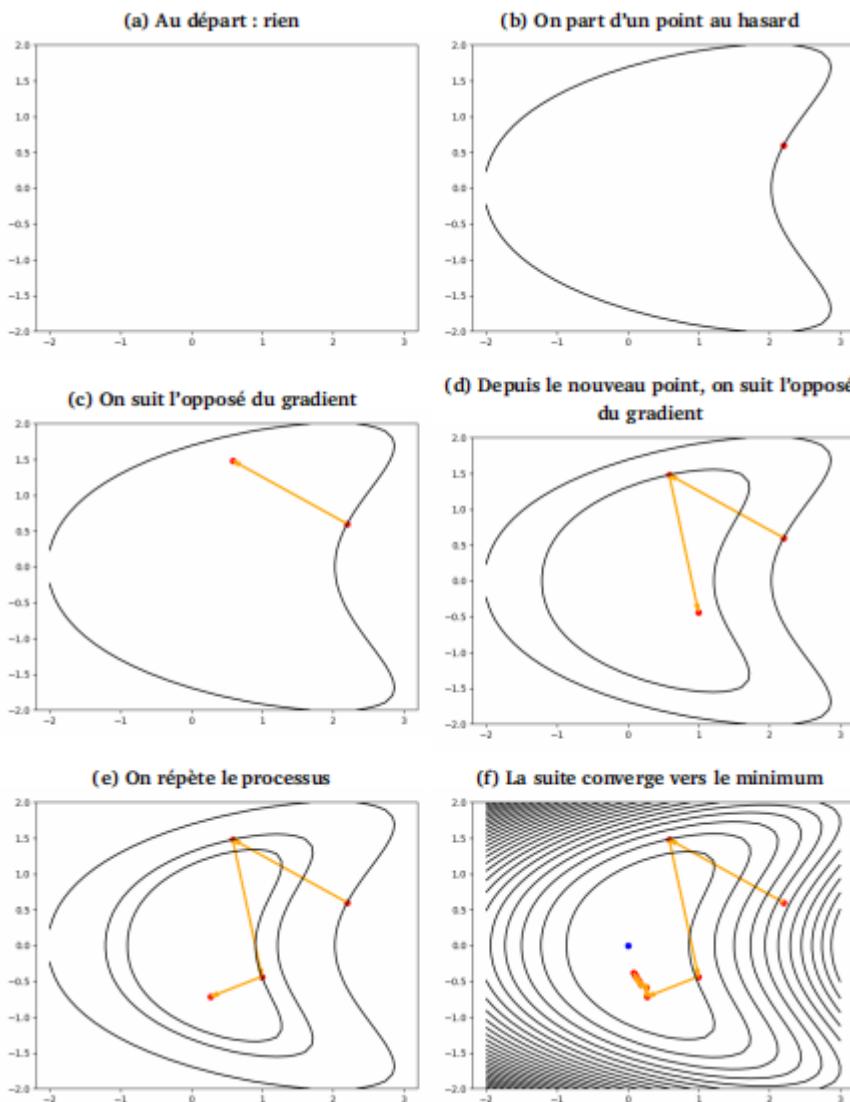
Figure 17 : recherche du minimum pour une fonction

Figure (a). Au départ nous n'avons aucune information globale sur f . La seule opération que l'on s'autorise c'est calculer $\text{grad } f(a, b)$ en certains points.

Figure (b). On choisit un point (a_0, b_0) au hasard. Si on note $c_0 = f(a_0, b_0)$ la valeur de f en ce point, on sait que la ligne de niveau $(f = c_0)$ passe par (a_0, b_0) .

Figure (c). On calcule en ce point le gradient de f . On trace l'opposé du gradient : $-\text{grad } f(a_0, b_0)$. On sait d'une part que la ligne de niveau est orthogonale à ce gradient et surtout que dans la direction de $\text{grad } f(a_0, b_0)$, les valeurs de f vont diminuer. On se dirige alors dans la direction opposée au gradient d'un facteur δ (par exemple $\delta = 0.1$). On arrive à un point noté (a_1, b_1) . Par construction, si δ est assez petit, la valeur $c_1 = f(a_1, b_1)$ est plus petite que c_0 .

Figure (d). On recommence depuis (a_1, b_1) . On calcule l'opposé du gradient en (a_1, b_1) , on se dirige dans cette nouvelle direction pour obtenir un point (a_2, b_2) où $c_2 = f(a_2, b_2) < c_1$.

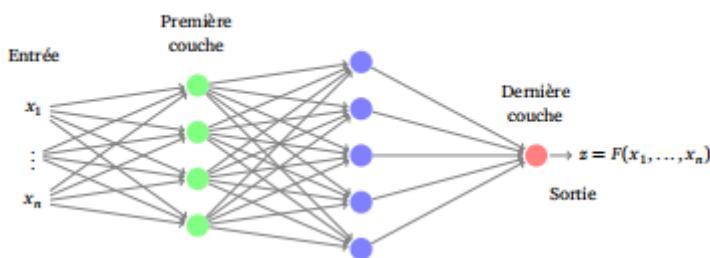
Figure (e). On itère le processus pour obtenir une suite de points (a_k, b_k) pour lesquels f prend des valeurs de plus en plus petites.

Figure (f). On choisit de s'arrêter (selon une condition préalablement établie) et on obtient une valeur approchée (a_N, b_N) du point (a_{\min}, b_{\min}) en lequel f atteint son minimum. Évidemment avec la vision globale de la fonction, on se dit qu'on aurait pu choisir un point de départ plus près et que certaines directions choisies ne sont pas les meilleures. Mais souvenez-vous que l'algorithme est « aveugle », il ne calcule pas les valeurs de f en les (a_k, b_k) et n'a pas connaissance du comportement de f au voisinage de ces points.

LA RETRO PROAGATION

La rétro propagation, c'est la descente de gradient appliquée aux réseaux de neurones. Elle étudie des problèmes variés et analyse les solutions produites par des réseaux de neurones. Partons de la figure suivante pour bien assimiler le terme :

Figure 18 :réseau de neurones pour classification



On dispose de données (X_i, z_i) (pour $i = 1, \dots, N$) où $X_i \in \mathbb{R}^n$ est une **entrée** (de la forme $X = (x_1, \dots, x_n)$) et $z_i \in \mathbb{R}$ est la **sortie attendue** pour cette entrée. Le but est de trouver les poids du réseau afin que la fonction F qui lui est associée vérifie : $F(X_i) \approx z_i$ pour tout $i = 1, \dots, N$.

Pour mesurer précisément la performance de l'approximation, on définit une **fonction erreur** :

Équation 5 : fonction d'erreur

$$E = \frac{1}{N} \sum_{i=1}^n E_i$$

La conception d'un réseau de neurones est réalisée en modélisant au mieux les données injectées. Mais l'objectif réel est de faire des prédictions pour de nouvelles valeurs, jamais rencontrées auparavant. La descente de gradient produit un ensemble de poids P qui définit complètement notre réseau R . Nous obtenons donc une fonction $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$, construite de sorte que $F(X_i) \approx z_i$. Nous pouvons évaluer cette fonction pour tout $X \in \mathbb{R}^n$, même pour des X différents des X_i .

CONVOLUTION

La convolution est une opération qui à partir d'un tableau de nombres et d'un motif produit un nouveau tableau de nombres.

On calcule la liste de sortie terme par terme :

- On centre le motif renversé sous la liste d'entrée, à la position à calculer,
- On multiplie terme à terme les éléments de la liste d'entrée et ceux du motif,
- La somme de tous ces produits est le terme de la liste de sortie.

Soient $(f(n))_{n \in \mathbb{Z}}$ et $(g(n))_{n \in \mathbb{Z}}$ deux suites de nombres réels. Le **produit de convolution** $f * g$ est la suite $(h(n))_{n \in \mathbb{Z}}$ dont le terme général est défini par :

Équation 6 : produit de convolution

$$f * g(n) = \sum_{k=-\infty}^{+\infty} f(n-k).g(k)$$

Il ne faut pas avoir peur de cette formule. En particulier, dans les situations rencontrées ici, il n'y a pas vraiment une infinité de termes à calculer. Voici une formule plus simple, lorsque l'on suppose que les termes de g sont nuls en dehors des indices appartenant à $[-K, +K]$:

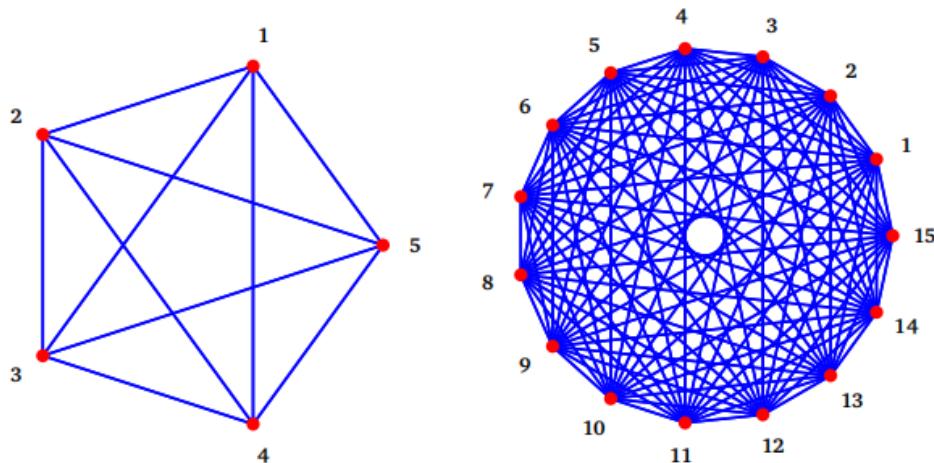
$$f * g(n) = \sum_{k=-K}^{+K} f(n-k).g(k)$$

LE PROBALITE

Dropout

Le dropout est une technique qui simule différentes configurations de liens entre les neurones et limite le sur-apprentissage. Imaginons n neurones. Quelle est la meilleure architecture pour les relier entre eux ? Bien sûr la réponse dépend du problème, ainsi on ne peut pas le savoir à l'avance. Une première façon de faire est de relier tous les neurones entre eux.

Équation 7 : liaison de point d'un réseau de neurones



Voici l'idée du dropout (avec paramètre $p = 0.5$). On part d'un réseau de n neurones, tous reliés les uns aux autres. Avant la première étape d'apprentissage, on décide de désactiver certains neurones. Cette décision est prise au hasard. Pour chaque neurone on lance une pièce de monnaie, si c'est « pile » on conserve le neurone, si c'est « face » on le désactive. Ensuite on effectue une étape de la descente de gradient, avec seulement une partie de nos neurones activés. Avant la deuxième étape de la descente de gradient, on reprend notre pièce et on choisit au hasard les neurones à désactiver, etc.

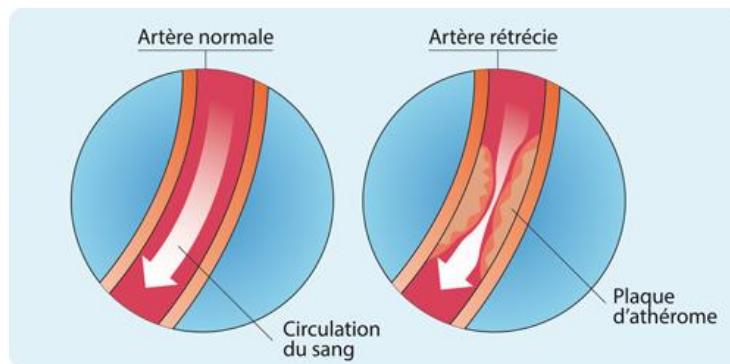
IV.3. SECTION DEUXIÈME : ANALYSE DE L'OBJET D'ÉTUDE

De nos jours, l'apprentissage automatique s'impose dans une grande variété de domaines d'application. Avec la généralisation du "Big Data", son utilisation s'est tellement répandue qu'il est aujourd'hui implanté dans la quasi-totalité des infrastructures urbaines de la vie quotidienne. Le domaine médical ne faisant pas exception, le concept de la "E-santé" est devenu de plus en plus performant et repose toujours plus sur des systèmes de santé tirant partie des prouesses de l'apprentissage automatique. Alors, de plus en plus d'appareils médicaux portables, comme les capteurs mesurant les signaux physiologiques, sont développés et ces signaux deviennent donc plus facilement mesurables dans un environnement quotidien (Injibar, 2021).

IV.3.1. MALADIE CARDIO-VASCULAIRE

Les maladies cardio-vasculaires (MCV) sont généralement décrites comme des conditions qui impliquent un rétrécissement ou un blocage (thrombose) des vaisseaux sanguins pouvant entraîner une cardiopathie ischémique (CI) (infarctus du myocarde, angine de poitrine, etc.) ou un accident vasculaire cérébral (AVC). Cela empêche le sang d'atteindre le cœur ou le cerveau. La raison la plus courante est une accumulation de dépôts graisseux sur les parois internes des vaisseaux sanguins qui forment le caillot bloquant les artères. Les AVC peuvent être causés par des caillots sanguins ou des saignements d'un vaisseau sanguin dans le cerveau (A. Freeman, 2016). La figure ci-dessous illustre la différence entre une artère normale et une artère rétrécie. Dans cette section, les maladies cardiovasculaires étudiées dans le cadre de ce mémoire sont brièvement présentées. Une maladie en particulier fut le centre d'attention d'une plus grande partie de mes travaux de recherche. Il s'agit de la fibrillation atriale qui sera présentée en premier lieu, ainsi que les résultats d'une analyse épidémiologique poussée. Les autres maladies cardiaques seront présentées sous une forme plus abrégée.

Figure 19 : la différence entre une artère normale et une artère rétrécie



IV.3.2. Cardiopathie hypertensive

Les maladies cardiaques hypertensives se réfèrent aux maladies cardiaques causées par une pression artérielle élevée. Le cœur travaillant sous une pression accrue provoque différents troubles cardiaques. Les maladies cardiaques hypertensives comprennent l'épaississement du muscle cardiaque, la maladie coronarienne et d'autres complications. Une cardiopathie hypertensive peut entraîner de graves problèmes de santé. C'est la

principale cause de décès par hypertension. En général, les problèmes cardiaques associés à l'hypertension artérielle sont liés aux artères et aux muscles du cœur.

On distingue deux types de cardiopathie hypertensive. Le premier est le rétrécissement des artères. Les artères coronaires transportent le sang vers le muscle cardiaque. Lorsque l'hypertension artérielle fait rétrécir les vaisseaux sanguins, le flux sanguin vers le cœur peut ralentir ou s'arrêter. Cette condition est connue sous le nom de maladie coronarienne. Cette dernière rend difficile le fonctionnement du cœur et l'approvisionnement en sang des autres organes. Cela peut exposer l'individu à un risque de crise cardiaque à cause d'un caillot de sang qui se coince dans l'une des artères rétrécies et coupe le flux sanguin vers le cœur. Le second type de cardiopathie hypertensive est l'épaississement et l'élargissement du cœur (S. Yusuf, 2014).

L'hypertension artérielle empêche le cœur de pomper correctement le sang dans l'organisme. Comme les autres muscles du corps, un travail acharné régulier entraîne l'épaississement et la croissance des muscles cardiaques. Cela modifie le fonctionnement du cœur. Ces changements se produisent généralement dans la chambre de pompage principale du cœur, le ventricule gauche. La condition est connue comme l'hypertrophie ventriculaire gauche (science, 2023).

IV.3.3. INFARCTUS AIGU DU MYOCARDE

L'infarctus aigu du myocarde est le nom médical d'une crise cardiaque. Une crise cardiaque est une maladie mortelle qui survient lorsque le flux sanguin vers le muscle cardiaque est brusquement interrompu, causant des dommages aux tissus. Ceci est généralement le résultat d'un blocage dans une ou plusieurs des artères coronaires. Un blocage peut se développer en raison d'une accumulation de plaque, une substance principalement constituée de matières grasses, de cholestérol et de déchets cellulaires. Bien que les symptômes classiques d'une crise cardiaque soient des douleurs thoraciques et un essoufflement, les symptômes peuvent être assez variés.

Les symptômes les plus courants d'une crise cardiaque comprennent une pression ou oppression dans la poitrine, une douleur dans la poitrine, le dos, la mâchoire et d'autres parties du haut du corps qui dure plus de quelques minutes ou qui disparaît et revient, un essoufflement, la transpiration, un sentiment de nausée, des vomissements, de l'anxiété, de la toux, des vertiges et une fréquence cardiaque rapide. Toutes les personnes qui ont une crise cardiaque ne présentent pas les mêmes symptômes ni la même gravité des symptômes. La douleur thoracique est le symptôme le plus souvent signalé chez les femmes et les hommes. Cependant, les femmes sont plus susceptibles que les hommes de ressentir un essoufflement, une douleur à la mâchoire ou dans le haut du dos, des étourdissements, de la nausée et des vomissements. Le cœur est le principal organe du système cardio-vasculaire, qui comprend également différents types de vaisseaux sanguins.

Certains des vaisseaux les plus importants sont les artères. Ils transportent du sang riche en oxygène vers le corps et tous les organes. Les artères coronaires transportent le sang riche en oxygène spécifiquement vers le muscle cardiaque. Lorsque ces artères deviennent bloquées ou rétrécies en raison d'une accumulation de plaque, le flux sanguin vers le cœur peut diminuer considérablement ou s'arrêter complètement. Cela peut provoquer une crise cardiaque. Plusieurs facteurs peuvent conduire à un blocage des artères coronaires. D'abord le mauvais cholestérol, également appelé lipoprotéine de basse densité, est l'une

des principales causes d'un blocage dans les artères. Le cholestérol est une substance incolore présente dans les aliments comestibles au quotidien. Le corps le produit aussi naturellement. Tout le cholestérol n'est pas mauvais, mais la lipoprotéine de basse densité peut coller aux parois des artères et produire de la plaque. La plaque est une substance dure qui bloque la circulation sanguine dans les artères (Sullivan, 2018).

IV.3.4. ANGINE DE POITRINE

L'angine de poitrine est un type de douleur thoracique qui résulte d'une diminution du flux sanguin vers le cœur. Un manque de circulation sanguine signifie que le muscle cardiaque ne reçoit pas suffisamment d'oxygène. La douleur est souvent déclenchée par l'activité physique ou le stress émotionnel. L'angine de poitrine stable est le type d'angine de poitrine le plus courant. Elle constitue un schéma prévisible de douleur thoracique. L'angine de poitrine instable est une autre forme d'angine de poitrine. Elle survient soudainement et s'aggrave avec le temps. Cela peut éventuellement conduire à une crise cardiaque. Bien que l'angine de poitrine stable soit moins grave que l'angine instable, elle peut être douloureuse et inconfortable. Les deux types d'angine de poitrine sont généralement des signes d'une affection cardiaque sous-jacente. L'angine de poitrine stable se produit lorsque le muscle cardiaque ne reçoit pas l'oxygène dont il a besoin pour fonctionner correctement. Le cœur travaille alors plus fort lors d'une activité physique ou d'un stress émotionnel. Certains facteurs, tels que le rétrécissement des artères (athérosclérose), peuvent empêcher le cœur de recevoir plus d'oxygène. Les artères peuvent devenir étroites et dures lorsque la plaque (une substance constituée de graisse, de cholestérol, de calcium et d'autres substances) s'accumule à l'intérieur des parois des artères.

Cependant, il est plus probable de ressentir des symptômes le matin. La sensation douloureuse qui se produit lors d'un épisode d'angine stable est souvent décrite comme une pression ou une plénitude au centre de la poitrine. La douleur peut se sentir comme un étau serrant votre poitrine ou comme un poids lourd reposant sur votre poitrine. Cette douleur peut se propager de votre poitrine à votre cou, vos bras et vos épaules. Les facteurs de risque de l'angine de poitrine stable sont nombreux (Sullivan, Stable Angina., 2018).

On en distingue surtout l'obésité, les antécédents de maladies cardiaques, un taux de cholestérol élevé ou une pression artérielle élevée, le diabète, le tabagisme, l'inactivité physique. Les repas copieux, des entraînements physiques vigoureux et un temps extrêmement chaud ou froid peuvent également déclencher une angine stable dans certains cas.

IV.3.5. ARYTHMIES CARDIAQUES

Une arythmie est un trouble du cœur qui affecte le rythme auquel le cœur bat. Une arythmie se produit lorsque les impulsions électriques, qui dirigent et régulent les battements cardiaques, ne fonctionnent pas correctement. Cela fait battre le cœur soit trop vite (tachycardie), soit trop lentement (bradycardie), soit trop tôt (contraction prématuée) ou soit de façon erratique (fibrillation). Presque tout le monde connaîtra au moins une fois un rythme cardiaque anormal. Les arythmies sont courantes et généralement inoffensives, mais certaines sont problématiques. Lorsqu'une arythmie interfère avec le flux sanguin vers le corps, elle peut endommager le cerveau, les poumons et d'autres organes vitaux. Si elles ne sont pas traitées, ces arythmies peuvent être mortelles. Le cœur est divisé en quatre chambres. Chaque moitié du cœur se compose d'une chambre supérieure (l'oreillette) et

d'une chambre inférieure (le ventricule). Les deux moitiés créent deux pompes, une de chaque côté du cœur. Dans un cœur qui bat correctement, les impulsions électriques suivent des voies précises à travers le cœur jusqu'à chaque pompe. Ces signaux coordonnent l'activité du muscle cardiaque afin que le sang pénètre dans le cœur et en sort. Toute interruption de ces voies ou impulsions peut provoquer un rythme cardiaque anormal (D. Sullivan, 2017).

Cela fait chuter la tension artérielle et diminue l'apport sanguin au corps et aux organes. La fibrillation ventriculaire est la première cause d'arrêt cardiaque soudain. Les battements cardiaques prématurés peuvent donner l'impression que le cœur a sauté un battement. En réalité, le rythme cardiaque normal a été interrompu par un battement trop tôt et un battement supplémentaire entre deux battements de cœur normaux est alors ressenti.

IV.3.6. INSUFFISANCE CARDIAQUE

L'insuffisance cardiaque se caractérise par l'incapacité du cœur à pomper une quantité suffisante de sang dans le corps. Sans un flux sanguin suffisant, toutes les principales fonctions du corps sont perturbées. L'insuffisance cardiaque est une affection ou un ensemble de symptômes qui affaiblissent votre cœur. Chez certaines personnes souffrant d'insuffisance cardiaque, le cœur a du mal à pomper suffisamment de sang pour soutenir d'autres organes du corps. D'autres personnes peuvent avoir un durcissement et un raidissement du muscle cardiaque lui-même, ce qui bloque ou réduit le flux sanguin vers le cœur. Il peut s'agir d'une affection aiguë (à court terme) ou chronique (à long terme). Dans l'insuffisance cardiaque aiguë, les symptômes apparaissent soudainement mais disparaissent assez rapidement. Cette condition survient souvent après une crise cardiaque. Cela peut également être le résultat d'un problème avec les valves cardiaques qui contrôlent le flux sanguin dans le cœur. Cependant, dans l'insuffisance cardiaque chronique, les symptômes sont continus et ne s'améliorent pas avec le temps. La grande majorité des cas d'insuffisance cardiaque sont chroniques. L'insuffisance cardiaque est une maladie grave qui nécessite un traitement. Un traitement précoce augmente les chances de guérison à long terme avec moins de complications. Les symptômes de l'insuffisance cardiaque peuvent inclure une fatigue excessive, la prise de poids soudaine, une perte d'appétit, une toux persistante, un pouls irrégulier, des palpitations cardiaques, un gonflement abdominal, un essoufflement, un gonflement des jambes et des chevilles et des veines cervicales saillantes (Sullivan, Heart Failure, 2020).

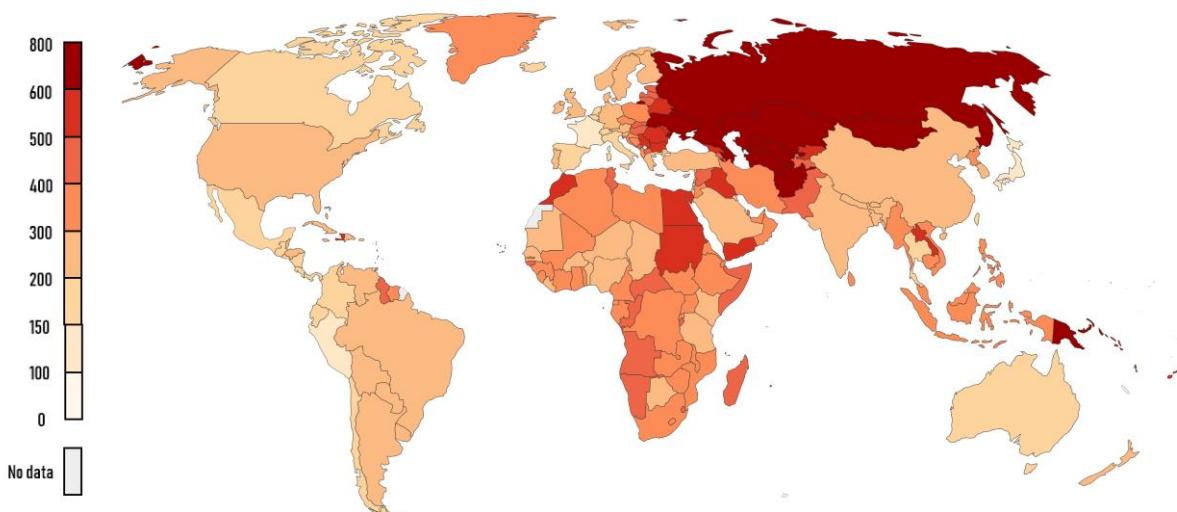
Les contractions du cœur sont nécessaires pour pomper le sang riche en oxygène vers le corps. Ce problème est connu sous le nom de dysfonction systolique, et il se développe généralement lorsque le cœur est faible et hypertrophié. L'insuffisance cardiaque systolique est plus fréquente chez les hommes que chez les femmes. Une insuffisance cardiaque diastolique et systolique peut survenir du côté gauche ou droit du cœur

IV.3.7. ANALYSE EPIDEMIOLOGIQUE

Les maladies cardio-vasculaires sont la principale cause de décès et d'invalidité (Metrics, 2020) dans le monde. Dix-huit millions de décès par an sont attribués aux maladies cardio-vasculaires dans le monde (C. J. Murray, 2012). L'incidence des événements cardio-vasculaires majeurs a été signalée comme étant la plus élevée dans les pays à faible revenu, malgré le fait que ces pays présentent des facteurs de risque plus faibles (S. Yusuf, 2014). Les données sur la mortalité sont couramment utilisées dans la surveillance des maladies. Le taux

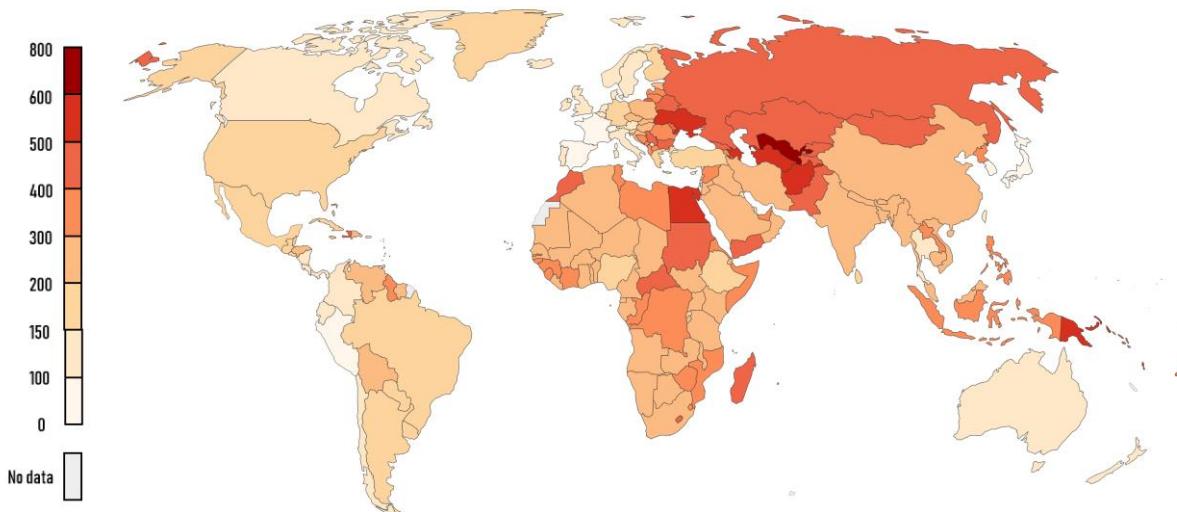
de mortalité d'une maladie peut fournir une indication de son influence au sein d'une population, en particulier pour les maladies chroniques telles que les maladies cardio-vasculaires. De plus, la mortalité est l'un des résultats les plus fiables. La déclaration des décès fait souvent partie intégrante du système de statistiques de l'état civil d'un pays. Dans la plupart des pays, les autorités sanitaires exigent que chaque décès dans la population soit signalé, y compris un enregistrement de la cause du décès. Ces enregistrements peuvent ensuite être compilés afin de produire des mesures de mortalité pour la population, y compris le nombre absolu de décès et les taux qui s'ajustent à la taille et à la distribution de la population. Bien que les systèmes d'enregistrement des décès soient universels, leur qualité et leur exhaustivité peuvent varier. Dans la plupart des pays développés, ce système est relativement complet et les taux de mortalité calculés à partir des données sont assez précis.

Figure 20 : Répartition des décès par maladies cardio-vasculaires standardisée selon l'âge par 100000 hab en 2000



Source : www.ouronlinedata.org

Figure 21 : Répartition des décès par maladies cardio-vasculaires standardisée selon l'âge par 100000 hab en 2017



IV.3.8. LES FACTEURS RISQUES

Il existe une relation linéaire entre les niveaux de pression artérielle et le risque d'accident vasculaire cérébral ou d'infarctus du myocarde (S. Lewington, 2003). Des enquêtes internationales montrent que le taux d'hypertension artérielle systolique (≥ 140 mmHg) a considérablement augmenté entre 1990 et 2015 avec des effets d'entraînement sur les années de vie ajustées sur l'incapacité et les décès attribuables à l'hypertension (M. H. Forouzanfar, 2017). L'étude INTERHEART a estimé que 22% des infarctus du myocarde en Europe sont liés à l'hypertension, ce qui double presque le risque par rapport aux personnes sans antécédents d'hypertension (S. Yusuf, 2014). Le traitement pour abaisser la tension artérielle offre une protection significative contre les événements cardio-vasculaires, avec des avantages supplémentaires de traitement plus intensif à l'égard des patients à haut risque (maladie vasculaire concomitante, maladie rénale ou diabète) (X. Xie, 2016). En 2015, la prévalence médiane standardisée selon l'âge de l'hypertension artérielle (définie comme la pression artérielle systolique ≥ 140 mmHg ou la pression artérielle diastolique ≥ 90 mmHg) dans les pays membres de l'ESC était de 24.8% (intervalle interquartile (IIQ) [19.8; 28.5] %).

MAUVAISES HABITUDES LIEES

Le tabac a été décrit comme « le plus grand risque sanitaire évitable dans l'Union Européenne » par la Direction générale de la santé et des consommateurs (DG SANTÉ). Le tabac est lié à de nombreuses formes de cancer et de maladies cardio-vasculaires. De plus, c'est la principale cause de décès de près de 6 millions de personnes par an. Par conséquent, l'UE a encouragé des mesures politiques liées à l'usage du tabac et à la commercialisation des dérivés du tabac au cours des 15 dernières années. Au cours de cette période, la prévalence du tabac a constamment diminué en Europe (S. J. Hoffman, 2015).

La consommation d'alcool est définie comme la quantité enregistrée (en litres) d'alcool pur consommée par adulte (plus de 15 ans) au cours d'une année civile et est un indicateur de santé de base européen. L'indicateur ne prend en compte que la consommation enregistrée à partir de la production, import, export et vente des données, souvent via la fiscalité. La consommation excessive d'alcool reste la principale cause de décès prématuré, où elle est responsable de 1 décès sur 10 chez les adultes en âge de travailler (M. Stahre, 2014). Dans l'UE, la consommation nocive d'alcool est la troisième cause de décès prématuré après le tabac et l'hypertension avec l'alcool. On estime que la dépendance est responsable de plus de 60% de la mortalité attribuable à l'alcool (Baumberg, 2006).

La faible consommation de fruits et légumes est un facteur de risque reconnu de maladies non transmissibles, comme le cancer et les maladies coronariennes (V. Miller, 2017). Dans une revue systématique et une méta-analyse dose-réponse de 95 études prospectives, des réductions progressives du risque de MCV et de la mortalité toutes causes confondues ont été observées jusqu'à un apport de 800 g/jour de fruits et légumes combinés, alors que pour le cancer total, aucune d'autres réductions du risque ont été observées au-dessus de 600g/jour. Comme toutes les études examinant les relations entre la nutrition et la maladie, une interprétation prudente est nécessaire, en partie en raison de la confusion par un mode de vie sain qui rend les contributions nutritionnelles à la réduction du risque difficiles à déterminer et en partie en raison du manque variable de précision dans la mesure diététique. Les dernières estimations disponibles pour la consommation de légumes et de fruits ont été obtenues en 2014 pour les personnes de plus de 15 ans et se limitaient à 22 pays

à revenu élevé et à un pays à revenu intermédiaire (Turquie). Dans ces 23 pays membres du CES, 52.2% des personnes consommaient au moins une portion de légumes par jour et 55% consommaient au moins une portion de fruits par jour

L'activité physique insuffisante est définie comme la proportion de la population atteignant moins de 150 minutes d'activité physique d'intensité modérée par semaine ou moins de 75 minutes d'activité physique d'intensité vigoureuse par semaine. L'inactivité augmente le risque de plusieurs maladies non transmissibles telles que la cardiopathie ischémique, le diabète de type 2, les cancers du sein et du côlon, et représente près de 10% de tous les décès dans le monde (I.-M. Lee, 2012). La promotion de l'exercice pendant les loisirs a constamment démontré qu'elle favorisait la santé cardio-vasculaire (M. Hamer, 2017). En 2016, la prévalence médiane normalisée selon l'âge de l'activité physique insuffisante auto-déclarée était de 29.32% chez les adultes âgés de plus de 18 ans en France. La prévalence était généralement plus élevée chez les femmes que chez les hommes, avec plus de 43% des femmes à Chypre, en Allemagne, en Italie, à Malte et au Portugal déclarant une activité physique insuffisante par rapport à plus de 35% des hommes dans les mêmes pays.

ANALYSE DE MORBIDITE

Les mesures de la morbidité jouent un rôle important dans la description de l'épidémiologie d'une maladie, car elles représentent le nombre de personnes qui en souffrent au sein d'une population (D. Coggon, 2009). Les mesures de prévalence décrivent le nombre d'individus qui souffrent actuellement d'une maladie particulière dans une population donnée. Dans le cas d'événements cardiovasculaires aigus, tels que l'infarctus du myocarde ou l'accident vasculaire cérébral, ils indiquent le nombre de personnes vivant dans la population, qui ont déjà souffert d'un tel événement. L'incidence définit le nombre de nouveaux cas dans une population au cours d'une période donnée, fournissant une mesure de l'occurrence de la maladie. Même pour les conditions avec un taux de mortalité relativement élevé, comme les accidents vasculaires cérébraux et les infarctus aigus du myocarde, la morbidité est une mesure importante à côté de la mortalité, car l'invalidité causée par de tels événements peut être élevée chez ceux qui y survivent et il est difficile de traiter les individus indépendamment de leur niveau de récupération. Cependant, par rapport à la collecte de données sur la mortalité, qui implique d'agrégner le nombre de décès enregistrés, la collecte de données sur la morbidité est plus difficile. Malgré la prolifération des systèmes d'enregistrement électronique au sein des services de santé, peu de pays utilisent ces systèmes pour signaler des enregistrements précis de morbidité. Il faut donc s'appuyer sur d'autres sources pour les statistiques de la morbidité.

Figure 22 : Incidence des maladies cardio-vasculaire en 2017

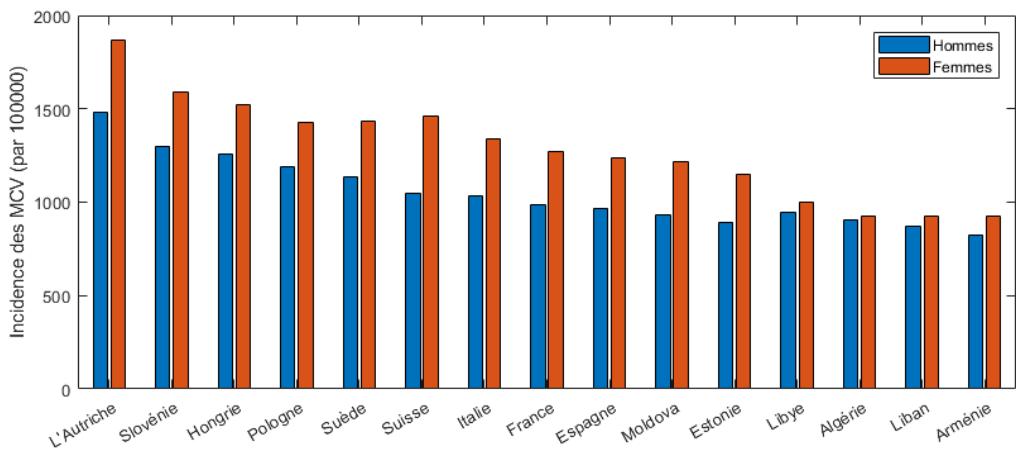
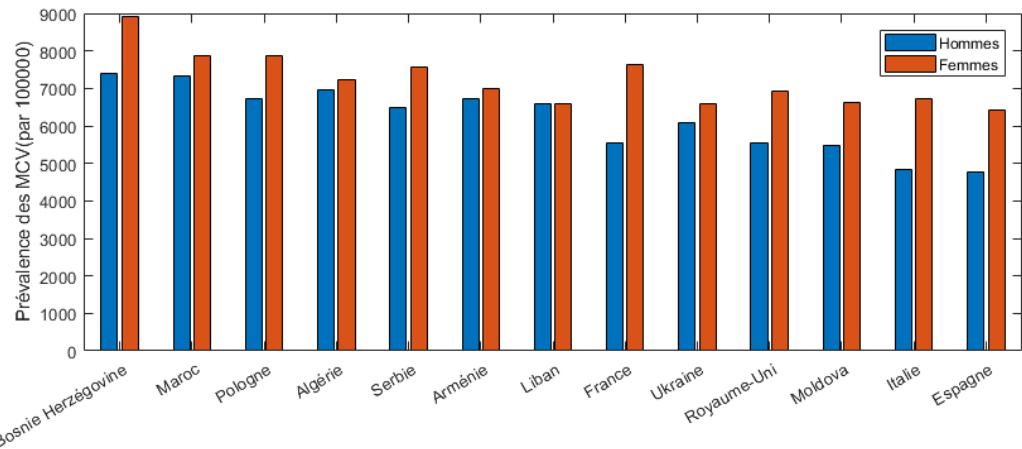


Figure 23 : Prévalence des maladies cardio-vasculaire en 2017



IV.4 CONCLUSION

Dans ce chapitre, nous avons mis l'accent sur les opérations de prétraitements effectuées sur les comptes rendus médicaux afin de rendre les connaissances qui y sont contenues accessibles. Nous avons démontré l'utilité de mathématique dans le processus de fouille de texte et du traitement automatique de langage naturel pour la détermination de certains paramètres dans la conception des algorithmes d'intelligence artificielle.

Après avoir présenté l'aspect technique autour de notre contexte d'étude nous nous sommes livrée à la présentation et l'analyse de l'objet d'étude. Dans le cas d'événements cardiovasculaires aigus, tels que l'infarctus du myocarde ou l'accident vasculaire cérébral, ils indiquent le nombre de personnes vivant dans la population, qui ont déjà souffert d'un tel événement. L'incidence définit le nombre de nouveaux cas dans une population au cours d'une période donnée, fournissant une mesure de l'occurrence de la maladie.

CONCLUSION DU DEUXIEME VOLET

Le domaine de santé a été l'un des secteurs qui a adopté le plus rapidement l'IoT, l'IA et les agents conversationnel. Surtout en période de saisonnalité et de pics, lorsque l'ensemble du secteur de santé est confronté à de nouveaux défis et à une demande très élevée par rapport aux ressources disponibles. Dans ce contexte, plusieurs possibilités d'application de l'IoT et de l'IA ont été examinées. Il comprend des consultations automatiques efficaces et le suivi à distance des patients, des diagnostics intelligents, le partage des EHR et la programmation prioritaire des patients. Plusieurs pays dans le monde sont confrontés à des défis dans des domaines tels que la fracture numérique et la disparité de la population ayant accès aux technologies numériques en matière de soins de santé. En dehors de cela, malgré d'autres défis tels que la puissance limitée dans le traitement des big data, l'interopérabilité des données de santé entre des parties prenantes hétérogènes et le manque de structure unifiée de mise en œuvre pour la santé en ligne, l'IA et l'IoT présentent un immense potentiel dans le secteur des soins de santé.

Pour améliorer l'efficience et l'efficacité des soins de santé, il convient de se concentrer sur le flux de patients, car il est au cœur de la capacité d'une organisation à fournir des services de santé, tandis que d'autres opérations apportent un soutien dans le but de créer les conditions d'un flux de patients efficient et efficace. Il est donc nécessaire d'avoir une vision holistique de l'ensemble du parcours du patient. La qualité globale des processus de soins de santé ne s'améliorera pas non plus tant que toutes les personnes impliquées dans la prise en charge d'un patient ne se considéreront pas comme faisant partie de la même équipe, indépendamment de leur affiliation organisationnelle. La nécessité d'une perspective systémique a également été mise en avant comme une condition préalable à la recherche de solutions possibles aux problèmes majeurs de la prestation de soins de santé. Suite à notre étude de l'ensemble des approches déjà utilisé nous allons expliquer dans ce qui suit les différents modèle, étapes et processus que nous avons pris pas à pas dans notre travail de recherche pour implémenter le système futur.

VOLET DE RESULTATS

Le système d'information est une réalité intrinsèque à l'hôpital, indépendante de toute informatisation. Il se situe au cœur du fonctionnement de l'établissement de santé, il couvre l'ensemble des informations utilisées dans cet hôpital. La mise en place d'un système informatique est souvent l'occasion de son réexamen, en vue de son automatisation plus ou moins complète, aboutissant à la constitution d'un SIH. Cette automatisation recouvre généralement les fonctions de mémorisation et de communication, voire de traitement.

Dans le présent volet, nous allons présenter en premiers lieux les manquements décellés dans les systèmes actuels, les failles observer. Signalons que nous n'aborderons pas le système de santé en général, comme nous cherchons dès le début de cette recherche à trouver des approches performant pour éviter engorgement aux urgences et proposer des outils d'aide à l'observance dans le milieux d'habitation nous allons spécifier le manquement lier aux services d'urgences hospitalier.

À la suite nous allons parler des outils utiles à la conception d'un nouveau système pour éclairer certains mauvais choix de technologie pris par nos prédecesseurs. Les données cliniques utiliser dans cette recherche sont certifiées par le CDC. Des modèles basés sur UML suivront pour expliciter la conception de notre système, nous finirons par l'interprétation de résultat par de graphiques de présentations des données traitées et des interfaces utilisateur.

CHAPITRE V. MISE EN ŒUVRE ET MODELISATION

V.I. INTRODUCTION

Une des contributions majeures de notre mémoire est la collecte et la création d'une armoire à base de données propre à notre système. En travaillant avec une équipe d'experts et de cliniciens, nous avons réussi à mettre en place huit base de données de plus de 300 milles patients acquis avec de dataset de différent pathologie, pandémie et épidémie.

En outre de la complexité du corps humain, et l'absence de signaux très informatifs, les signaux physiologiques peuvent être inaccessibles par moment et contenant des imperfections liées principalement à leurs modes d'acquisition. Ainsi, l'objectif scientifique est la conception d'une approche de classification sous contrainte de données imprécises, hétérogènes, manquantes par moment et indirectement corrélées au phénomène suivi. Cela implique la construction d'un classifier dont l'entrée est de taille variable et qui fusionne une grande quantité de données hétérogènes moyennement informatives pour en tirer une information efficace sur l'état surveillé. L'objectif applicatif est de pouvoir atteindre des performances de prédiction acceptables en utilisant des dispositifs sans électrode, et plus généralement, tout dispositif multifonctionnel, permettant aux personnes d'être prévenues et traitées prématûrement lorsqu'un risque est signalé

V.2. MISE EN ŒUVRE ET AQUISITION DE DONNEES

Une approche novatrice de modélisation de l'information est ainsi proposée, en définissant une fonction de masse originale représentant précisément les connaissances et les lacunes des données. En considérant un modèle par caractéristique, nous surmontons la forte variabilité de la disponibilité des signaux, et le rendons applicable même avec peu de données disponibles. Afin de construire les modèles, la base de données médicale MIMIC III est utilisée en sélectionnant les signaux de patients atteints de FA et les signaux d'autres patients sans FA. Les modèles disponibles sont ensuite combinés conduisant à un modèle plus informatif, qui nous donne une décision en guise de sortie du classifier. Dans la suite, nous commençons par présenter la base de données MIMIC III et la sélection des sujets et des signaux d'entrée.

V.2.1. BASE DE DONNEES MEDICALE MIMIC III

La base de données "**Medical Information Mart for Intensive Care III**" ou MIMIC III est une immense base de données librement accessible comprenant des données sur la santé associée à plus de quarante mille patients anonymisés qui ont séjourné dans des unités de soins intensifs (USI) d'un hôpital entre 2001 et 2012 (MIT-LCP, 2023). La base de données comprend des informations telles que les données démographiques, les mesures des signaux vitaux effectuées au chevet du patient (1 point de données par heure), les résultats des tests de laboratoire, les procédures, les médicaments, les notes du soignant, les rapports d'imagerie et la mortalité (à l'intérieur et à l'extérieur de l'hôpital). La base de données est hébergée sur la plateforme PHYSIONET (physionet, 2023). Elle est divisée en deux grandes parties qui sont les données cliniques et les données temporelles. La première regroupe les données cliniques des patients et contient notamment des informations sur les admissions, les sorties, les diagnostics, les transferts au sein de l'hôpital, le personnel médical, les procédures médicales, les analyses de laboratoire, les notes des infirmiers, les séjours à l'unité des soins

intensifs, etc. Les données cliniques de MIMIC III forment un ensemble de 40 tables, faisant un total de 534 colonnes et 728 556 685 lignes. La deuxième contient les données cartographiques enregistrées pour chaque patient. Ces données comprennent deux types de séries temporelles, l'une échantillonnée à haute fréquence (60 Hz) qui inclut l'electrocardiogramme et la pression artérielle continue, et l'autre qui enregistre en général un point par minute et qui comprend notamment le rythme cardiaque, la pression artérielle, la saturation pulsée en oxygène, etc. La FIGURE 19 montre un extrait d'une table de la base de données clinique qui liste les patients ayant fréquenté l'hôpital. Seules quelques colonnes de la table y sont exposées. On peut voir notamment l'identifiant du sujet "SUBJECT_ID", le genre "GENDER", la date de naissance "DOB" et la date de décès à l'hôpital "DOD_HOSP". On remarque que les informations temporelles sont toutes décalées d'une durée aléatoire et propre à chaque patient pour des raisons de confidentialité médicale

Figure 24: Extrait de la table "ADMISSIONS" de la base de données clinique

File: <base> / ADMISSIONS.csv (26,823 bytes)											
row_id	subject_id	hadm_id	admittime	dischtime	deathtime	admission_type	admission_location	discharge_location	insurance	lang	
12258	10006	142345	2164-10-23 21:09:00	2164-11-01 17:15:00		EMERGENCY	EMERGENCY ROOM ADMIT	HOME HEALTH CARE	Medicare		
12263	10011	105331	2126-08-14 22:32:00	2126-08-28 18:59:00	2126-08-28 18:59:00	EMERGENCY	TRANSFER FROM HOSP/EXTRAM	DEAD/EXPIRED	Private		
12265	10013	165520	2125-10-04 23:36:00	2125-10-07 15:13:00	2125-10-07 15:13:00	EMERGENCY	TRANSFER FROM HOSP/EXTRAM	DEAD/EXPIRED	Medicare		
12269	10017	199207	2149-05-26 17:19:00	2149-06-03 18:42:00		EMERGENCY	EMERGENCY ROOM ADMIT	SNF	Medicare		
12270	10019	177759	2163-05-14 20:43:00	2163-05-15 12:00:00	2163-05-15 12:00:00	EMERGENCY	TRANSFER FROM HOSP/EXTRAM	DEAD/EXPIRED	Medicare		
12277	10026	103770	2195-05-17	2195-05-24		EMERGENCY	EMERGENCY ROOM ADMIT	REHAB/DISTINCT PART HOSP	Medicare		

V.2.2. SELECTION DES SUJETS ET DES SIGNAUX

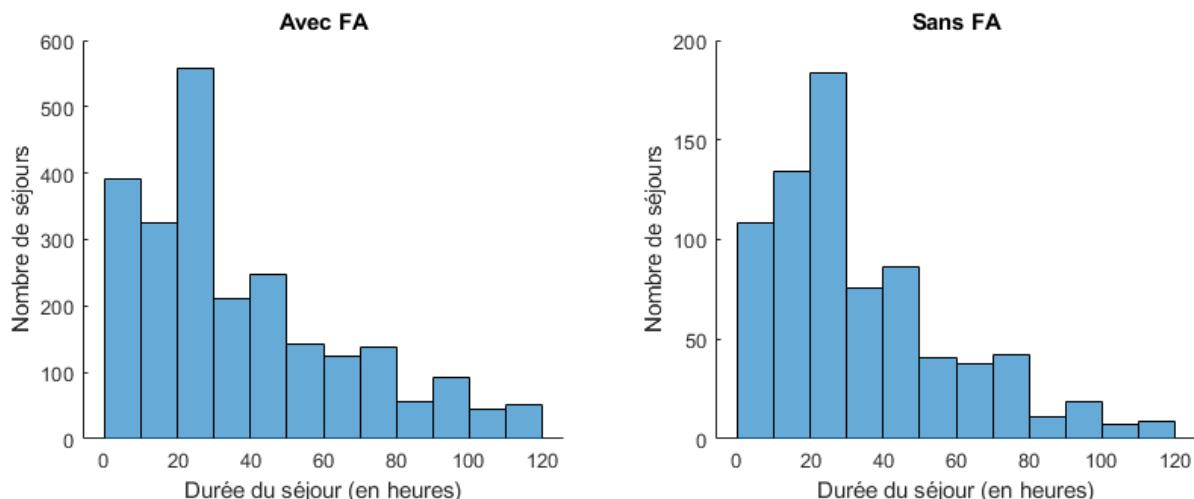
Pour étudier l'effet de la FA sur les signaux physiologiques avant la phase d'épisode, il est judicieux d'analyser les enregistrements physiologiques des patients diagnostiqués avec une FA. Pour cela, on sélectionne de la base de données médicale les patients en se basant sur la table des diagnostics appelée "DIAGNOSES_ICD" qui contient essentiellement les différents éléments diagnostiqués pour chaque admission à l'hôpital. La table "DIAGNOSES_ICD" possède 5 champs ou colonnes :

- ❖ **ROW_ID** : Entier unique pour chaque ligne ou enregistrement de la table;
- ❖ **SUBJECT_ID** : Entier identifiant le sujet ou le patient concerné par le diagnostic;
- ❖ **HADM_ID** : Entier identifiant l'admission à l'hôpital concernée par le diagnostic. Cet entier doit être conforme au champ de "SUBJECT_ID" selon la table "ADMISSIONS";
- ❖ **SEQ_NUM** : Entier indiquant l'ordre dans lequel les diagnostics se rapportent au patient. Les diagnostics sont classés par priorité et l'ordre a un impact sur le remboursement du traitement;
- ❖ **ICD9_CODE** : Chaîne de caractère représentant un code du classement international des maladies "ICD9" et indiquant l'objet du diagnostic.

Le "ICD9" pour "International Classification of Diseases 9" est un registre de classification des maladies qui attribue à chaque pathologie un code unique. Il est généralement constitué d'un nombre à 3 chiffres suivi d'un point puis d'un autre nombre à deux chiffres au plus.

Ainsi, on obtient la liste des admissions pour les patients atteints de FA⁴⁰(fibrillation auriculaire). En parallèle, nous sélectionnons au hasard des patients sans cardiopathies et avec peu de pathologies diagnostiquées, pour les sujets témoins. Cela évite d'avoir des similitudes entre les états ou les modèles des signaux physiologiques avec ceux des patients atteints de FA. Dans le groupe de sujets témoins, seuls les patients atteints de troubles mentaux, qui ont reçu un diagnostic de ICD9 entre 290 et 319, sont admis. Nous nous retrouvons alors avec 3605 séjours ICU pour les sujets témoins, 799 d'entre eux ayant des enregistrements associés dans la base de données cartographique. Chaque enregistrement est un ensemble de signaux physiologiques mesurés et possède sa propre durée qui représente la durée du séjour du patient en réanimation. On note qu'une durée du séjour élevée permet de faire des prédictions plus tôt car elle donne l'opportunité d'explorer le comportement du métabolisme bien avant l'épisode de FA. Sur la FIGURE 20, les histogrammes des différentes durées de séjour disponibles dans la base de données médicale sont présentés à la fois pour le groupe de patients FA et pour le groupe témoins. La durée du séjour moyenne pour tous les sujets (avec et sans FA) est de 41h avec un écart-type de 37h.

Figure 25 : Histogrammes des durées de séjour pour les patients avec et sans FA.



Les types de signaux considérés doivent également être liés à la FA, plus généralement, ils doivent être liés à la circulation sanguine car elle est assurée par le cœur. De plus, les valeurs prises par le signal doivent être continues car une application d'analyse statistique est prévue à ce signal. Par conséquent, tous les types de signaux avec des valeurs discrètes telles que les signaux d'état sont exclus de notre étude. En tenant compte de tous ces facteurs, une sélection de 10 types de signaux semblent la plus adéquate :

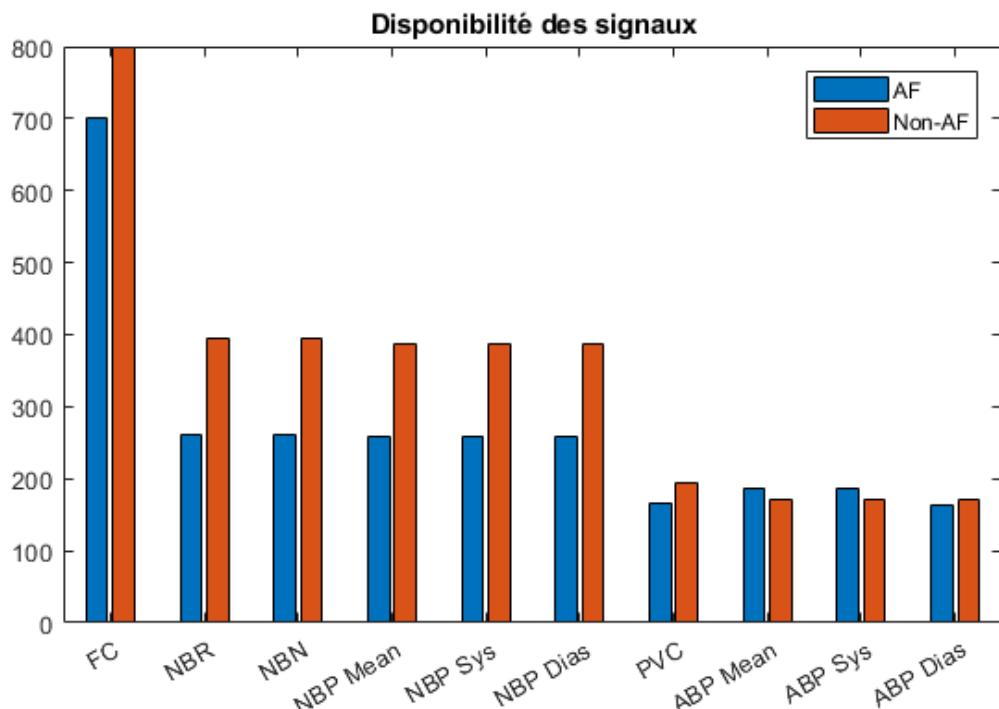
- ❖ Fréquence cardiaque (FC),
- ❖ Nombre de battements rythmés (NBR),
- ❖ Nombre de battements normaux (NBN),
- ❖ Moyenne de la tension artérielle non invasive (NBP Mean),

⁴⁰ La fibrillation auriculaire, aussi appelée fibrillation atriale, est un trouble du rythme cardiaque. En temps normal, oreillettes et ventricules cardiaques se contractent sous l'effet d'une impulsion électrique, au rythme régulier de 60 à 100 battements par minute (au repos)

- ❖ Tension artérielle systolique non invasive (NBP Sys),
- ❖ Tension artérielle diastolique non invasive (NBP Dias),
- ❖ Pression veineuse centrale (PVC),
- ❖ Moyenne de la tension artérielle (ABP Mean),
- ❖ Tension artérielle systolique (ABP Sys),
- ❖ Tension artérielle diastolique (ABP Dias)

La FIGURE 21 montre la disponibilité pour chaque type de signal. Le nombre moyen de types de signaux disponibles par sujet est de 4.06 pour les sujets atteints de FA et de 4.32 pour les sujets témoins, 102 sujets ayant tous leurs signaux disponibles. On rappelle que, pour les sujets du groupe FA, seules les parties des signaux précédant le début de l'épisode de FA sont conservées pour être traitées ultérieurement, c'est-à-dire pour effectuer la prédiction et non le diagnostic de la FA.

Figure 26 : Statistiques de disponibilité des signaux physiologiques.

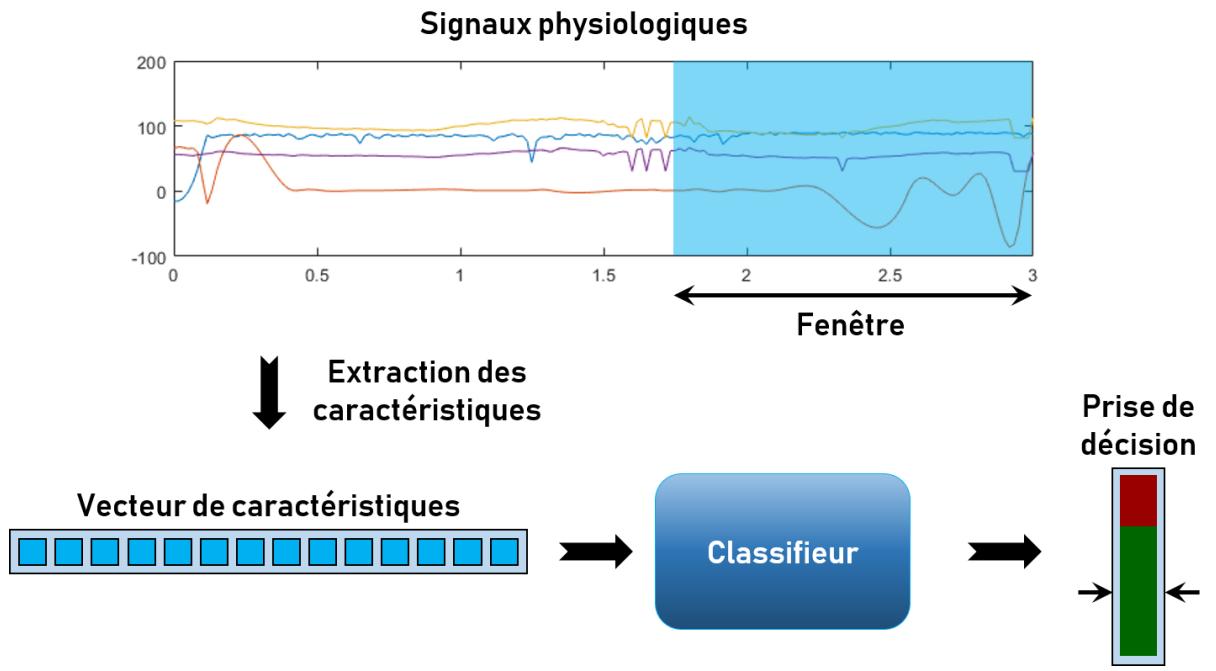


V.2.3. PRESENTATION DE L'APPROCHE

La méthode présentée dans ce chapitre consiste à analyser des signaux physiologiques d'un sujet dans le but d'en forger une prédiction sur l'avenir sanitaire de celui-ci. Cette méthode est composée de plusieurs étapes. La première consiste à extraire des caractéristiques des signaux pour adapter les informations à l'entrée du classifier. Cette étape permet au classifier de traiter des caractéristiques pré calculées des signaux au lieu des signaux eux-mêmes et sera détaillée dans la section. La deuxième étape consiste à se servir de la théorie des fonctions de croyance dans l'optique de construire un classifier ou prédicteur. Dans ce cadre, chaque caractéristique précédemment extraite sera utilisée, en se référant à une base d'apprentissage, comme une source d'information représentée par une distribution de masses. La troisième étape consiste à combiner les différentes distributions de masses obtenues des caractéristiques pour en déduire une nouvelle plus informative et rassemblant les connaissances apportées par toutes les caractéristiques calculées à partir

des signaux physiologiques. La dernière étape consiste à prendre une décision basée sur la distribution de masse obtenue qui est assurée par une transformation de l'approche évidentielle en une approche probabiliste. La méthode globale est illustrée dans la FIGURE 22.

Figure 27 : Processus de prédiction de la FA.



A. L'EXTRACTION DE CARACTERISTIQUES

L'extraction de caractéristiques est une étape-clé des problèmes de reconnaissance de formes. Elle est notamment une étape préalable à la classification, permettant de décrire l'objet à classer grâce à un ensemble d'attributs discriminants (I. Guyon, 2008). Ce problème réside dans une large variété de domaines d'applications allant de la reconnaissance vocale dans le domaine de la robotique au traitement du signal et de l'image dans le domaine médical.

Dans ce travail de mémoire, nous nous intéressons au domaine médical, et plus particulièrement à la représentation de signaux physiologiques pour le pronostic et le diagnostic médical. Pour mieux analyser les signaux physiologiques d'un sujet, il est nécessaire d'extraire des caractéristiques de ces signaux. Il en résulte un vecteur de caractéristiques dont chaque composante représente un critère numérique bien défini décrivant l'état du sujet observé. Dans ce chapitre, nous nous intéressons aux indicateurs statistiques, choisis pour leur forte dépendance aux différents schémas des signaux physiologiques de l'être humain. Ainsi, les sursauts, les augmentations ou diminutions prolongées, les fortes variations ou la stabilité des signaux physiologiques sont facilement détectés et quantifiés par ces caractéristiques. Les caractéristiques extraites sont les suivantes :

Équation 8 : moyenne de donnée extraite

$$x_{i,F(j-1)+1}(t) = \frac{1}{N} \sum_{r=1}^N s_{i,j}(t-N+r),$$

Équation 9 : Ecart-type

$$x_{i,F(j-1)+2}(t) = \sqrt{\frac{1}{N} \sum_{r=1}^N (s_{i,j}(t-N+r) - x_{i,F(j-1)+1}(t))^2},$$

Équation 10 : Skewness⁴¹

$$x_{i,F(j-1)+3}(t) = \frac{1}{N} \sum_{r=1}^N \left(\frac{s_{i,j}(t-N+r) - x_{i,F(j-1)+1}(t)}{x_{i,F(j-1)+2}(t)} \right)^3,$$

Équation 11 : Kurtosis⁴²

$$x_{i,F(j-1)+4}(t) = \frac{1}{N} \sum_{r=1}^N \left(\frac{s_{i,j}(t-N+r) - x_{i,F(j-1)+1}(t)}{x_{i,F(j-1)+2}(t)} \right)^4,$$

Équation 12 : Valeur absolue de la différence

$$x_{i,F(j-1)+5}(t) = \frac{1}{N} \sum_{r=2}^N \left| s'_{i,j}(t-N+r) \right|,$$

Équation 13 : Écart-type de la différence

$$x_{i,F(j-1)+6}(t) = \sqrt{\frac{1}{N} \sum_{r=2}^N \left(s'_{i,j}(t-N+r) - \frac{s_{i,j}(t) - s_{i,j}(t-N+1)}{N} \right)^2},$$

Équation 14 : Skewness de la différence

$$x_{i,F(j-1)+7}(t) = \frac{1}{N} \sum_{r=2}^N \left(\frac{Ns'_{i,j}(t-N+r) - s_{i,j}(t) + s_{i,j}(t-N+1)}{Nx_{i,F(j-1)+6}(t)} \right)^3,$$

Équation 15 : Kurtosis de la différence

⁴¹ En théorie des probabilités et statistique, le coefficient d'asymétrie (**skewness** en anglais) correspond à une mesure de l'asymétrie de la distribution d'une variable aléatoire réelle.

⁴² En théorie des probabilités et en statistique, le **kurtosis**, aussi traduit par coefficient d'acuité1, coefficient d'aplatissement et degré de vousure, est une mesure directe de l'acuité et une mesure indirecte de l'aplatissement de la distribution d'une variable aléatoire réelle

$$x_{i,F(j-1)+8}(t) = \frac{1}{N} \sum_{r=2}^N \left(\frac{Ns'_{i,j}(t-N+r) - s_{i,j}(t) + s_{i,j}(t-N+1)}{Nx_{i,F(j-1)+6}(t)} \right)^4,$$

Équation 16 : Moyenne de la valeur minimale sur une fenêtre glissante

$$x_{i,F(j-1)+9}(t) = \frac{1}{N - [N/5]} \sum_{r=1}^{N-[N/5]} \min_{0 \leq \epsilon \leq [N/5]} (s_{i,j}(t-N+r+\epsilon)),$$

B. LA CLASSIFICATION

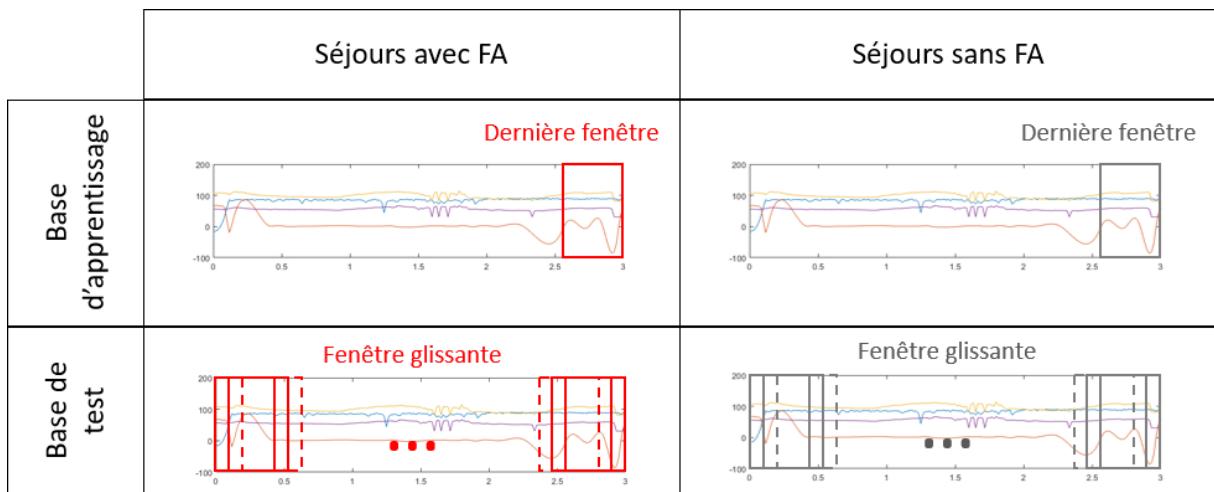
La classification est un domaine de l'intelligence artificielle où le but est d'attribuer un label à une observation. Ce label est choisi parmi un certain nombre de classes ou clusters possibles. On distingue la classification non-supervisée et la classification supervisée. La classification non-supervisée ou clustering consiste à regrouper des instances en se basant sur un critère de dissimilarités entre elles.

Dans notre étude, il s'agit d'un cas de classification supervisée puisque l'on dispose d'une base d'apprentissage avec des sujets labellisés. Le processus de prédiction en temps réel de l'épisode de la FA commence par la mesure des signaux physiologiques du sujet. A chaque instant t et pour chaque signal mesuré, la fenêtre de longueur N minutes se terminant à t est envoyée à la tâche d'extraction de caractéristiques qui donne un vecteur de caractéristiques $\mathbf{x}(t)$. Ce vecteur de caractéristiques est ensuite utilisé comme entrée au classifier qui génère un niveau de risque d'occurrence d'un épisode de la FA. Deux classes sont considérées, la classe "Sujet témoin" non prévu de subir un épisode de FA, labellisée 0, et la classe "Sujet FA" prévu de subir un épisode de FA, labellisée 1. Le classifier proposé est défini dans le cadre de la théorie des fonctions de croyance en utilisant une base de données d'apprentissage. L'avantage de cette théorie réside dans sa capacité à modéliser le problème de manière décente même avec des données manquantes ou peu fiables, sans nécessité d'une base d'apprentissage totalement équilibrée. La construction de la base d'apprentissage et la conception du classifier sont présentées ci-dessous :

B.1 DIVISION DE LA BASE DE DONNEES

Afin d'entraîner le classifier, une base d'apprentissage est construite en sélectionnant au hasard une portion $\delta\%$ ($0 \leq \delta \leq 100$) de l'ensemble des 701 séjours à l'USI des sujets ayant eu la FA et une portion pareille des 799 séjours de ceux ne l'ayant pas eue. La base de données A est utilisée dans le domaine de la théorie des fonctions de croyance pour construire le classifier. Dans la phase de test, les sujets des indices dans I_T sont considérés avec une fenêtre glissante. A chaque instant t , le vecteur de caractéristiques $\mathbf{x}_i(t)$, $i \in I_T$, calculé à t est utilisé comme entrée du classifier, comme illustré dans la FIGURE 3.7

Figure 28 : Répartition des fenêtres selon les bases d'apprentissage et de test.



B.2. ATTRIBUTION DES MASSES

La méthode de classification proposée est une méthode évidentielle qui consiste à utiliser chaque caractéristique comme une source d'information indépendante dans le cadre de la théorie des fonctions de croyance. On adapte alors une distribution normale pour chaque caractéristique et pour chaque classe. Cette étape a pour but de concrétiser les données pour une mesure de similarité plus simple avec un vecteur de caractéristiques à classifier. Cette adaptation est assurée par la méthode de la maximisation de la fonction de vraisemblance et elle consiste à estimer les paramètres d'une distribution normale, à savoir, l'espérance et l'écart-type. On obtient alors les paramètres suivants $\forall u \in \{1, 2, \dots, 190\}$,

Équation 17 : démonstration de la distribution normale

$$\hat{\mu}_u^{(1)} = \frac{\sum_{i \in I_u} x_{i,u}(T_i) y_i}{\sum_{i \in I_u} y_i},$$

$$\hat{\mu}_u^{(0)} = \frac{\sum_{i \in I_u} x_{i,u}(T_i) (1 - y_i)}{\sum_{i \in I_u} 1 - y_i},$$

$$\hat{\sigma}_u^{(1)} = \sqrt{\frac{\sum_{i \in I_u} (x_{i,u}(T_i) - \hat{\mu}_u^{(1)})^2 y_i}{\sum_{i \in I_u} y_i}},$$

$$\hat{\sigma}_u^{(0)} = \sqrt{\frac{\sum_{i \in I_u} (x_{i,u}(T_i) - \hat{\mu}_u^{(0)})^2 (1 - y_i)}{\sum_{i \in I_u} 1 - y_i}},$$

B.3. SELECTION DES CARACTERISTIQUES

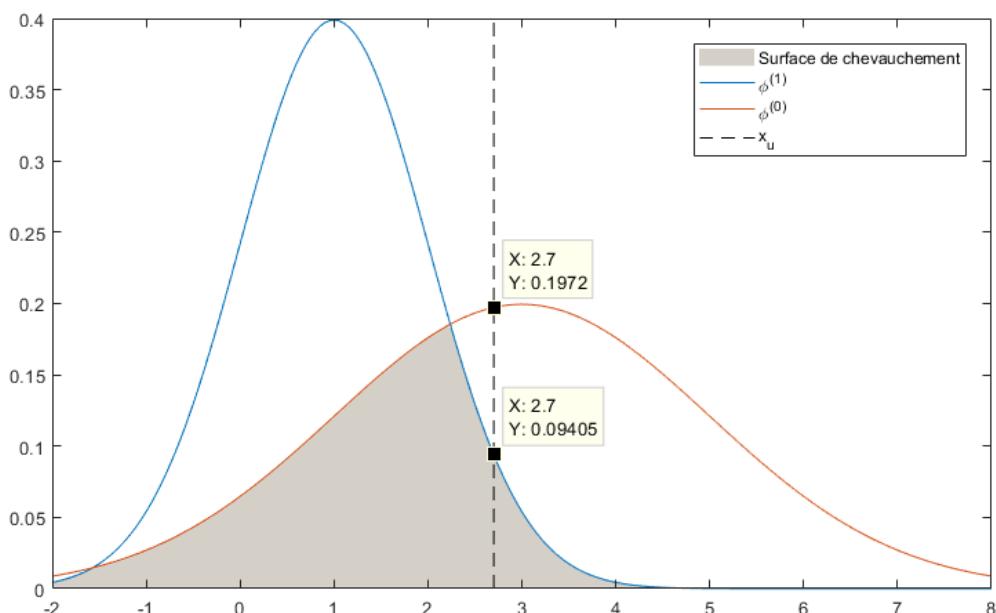
Dans le but d'améliorer la modélisation des données statistiques lors de l'attribution des masses, la valeur de l'ignorance peut être modifiée selon la disponibilité et la quantité d'information traduite par la base d'apprentissage. Certaines sources d'information présentent une masse d'ignorance très importante. Dans le cas de notre étude, la majorité des sources d'information ne sont utiles que sur de rares plage de valeurs des

caractéristiques. Ainsi, une masse d'ignorance élevée permet d'évanouir leur participation lors du processus de combinaison des sources d'information et dans la construction de la fonction de masse finale. Il n'en demeure pas moins vrai que leur nombre élevé finit quand même par perturber les informations issues des sources fiables dans la combinaison. Pour pallier à ce mal, un terme β ($0 \leq \beta \leq 1$) est introduit. Les sources ayant une masse d'ignorance supérieure à ce seuil β seront exclues de la combinaison et n'auront donc aucun effet sur la masse finale. Mathématiquement, ceci est équivalent à modifier cette masse d'ignorance et la remettre à 1 ce qui donne l'équation suivante :

Équation 18 : la fonction masse

$$m_u^{i,t}(\Omega) = \begin{cases} (\mathcal{S}_u)^\alpha & \text{si } (\mathcal{S}_u)^\alpha < \beta, \\ 1 & \text{si } (\mathcal{S}_u)^\alpha \geq \beta, \end{cases}$$

Figure 29 : Exemple de distribution des masses.



En effet, une fonction de masse qui attribue l'unité entière à la masse de l'ignorance totale n'apporte aucune information et se trouve donc l'élément neutre de l'opération de combinaison. On note que la sélection des caractéristiques peut être ignorée en considérant $\beta = 1$.

B.4. COMBINAISON ET PRISE DE DECISION

Dans le chapitre précédent, plusieurs méthodes de combinaison ont été introduites. Le choix de la méthode de combinaison est une étape cruciale pour le bon fonctionnement du classifier. Elle définit comment les informations seront mises en commun et influe sur la certitude prononcée par la fonction de masse finale. Dans le cas de notre étude, étant donné que deux événements (0 et 1) seulement sont possibles, la distribution des masses concerne les ensembles d'événements suivants : ;, {0}, {1} et $\Omega = \{0, 1\}$. De plus, on tient à éliminer l'impossibilité pour éviter les conflits souvent trop élevés dans l'analyse de certains

types de données et surtout médicaux. Ainsi, la distribution de masse est assimilée à un triplé de valeurs réelles comprises entre 0 et 1 et dont la Somme vaut l'unité.

A chaque itération, la règle de combinaison de YAGER permet de combiner deux fonctions de masses différentes qui seront représentées par deux règles différentes que l'on va placer de manière perpendiculaire l'une à l'autre dans le but de mieux illustrer la méthode de combinaison. Le calcul des masses de la source d'information résultante de la combinaison est essentiellement basé sur les produits des masses de chacune des deux sources La prise de décision est basée sur la transformation de la fonction de masse finale en distribution de probabilité. Ceci est fait en utilisant la transformation pignistique, définie au chapitre précédent, qui consiste à répartir la masse d'un ensemble équitablement sur ses éléments.

V.3. MODELISATION D'UN SYSTEME

Pendant plusieurs décennies, le monde informatique a toujours rêvé d'un processus qui puisse garantir le développement efficace de logiciels de qualité, valable quel que soit la grandeur et la complexité du projet, et présentant de bonnes pratiques adaptées à la méthode en question, surtout que, de nos jours, les logiciels demandés sont de plus en plus imposants et exigeants qu'auparavant. Le processus de modélisation vise à obtenir une solution acceptable du système informatique. La solution finalement retenue n'est pas obtenue en une seule itération. Plusieurs étapes sont nécessaires ; ces étapes successives permettent de raffiner le niveau de détails du système à réaliser. Les premières étapes donnent une vision à très gros grains et permettent d'avancer dans la compréhension du problème.

V.3.1. LE PROCESSUS UNIFIE

Le processus uniifié (PU), ou « ***unified process (UP)*** » en anglais, ou « ***Unified Software Development Process (USDP)*** » est un processus de développement logiciel itératif, centré sur l'architecture, piloté par des cas d'utilisation et orienté vers la diminution des risques. C'est un patron de processus pouvant être adapté à une large classe de systèmes logiciels, à différents domaines d'application, à différents types d'entreprises, à différents niveaux de compétences et à différentes tailles de l'entreprise. Le document suivant présente sous la forme d'une note les concepts associés à ce processus (developpez.com, 2023)

Elle se caractérise par une démarche itérative et incrémentale, pilotée par les cas d'utilisation, et centrée sur l'architecture et les modèles UML. Elle définit un processus intégrant toutes les activités de conception et de réalisation au sein de cycles de développement composés d'une phase de création, d'une phase d'élaboration, d'une phase de construction et d'une phase de transition, comprenant chacune plusieurs itérations (Roques, 2017).

Le processus uniifié est une méthode de développement de logiciel caractérisée par :

- ❖ un pilotage par les cas d'utilisation,
- ❖ une démarche centrée sur l'architecture,
- ❖ une approche basée sur les modèles, et en particulier les modèles UML,
- ❖ Une approche itérative et incrémentale visant en priorité à réduire les incertitudes.

La méthode UP se base sur quatre phases :

❖ **Analyse des besoins** : L'analyse des besoins donne une vue du projet sous forme de produit fini. Cette phase porte essentiellement sur les besoins principaux (du point de vue de l'utilisateur), l'architecture générale du système, les risques majeurs, les délais et les coûts

On met en place le projet. Elle répond aux questions suivantes :

- Que va faire le système ? Par rapport aux utilisateurs principaux, quels services va-t-il rendre ?
- Quelle va être l'architecture générale (cible) de ce système ?
- Quels vont être : les délais, les coûts, les ressources, les moyens à déployer ?

❖ **Élaboration** : L'élaboration reprend les éléments de la phase d'analyse des besoins et les précise pour arriver à une spécification détaillée de la solution à mettre en œuvre. L'élaboration permet de préciser la plupart des cas d'utilisation, de concevoir l'architecture du système et surtout de déterminer l'architecture de référence.

Au terme de cette phase, les chefs de projet doivent être en mesure de prévoir les activités et d'estimer les ressources nécessaires à l'achèvement du projet. Les tâches à effectuer dans la phase élaboration sont les suivantes :

- Créer une architecture de référence ;
- Identifier les risques, ceux qui sont de nature à bouleverser le plan, le coût et le calendrier ;
- Définir les niveaux de qualité à atteindre ;
- Formuler les cas d'utilisation pour couvrir les besoins fonctionnels et planifier la phase de construction ;
- Elaborer une offre abordant les questions de calendrier, de personnel et de budget.

❖ **Construction**: La construction est le moment où l'on construit le produit.

L'architecture de référence se métamorphose en produit complet.

Le produit contient tous les cas d'utilisation que les chefs de projet en accord avec les utilisateurs ont décidé de mettre au point pour cette version. Transformer l'architecture de référence en produit exécutable tout en veillant à respecter son intégrité.

❖ **Transition** Le produit est en version bêta. Un groupe d'utilisateurs essaye le produit et détecte les anomalies et défauts. Cette phase suppose des activités comme la formation des utilisateurs clients, la mise en œuvre d'un service d'assistance et la correction des anomalies constatées

Chaque phase est constituée d'une succession d'activités. Les activités de la méthode UP sont :

❖ **Expression des besoins** : L'expression des besoins comme son nom l'indique, permet de définir les différents besoins :

- Inventorier les **besoins principaux** et fournir une liste de leurs fonctions ;
- Recenser les **besoins fonctionnels** (du point de vue de l'utilisateur) qui conduisent à l'élaboration des modèles de cas d'utilisation ;
- Appréhender les **besoins non fonctionnels** (technique) et livrer une liste des exigences.

Le modèle de cas d'utilisation présente le système du point de vue de l'utilisateur et représente sous forme de cas d'utilisation et d'acteur, les besoins du client

- ❖ **Analyse** : L'objectif de l'analyse est d'accéder à une compréhension des besoins et des exigences du client. Il s'agit de livrer des spécifications pour permettre de choisir la conception de la solution. Un modèle d'analyse livre une spécification complète des besoins issus des cas d'utilisation et les structures sous une forme qui facilite la compréhension (scénarios), la préparation (définition de l'architecture), la modification et la maintenance du futur système. Il s'écrit dans le langage des développeurs et peut être considéré comme une première ébauche du modèle de conception.
- ❖ **Conception** : La conception permet d'acquérir une compréhension approfondie des contraintes liées au langage de programmation, à l'utilisation des composants et au système d'exploitation. Elle détermine les principales interfaces et les transcrit à l'aide d'une notation commune. Elle constitue un point de départ à l'implémentation :
 - Elle décompose le travail d'implémentation en sous-système ;
 - Elle crée une abstraction transparente de l'implémentation.
- ❖ **Implémentation** : L'implémentation est le résultat de la conception pour implémenter le système sous forme de composants, c'est-à-dire, de code source, de scripts, de binaires, d'exécutables et d'autres éléments du même type. Les objectifs principaux de l'implémentation sont de planifier les intégrations des composants pour chaque itération, et de produire les classes et les sous-systèmes sous forme de codes sources.
- ❖ **Tests** : Les tests permettent de vérifier des résultats de l'implémentation en testant la construction. Pour mener à bien ces tests, il faut les planifier pour chaque itération, les implémenter en créant des cas de tests, effectuer ces tests et prendre en compte le résultat de chacun.

V.3.2. LANGUAGE DE MODELISATION

Les langages de modélisation ont été créés dans le but de simplifier la résolution de problèmes de programmation mathématique. Le principe de base sous-jacent aux langages de modélisation traditionnels est la découverte que de nombreux problèmes de programmation mathématique peuvent être exprimés dans un langage informatique dont la syntaxe est proche de la présentation standard de ces problèmes dans les livres et les articles scientifiques. Ces langages fournissent généralement plusieurs types de données tels que des tableaux et des ensembles, ainsi que des équivalents des langages informatiques pour les notations algébriques traditionnelles (IBM, 2023)

A. LE LANGUAGE UML : UNIFIED MODELING LANGUAGE

Avant même l'introduction d'UML dans le développement logiciel, le domaine de la programmation orientée objet (OOP) était déjà en plein essor. Ce style de programmation est basé sur le concept que tout est un objet : les éléments constitutifs d'un programme sont des objets qui interagissent les uns avec les autres. Les messages envoyés dans les deux sens sont également constitués d'objets. Chaque objet individuel est un exemple de sa classe supérieure. La classe elle-même agit également comme un objet et détermine le comportement des instances d'objet qu'elle contient. Les objets sont constitués de données

et de code. L'objet organise les données en zones, également appelées attributs. Le code détermine leur procédure ou leur méthode.

De la fin des années 1980 aux années 1990, de nombreuses méthodes et langages pour la représentation de la POO ont été développés et mis en œuvre. Il en est résulté une variété de méthodes très dissemblables. Pour unifier ces langages, les trois développeurs James **Rumbaugh, Grady Booch** et **Ivar Jacobson** ont décidé de fusionner plusieurs langages existants en un langage commun et standardisé.

Les trois avaient déjà créé leurs propres méthodes de développement logiciel orienté objet :

- ❖ La méthode Booch
- ❖ L'Object modeling technique (OMT)
- ❖ L'object-oriented software engineering method (OOSE)

UML devrait définir la sémantique pour la représentation de ces méthodes comme langage de modélisation. Sous le nom de "UML Partners", les développeurs ont commencé à travailler sur l'achèvement d'UML dans une équipe en 1996. Ils l'ont ensuite remis au Object Management Group (OMG), qui a introduit la version 1.1 du langage de modélisation unifié comme norme en 1997.

V.4. PRESENTATION DES ARCHITECTURE DU NOUVEAU SYSTEME

Cette partie contiendra les détails des étapes de modélisation et d'extraction de modèle précis avec le formalisme du data-minig en faisant l'analyse avec UML, donnant ainsi un squelette du système à réaliser en détaillant le fonctionnement de celui-ci par différents schémas. À partir de l'architecture obtenue, il s'agirait d'une implémentation physique du système pour simuler certaines des fonctionnalités du nouveau système. Pour la présentation de l'architecture UML, nous aurons à présenter le diagramme d'objet, le diagramme des composant et celui du déploiement vu que ce trois donne déjà une vue d'ensemble sur la constitution du système.

V.4.1. EXIGENCES FONCTIONNELLES

Il s'agit d'une étape essentielle au début de chaque processus de développement, elle vise à assurer la création d'un système adéquat, et son but est la description générale des caractéristiques du système. Le système à concevoir doit répondre aux exigences matérielles et aussi logicielles, il s'agira de la mise en place d'une surveillance et contrôle automatique des patients centrée autour de l'intelligence artificiel, pilotée par une application installée sur le serveur.

Exigences fonctionnelles matérielles

Les exigences que les parties matérielles doivent prendre en compte comprennent :

❖ **Détection de la détérioration de l'état de santé du patient:** en fonction des informations reçues du patient, les algorithmes décident oui ou non de la détérioration de l'état de santé du patient. En tant que cœur de la partie matérielle, les appareils connectés envoient l'information au logiciel de contrôle, qui à son tour alerte un soignant sur la prise en charge du patient.

Cette détection de la détérioration de l'état de santé du patient s'effectue à travers quatre activités dont :

Prélèvement de température : Il mesure la température à des intervalles de temps très courts et les transmet au microcontrôleur, et déclenche une demande d'aide en cas d'augmentation ou de baisse de la température du patient par rapport à la température normale.

Prise du rythme cardiaque : mesure la fréquence cardiaque et envoie ses informations au microcontrôleur, et déclenche la demande d'aide en cas de dysfonctionnement.

Aide : un appui sur le bouton d'urgence envoie une commande au microcontrôleur signifiant qu'il y a un besoin urgent, de l'aide de l'infirmière et le microcontrôleur envoie également un ordre au logiciel de contrôle, qui affiche les coordonnées du patient et déclenche également l'alarme.

Affichages : récupération des informations patient à partir du serveur y compris sa localisation au sein de l'hôpital ou dans son habitat, et ainsi permettre au service d'urgence de l'hôpital d'appréhender bien avant l'arrivée du patient au lieu de soin.

Exigences fonctionnelles du logiciel

Prise en charge du patient : à son arrivée à l'hôpital, le patient est d'abord inscrit dans le système hospitalier, selon son état il peut être hospitalisé à l'hôpital dans une chambre selon ses moyens et selon son cas. Lorsque son état se stabilise, il libère la chambre d'hôpital.

Prise en charge infirmière : lorsqu'il est engagé à l'hôpital, il doit être inscrit dans le système hospitalier, il peut alors démarrer le planning de suivi du patient.

Gestion des médecins : le médecin également une fois engagé il doit être inscrit dans le système de l'hôpital, il peut alors via son téléphone surveiller les patients et recevoir les notifications demandant son intervention pour certains cas de patient.

Surveillance des patients : à des moments précis, l'administrateur système surveille les patients, puis il fait ce que nous appelons les chambres tournantes à travers différentes caméras placées dans les chambres des patients en même temps, il peut observer la température et la fréquence cardiaque du patient concerné, et faire ressortir un rapport sur l'évolution de l'état de santé pour prendre une bonne décision.

Notification par SMS : après que l'administrateur a élaboré l'horaire de travail des infirmières selon différents critères, les messages seront envoyés aux différentes infirmières afin qu'elles aient l'attribution des tâches pendant une période donnée.

Il est ainsi possible de faire des statistiques sur une maladie récurrente et de déterminer les régions touchées par celle-ci, selon la source, afin de vulgariser les stratégies pour la combattre.

V.5. CONCLUSION

Dans ce chapitre, nous avons passé en revue une grande partie des techniques proposées pour traiter les problèmes de la détection et de la quantification des lésions des artères coronaires ainsi que l'extraction des structures vasculaires. Nous avons départagé les différents techniques en nous appuyant sur les caractéristiques qu'elles utilisent pour accomplir ces tâches.

Malgré la diversité des méthodes proposées dans l'état de l'art pour l'aide au diagnostic des maladies coronariennes, ce domaine n'a pas encore fourni des schémas entièrement automatiques et fiables. Avec l'introduction des techniques d'apprentissage profond, nous constatons une meilleure caractérisation des structures vasculaires et les lésions coronariennes.

Dans la suite de ce manuscrit, nous présenterons nos contributions scientifiques. Nous détaillerons notre méthode d'extraction des artères coronaires et notre approche pour la détection de la sténose et la classification des plaques. Nous serons aussi amenés à tirer profit des multiples méthodes de l'état de l'art que nous avons présenté dans ce chapitre, pour fournir un pipeline complet et automatisé pour la prévention des risques cardiovasculaires.

CHAPITRE VI. LE RESULTAT : CLASSIFICATION MULTI-LABEL PAR ENSEMBLE POUR LE SUIVI MEDICAL

VI.1. INTRODUCTION

La classification multi-label fait partie de l'apprentissage automatique où la même instance peut appartenir à plusieurs classes simultanément. Dans ce cas, les classes ne sont pas exclusives. L'une des méthodes de transformation les plus populaires est la méthode "Binary Relevance" (BR) (M.-L. Zhang, 2018), où chaque label génère un problème de classification binaire. En conséquence, un problème de classification multi-label est donc transformé en différentes tâches de classification binaire dont le nombre est égal au nombre de labels. Cependant, un inconvénient majeur de cette méthode est qu'elle gère les labels séparément et qu'elle ignore donc les dépendances entre eux.

Cela peut entraîner certaines conséquences, telles que la dégradation des performances et la perte d'informations importantes. Si la méthode du "Label Powerset" (LP) prend en compte les dépendances existantes entre les labels, son utilisation sur des bases de données caractérisées par un grand nombre de labels conduit à certains inconvénients. En effet, à mesure que le nombre de labels augmente, le nombre de sous-ensembles de labels augmente de façon exponentielle. Afin de surmonter ce problème, la méthode "Radom k-label sets" (RAkEL) est introduite (G. Tsoumakas, 2010). Cette méthode considère un certain nombre de sous-ensembles aléatoires de labels de taille k et utilise la méthode LP sur chacun de ces sous-ensembles. Avec un nombre approprié de k -label sets ou ensemble de labels de taille k , cette méthode a tendance à gérer les corrélations entre les labels.

Dans la littérature, plusieurs études sont présentées sur les méthodes d'ensemble, certains d'entre elles utilisent la corrélation entre les labels, d'autres sont basées sur la sélection de caractéristiques (C. Shi, 2011).

VI.2. APPROCHE PROPOSEE : COLEX

L'apprentissage des relations entre les labels est un problème difficile en classification multi label. En effet, les classes se chevauchent et sont corrélées, comme l'association d'un label à une instance peut fournir des informations sur l'appartenance de cette instance à d'autres labels. En l'absence de relations entre les labels, l'aspect multi-label devient sans intérêt et les données peuvent être traitées par une méthode mono-label sans perte de généralité. Différentes approches ont été proposées dans la littérature pour modéliser la corrélation entre les labels (C. Shi, 2011).

Les relations entre les labels peuvent être considérées sous forme de relations binaires entre chaque paire d'labels, ou bien des relations d'ordre supérieur entre chaque label et les labels restants. Les relations d'ordre supérieur sont plus complexes à représenter que les relations binaires, qui peuvent être mesurées, e.g., par la probabilité conditionnelle d'un label étant donnée l'existence d'une autre.

VI.2.1. LA METHODE DE RAKEL

On rappelle que la méthode Label Powerset (LP) est une méthode de transformation qui considère chaque sous-ensemble de labels comme une classe unique dans un problème

mono-label équivalent. Ainsi, s'il y a q labels dans l'espace des labels Y du problème multi-label, on peut obtenir un maximum de 2^q classes dans le problème correspondant en mono-label. Pour une nouvelle instance à classifier, toute méthode de classification mono-label peut être utilisée pour prédire sa classe. La sortie du classifier mono-label est alors remplacée par le sous-ensemble équivalent de labels pour obtenir la sortie du classifier multi-label. La méthode LP offre un moyen simple d'utiliser la méthode mono-label pour résoudre les problèmes de classification multi-label. Cependant, elle présente quelques défauts, en particulier lorsque l'espace des labels contient un nombre important de labels. Dans ce cas, le nombre d'instances associées à chaque classe diminue et rend la méthode mono-label incapable de généraliser correctement pour les nouvelles instances.

VI.2.2. LE COEFFICIENT DE CORRELATION BINAIRE Φ

Afin de mieux introduire le coefficient de corrélation binaire, on représente les labels de chaque instance x_i de la base de données multi-label par un vecteur binaire v_i contenant une composante $\lambda_{i,j}$ pour chaque label du domaine des labels Y qui prend la valeur $\lambda_{i,j} = 1$ si cette instance est labélisée par j et 0 sinon. Dans le cadre de variables binaires, la mesure de la corrélation entre les variables est généralement effectuée en utilisant le coefficient φ . C'est l'équivalent du coefficient de Pearson appliqué à la référence des variables binaires. Soit D une base de données multi-label étiquetée telle que

Équation 19 : coefficient de corrélation binaire

$$\mathcal{D} = \{(x_i, Y_i) \mid x_i \in \mathcal{X}, Y_i \subseteq \mathcal{Y}, 1 \leq i \leq M\}$$

En utilisant ces notations, le coefficient de corrélation binaire φ entre les labels j et j_0 peut être calculé par :

Équation 20 : coefficient de corrélation binaire entre label

$$\Phi_{j,j'} = \frac{c_{(j,j')} c_{(-j,-j')} - c_{(-j,j')} c_{(j,-j')}}{\sqrt{(c_{(j,j')} + c_{(-j,j')})(c_{(j,-j')} + c_{(-j,-j')})(c_{(j,j')} + c_{(j,-j')})(c_{(-j,j')} + c_{(-j,-j')})}}.$$

Sous l'hypothèse de variables indépendantes, ce coefficient est égal à zéro. En fait, plus il est proche de zéro, moins les labels sont corrélés. Contrairement au coefficient de Pearson, le coefficient φ n'est pas limité entre -1 et 1. Dans notre travail, nous utilisons ce coefficient comme indicateur de corrélation pour chaque paire de labels possible dans l'espace Y , comme il sera montré ci-dessous.

VI.2.3. FACTEUR DE CORRELATION GLOBALE

La méthode proposée dans ce chapitre consiste à sélectionner un certain nombre de k label sets ayant une forte corrélation entre leurs labels. En effet, avec des corrélations élevées, ces labels sont plus susceptibles de tomber ensemble pour les nouvelles instances. En les prenant ensemble, nous évitons d'avoir des prédictions contradictoires par différents prédicteurs. Afin de mesurer la corrélation des labels au sein d'un k -label set particulier, nous calculons

d'abord la corrélation entre toute les paires de labels dans celui-ci en utilisant le coefficient ϕ présenté dans la section précédente.

Ce coefficient est égal à zéro lorsqu'il n'y a pas de corrélation entre les deux labels considérés et a une valeur absolue élevée lorsqu'il existe une forte corrélation entre eux. Une matrice de corrélation symétrique Φ est construite à partir de la valeur absolue de ce coefficient de corrélation :

Équation 21 : Facteur de corrélation globale

$$\Phi(\ell_j, \ell_{j'}) = |\phi_{j,j'}| \quad \forall j, j' \in \{1, 2, \dots, q\}, j \neq j'$$

$$\Phi(\ell_j, \ell_j) = 0 \quad \forall j \in \{1, 2, \dots, q\},$$

$$\Phi = \begin{pmatrix} 0 & |\phi_{1,2}| & \cdots & |\phi_{1,q}| \\ |\phi_{2,1}| & 0 & \cdots & |\phi_{2,q}| \\ \vdots & & \ddots & \\ |\phi_{q,1}| & |\phi_{q,2}| & \cdots & 0 \end{pmatrix},$$

VI.2.4. STRATEGIE D'ECHANGE : METHODE COLEX

La méthode CoLEx ("Correlation based Labelsets by Exchange strategy") est proposée dans le but d'optimiser la sélection des n différents k -labelsets sur lesquels la méthode LP sera appliquée individuellement. Comme son nom l'indique, cette méthode est basée sur des échanges de labels entre les k -labelsets et l'objectif de ces échanges est d'aboutir à une nouvelle sélection des n k -labelsets ayant le facteur de corrélation ζ , le plus élevée.

Concrètement, la méthode consiste à tester chaque échange de labels possible, c'est à dire chaque paire de labels issues de différents labelsets et de déterminer la variation du facteur de corrélation globale $\Delta\zeta$ qui s'en suit. Si ce dernier est strictement positif, l'échange est appliqué et l'algorithme est lancé à nouveau, sinon, on passe au test de l'échange suivant. Il s'agit d'un moyen rapide de trouver une solution pseudo-optimale car l'optimisation absolue est coûteuse en puissance de calcul. De plus, compte tenu du fait que la méthode est basée sur des échanges de labels uniquement, le nombre d'occurrence totale de chaque label dans l'ensemble des n des k -labelsets reste invariant au cours de l'optimisation ce qui empêche d'aboutir à une sélection déséquilibrée entre labels. Cette matrice est initialisée à $S(0)$ en premier lieu en suivant l'algorithme. Elle est construite ligne par ligne en partant de la première colonne et en affectant les labels de sorte que $S(0)_{u,v} = r \quad \forall 1 \leq u \leq n \text{ et } 1 \leq v \leq k$ avec $r = ((u-1).k + v)\%q + 1$ et $\%$ étant l'opérateur modulo. En initialisant cette matrice comme indiquée, l'équilibre entre les labels est garanti et maintenu vu qu'elle contient des occurrences de labels dont l'écart est au plus 1.

Figure 30 : algorithme de la matrice avec la stratégie d'échange

FONCTION Matrice d'initialisation(n, k, q)

S \leftarrow Matrice($n \times k$)

r \leftarrow 1

Pour $u \in 1..n$

Pour $v \in 1..k$

$s_{(u,v)} \leftarrow \ell_r$

$r \leftarrow r + 1$

Si $r = q+1$

$r \leftarrow 1$

FinSi

FinPour

FinPour

Retourner S

La première matrice est issue d'un échange entre les labels `1 et `5 des lignes respectives 1 et 2, mais `1 existe déjà sur la ligne 2 donc cet échange est inutile car il conduit à une configuration illogique. La deuxième est issue d'un échange entre les labels `2 et `5 des lignes respectives 1 et 2, par contre, cet échange aboutit à la variation de la corrélation globale $\Delta\zeta(1,2) \Leftrightarrow (2,1); \mathbf{S}(0)\varphi = -0.3$ ce qui ne satisfait pas les conditions d'application de l'échange. Le troisième échange introduit un doublé de `1 sur la première ligne; cet échange est donc ignoré. Enfin, pour la dernière matrice, on a $\Delta\zeta(3,2) \Leftrightarrow (4,3); \mathbf{S}(0)\varphi = +0.2$; cet échange est pris en compte et on aura alors :

Figure 31 : algorithme de la matrice avec la stratégie d'échange optimisée

FONCTION Optimisation matrice(**S**)

unstable \leftarrow true

Tant que **unstable**

unstable \leftarrow false

Pour $u \in 1..n - 1$

Pour $u' \in u + 1..n$

Pour $v \in 1..k$

Pour $v' \in 1..k$

Si $s_{(u,v)} \notin S_{u'}$ **Et** $s_{(u',v')} \notin S_u$ **Et** $\Delta\zeta_{(u,v) \leftrightarrow (u',v')}(S) > 0$

$\leftarrow s_{(u,v)}$

$s_{(u,v)} \leftarrow s_{(u',v')}$

$s_{(u',v')} \leftarrow aide$

$\leftarrow s_{(u,v)}$

$unstable \leftarrow true$

FinSi

FinPour

FinPour

FinPour

FinTantque

Retourner S

VI.3 RESULTATS ET DISCUSSIONS

VI.3.1 Application de la méthode Colex à la base de données médicale

Dans le chapitre précédent, une base de données multi-label a été construite en se basant sur les données médicales de MIMIC III. Cette base regroupe un ensemble d'admissions à l'unité de soin intensif d'un hôpital pour lesquels un diagnostic incluant une pathologie cardiaque a été enregistré. En préservant la même notation, cette base de donnée sera notée H . Vu que cette base de données a été configurée pour convenir à l'apprentissage du modèle évidentiel, son fort niveau d'indisponibilité de l'ordre de 47% ne présentait pas un inconvénient insurmontable. Cependant, dans une optique d'apprentissage basée sur la méthode SVM, ce n'est plus le cas. En effet, toutes les instances de la base de données doivent avoir toutes leurs caractéristiques disponibles. En éliminant les instances ayant au moins une caractéristique indisponible, la base de donnée se vide littéralement. Alors, pour palier à ce problème, nous avons opté pour une adaptation de la base de donnée qui se fait en deux étapes. La première consiste à sélectionner des caractéristiques selon leurs niveaux de disponibilité respectives et par la suite, la deuxième comprend la suppression de toutes les instances n'ayant pas toutes les caractéristiques disponibles.

VI.3.2 ILLUSTRATION DES PERFORMANCES DE LA METHODE PROPOSEE SUR DES BASES DE DONNEES REELLES MULTI-LABEL

Afin de valider l'efficacité de l'approche CoEx proposée, nous considérons différentes bases de données multi-label bien connues que sont Yeast, Genbase, Scene, Medical, Enron et Emotions (J. Alcalá-Fdez, 2011). La première base de données Yeast contient des informations sur un ensemble de cellules de levure, le but étant de déterminer le site de localisation de chaque cellule. Dans la deuxième, appelée Genbase, il s'agit de déterminer la famille protéique de chaque protéine. Le troisième, Scene, fait référence à des scènes pour lesquelles plusieurs étiquettes peuvent être associées, comme par exemple, une même photographie peut contenir une montagne, la mer et la plaine. Pour la base de données Medical, il s'agit d'associer des labels à un texte libre de rapports médicaux

Tableau 1 : Caractéristiques des bases de données multi-label du monde réel

	Nombre d'instances	Nombre de caractéristiques	Nombre de labels	LCard	LDens	LDiv	PLDiv
Yeast	2417	103	14	4.24	0.30	198	0.082
Genbase	662	1185	27	1.25	0.05	32	0.048
Scene	2407	294	6	1.07	0.18	15	0.006
Medical	978	1449	45	1.25	0.03	94	0.096
Enron	1702	1001	53	3.38	0.06	753	0.442
Emotions	593	72	6	1.87	0.31	27	0.046

Dans ce qui suit, nous présentons les performances de notre méthode et ses avantages par rapport à la méthode RAKEL standard et la méthode exacte pour maximiser la corrélation des label sets. Nous donnons les résultats obtenus pour les deux ensembles de données à des valeurs k et n différentes. Dans le TABLEAU 2, plusieurs métriques de performances sont affichées pour les deux méthodes RAKEL et celle proposée pour la base de données Genbase avec $n = 9$ label sets et avec $k = 3$ labels par label set. Nous pouvons constater une amélioration significative de l'efficacité globale de la tâche de classification dans notre méthode. Cela montre que notre méthode tire parti des corrélations entre labels et réussit à créer un meilleur groupe de label sets. Dans le TABLEAU 3, nous avons utilisé $n = 18$

label sets au lieu de 9 et nous remarquons toujours de meilleurs résultats en utilisant notre méthode. Au contraire, la méthode RAKEL donne de meilleures performances avec ce nombre plus élevé de label sets. Cela peut s'expliquer par le fait que RAKEL a plus de chance d'attraper les corrélations entre labels aléatoirement lorsqu'il y a plus de label sets.

Tableau 2 : Comparaison entre RAKEL et COLEX pour la base de données Genbase k = 3 et n = 9

Métriques	RAKEL	COLEX
↓ Hloss	0.0252 ± 0.0057	0.0112 ± 0.0014
↓ One-Error	0.4154 ± 0.1271	0.1608 ± 0.0228
↓ Coverage	3.8613 ± 1.0213	1.8556 ± 0.2827
↓ Rloss	0.0000 ± 0.0000	0.0000 ± 0.0000
↑ Subset Accuracy	0.4422 ± 0.1470	0.8080 ± 0.0214
↑ Recall	0.4769 ± 0.1460	0.8142 ± 0.0225
↑ Average Precision	0.6534 ± 0.1055	0.8684 ± 0.0165

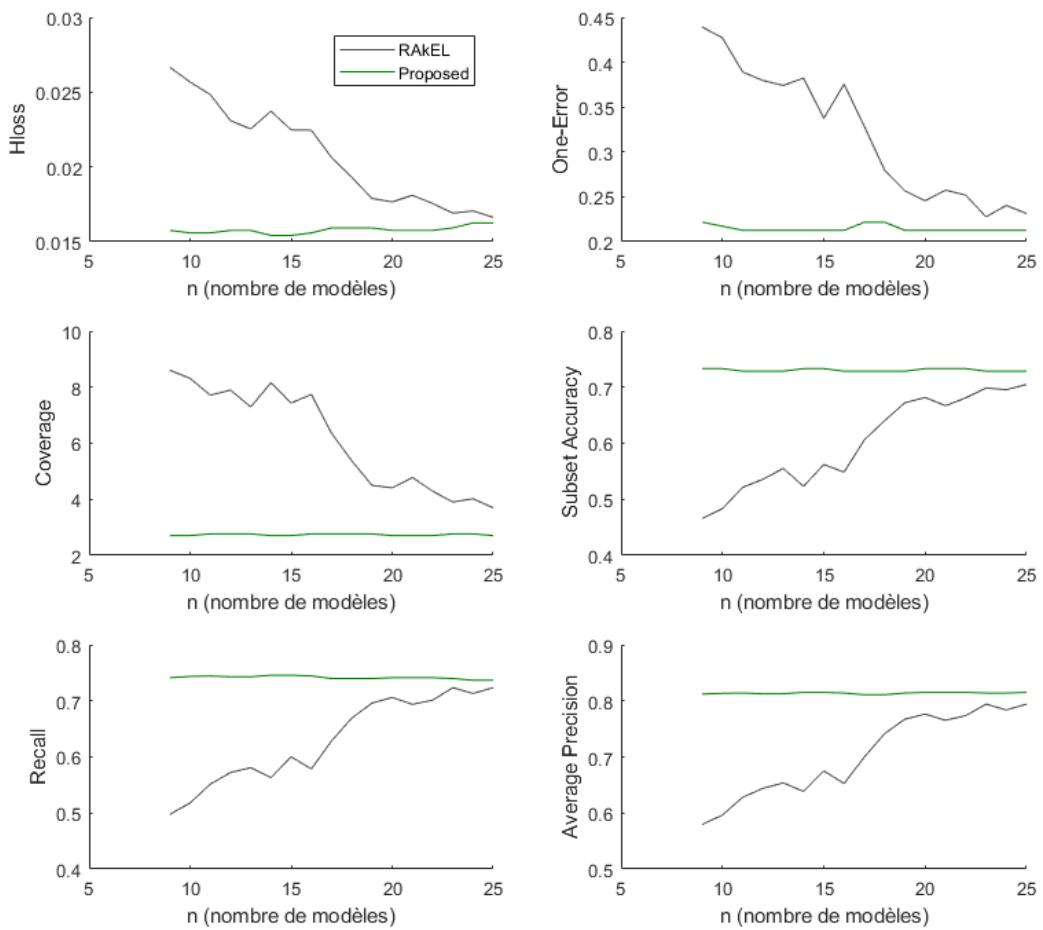
Tableau 3 : Comparaison entre RAKEL et COLEX pour la base de données Genbase k = 3 et n = 18

Métriques	RAKEL	COLEX
↓ Hloss	0.0168 ± 0.0047	0.0121 ± 0.0024
↓ One-Error	0.2670 ± 0.1325	0.1799 ± 0.0261
↓ Coverage	2.6402 ± 0.5702	2.0352 ± 0.4188
↓ Rloss	0.0000 ± 0.0000	0.0000 ± 0.0000
↑ Subset Accuracy	0.6422 ± 0.1269	0.7943 ± 0.0293
↑ Recall	0.6727 ± 0.1291	0.8035 ± 0.0275
↑ Average Precision	0.7777 ± 0.1046	0.8545 ± 0.0227

VI.3.4 IMPACT DES PARAMETRES K ET N

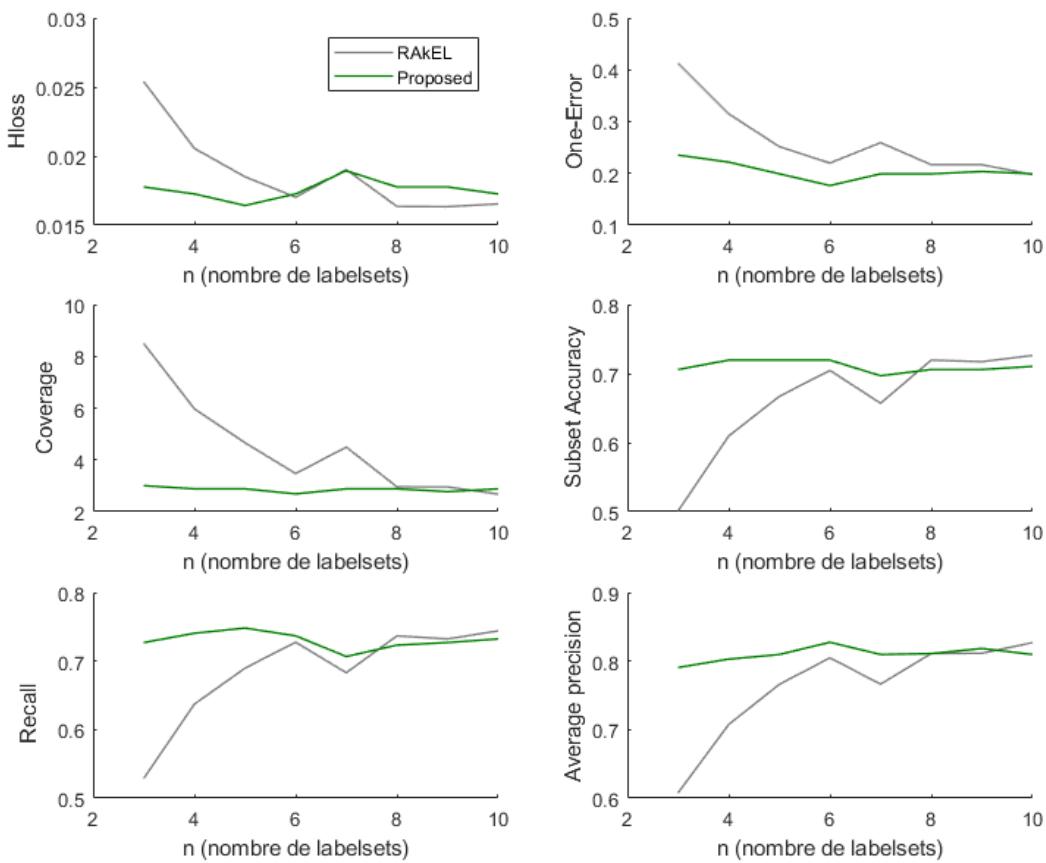
Dans la FIGURE 32, nous montrons 6 graphiques différents représentant chacun une métrique d'évaluation pour la base de données Genbase et pour $k = 3$ par rapport au nombre de labelsets n . Ces graphiques montrent une amélioration significative des performances globales de classification lors de l'utilisation de notre méthode au lieu de RAKEL. Nous remarquons sur cette figure que la méthode RAKEL donne de meilleurs performances lorsque n augmente, expliquant un besoin d'un grand nombre de labelsets pour atteindre les mêmes performances que notre méthode. Un autre fait est que notre méthode est insensible aux variations de n . Par conséquent, elle conserve les mêmes performances pour tous les nombres de labelsets, ce qui nous permet d'utiliser le nombre minimum de labelsets pour réduire le coût en temps de calcul sans sacrifier la valeur des performances. On peut également noter que même avec deux fois plus de labelsets, notre méthode donne quand même de meilleures performances que la méthode RAKEL.

Figure 32 : – Comparaison entre RAKEL et COLEX pour différentes valeurs de n avec k = 3 pour la base de données Genbase



Dans la FIGURE 32, nous faisons le même processus que précédemment mais pour $k = 10$, ce qui diminue l'effet de notre méthode puisque nous n'avons que 14 labels dans cette base de données et les labels corrélés sont plus susceptibles d'être capturés dans les mêmes labelsets sans l'intervention de la méthode proposée. Cependant, nous pouvons remarquer sur cette figure que notre méthode fonctionne toujours mieux que RAkEL jusqu'à $n = 8$ après quoi les performances des deux méthodes convergent presque au même niveau.

Figure 33 : –Comparaison entre RAkEL et COLEX pour différentes valeurs de n avec $k = 10$ pour la base de données Genbase.

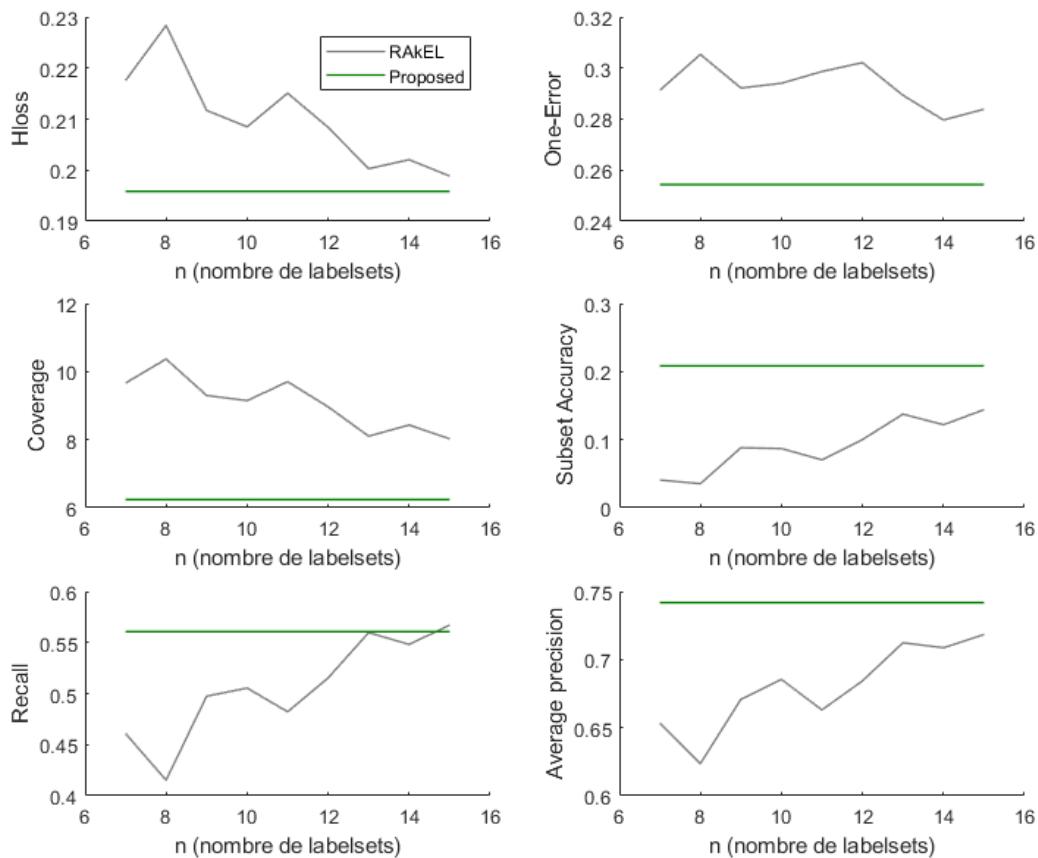


VI.3.5. DISCUSSION

Dans cette premier parti de notre chapitre, une méthode de transformation, appelée CoLEx pour "Correlation based Labelsets by Exchange strategy", et se basant sur l'optimisation du choix des k -labelsets en considérant les corrélations existantes entre les labels, a été proposée. Un indicateur de corrélation a été définie, ainsi qu'une procédure d'optimisation qui permet de choisir la meilleure combinaison des n k -labelsets. La méthode utilise la stratégie d'échange en maximisant les corrélations entre les labels d'un même k -labelsets. Cette stratégie fonctionne en échangeant des labels entre les différents k -labelsets. Ensuite, on a entraîné un classifier pour chaque k -labelsets en se servant de la méthode LP. On a utilisé tous ces classifier entraînés, en plus d'un système de vote pour donner une décision sur chacun des q labels. Des simulations sur la base de données médicale du chapitre précédent ainsi que sur des bases de données multi-label du monde réel ont montré l'efficacité de notre méthode par rapport à la méthode RAKEL et à d'autres méthodes de la littérature.

Dans cette section, la méthode CoLEx est comparée à plusieurs méthodes de classification multi-label. Le tableau qui fait la référence à toutes les méthodes utilisées pour la comparaison avec leurs abréviations correspondantes. EME est un ensemble évolutif multi-label qui utilise un algorithme génétique pour optimiser le choix des labelsets. EML est un ensemble de classifier multi-label utilisant différentes méthodes pour combiner les sorties de chaque classifier.

Figure 34 : Comparaison entre RAKEL et COLEX pour différentes valeurs de n avec k = 2 pour la base de données Yeast.



ECC génère une prédiction en sommant les sorties de différents ordres de chaînes aléatoires et effectue l'apprentissage classique des classifier en chaîne, c'est-à-dire que chaque label est entraîné seul avec les prédictions des labels précédents utilisés comme caractéristiques supplémentaires. CLR considère la classification binaire par paires où chaque paire de labels est confrontée aux autres par un classifier approprié. IBLR considère les labels des instances voisines comme des caractéristiques d'échantillons de test et réduit ainsi l'apprentissage basé sur les instances à la régression logistique. RF-PCT applique une forêt aléatoire sur les arbres de regroupement prédictifs. HOMER travaille sur une hiérarchie de classificateurs multi-label. TSA est une architecture en deux étapes utilisant deux couches d'apprentissage différentes telles que le *BinaryRelevance* et CLR. De plus, une méthode d'ensemble d'arbres multilabel pour la classification multi-label appelée ML-Forest est utilisée pour comparer les méthodes de l'état de l'art présentées dans cet article. Les comparaisons sont effectuées à l'aide de plusieurs mesures d'évaluation sur différentes bases de données. Des statistiques sur des bases de données utilisées sont données dans le TABLEAU 4

Tableau 4 : méthodes multi-label utilisées dans la partie expérimentale

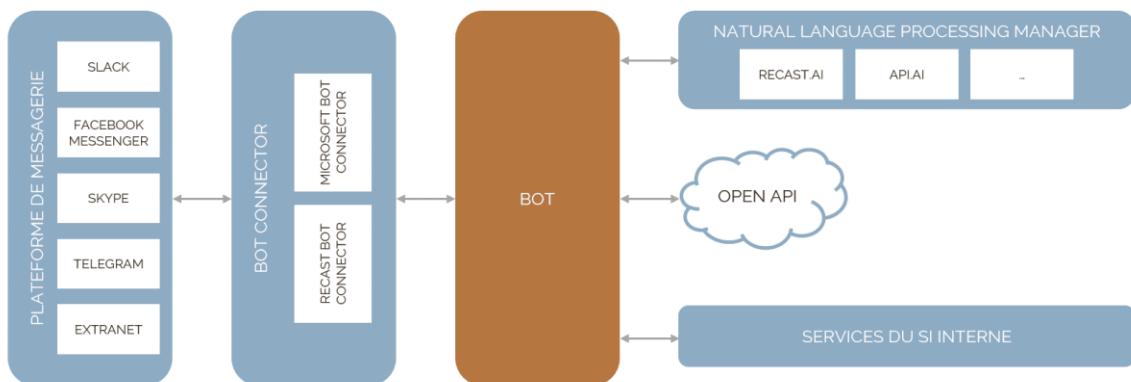
Méthode	Abréviation
Evolutionary multi-label Ensemble	EME
Ensemble of multi-label classifiers with mean rule	EML _A
Ensemble of multi-label classifiers with threshold selection	EML _T
Ensemble of multi-label classifiers with static weighting	EML _S
Ensemble of multi-label classifiers with Shepard's rule	EML _P
Ensemble of multi-label classifiers with Dudani's rule	EML _D
Ensemble of Classifier Chains	ECC
Calibrated Label Ranking	CLR
Instance Based Logistic Regression	IBLR
Random Forest of Predictive Clustering Trees	RF-PCT
Hierarchy Of multi-label classifiERs	HOMER
Two Stage Architecture	TSA
multi-label Forest	ML-F
Correlation based labelsets optimized by exchange strategy (proposed method)	CoLEx

VI.4. IMPLEMENTATION : TECHNOLOGIE, ARCHITECTURES, MODELES ET DATASET VIEW

VI.4.1. ETUDE DES TECHNOLOGIE

Dans le numérique encore plus que dans les environnements traditionnels, la confiance est impérative. C'est elle qui permet de passer outre un degré d'incertitude élevé renforcé par l'absence totale d'interaction humaine et de communication non verbale. Or la confiance est un concept polymorphe qui mobilise à la fois des dimensions cognitives et des dimensions affectives. De plus la confiance ne se décrète pas, mais se construit. Comment donc influencer favorablement la confiance placée dans un outil de e-santé sur tout si ce dernier intègres des aspects technologiques non rependu ? voilà pourquoi nous avons opéré un choix minutieux de technologies et outils a utilisé pour être fiable, viable et vraie dans notre déploiement.

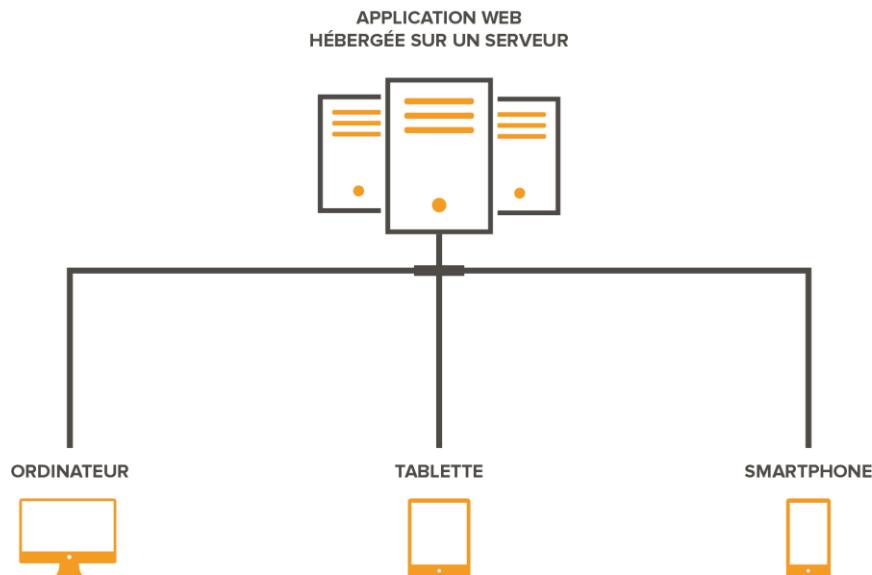
Figure 35 : Architecture du système AMEN



ASPECT WEB

Une **application web** désigne un logiciel applicatif hébergé sur un serveur et accessible via un navigateur web. Contrairement à un logiciel traditionnel, l'utilisateur d'une application web n'a pas besoin de l'installer sur son ordinateur. Il lui suffit de se connecter à l'application à l'aide de son navigateur favori.

Figure 36 : Application web présentation



OUTILS DE DEVELOPPEMENT FRONT-END :

En développement web, la notion de « front end » fait référence à l'ensemble des éléments visibles et accessibles directement sur un site, en un mot le front-end représente les interfaces utilisateurs. Nous avons utilisé comme outils certains langages descriptifs et de Framework :

❖ **Framework :**

- Tailwindcss
- AlpineJs
- ViteJs
- livewire

❖ **Language:**

- HTML5
- CSS
- JavaScript

OUTILS DE DEVELOPPEMENT BACK-END :

Cette partie côté serveur fait référence aux fonctions qui s'exécutent sur le serveur, comme l'accès aux données, la gestion des bases de données, les notifications, l'authentification, les fonctions d'application serveur et toutes les autres fonctions. On peut dire que toutes les fonctions auxquelles on accède par le côté client sont exécutées ici.

❖ **Framework :**

- FastApi

- Flask
 - Laravel
- ❖ **Language :**
- Python
-

ASPECT MOBILE

On parle d'application mobile quand on fait référence à un programme ou à un logiciel applicatif téléchargeable sur un support mobile (tablette ou smartphone). Lorsqu'elle est téléchargée, elle s'exécute grâce à un système d'exploitation : iOS pour l'environnement Apple et Android pour les supports mobiles de chez Samsung, Xiaomi ou Google.

OUTILS DE DEVELOPPEMENT FONT-END : UI/UX

UI est traduit par User Interface, soit, l'interface utilisateur. L'UI est l'élément faisant le lien entre l'humain et la machine, c'est en quelque sorte la partie visible. C'est ce qui permet à l'utilisateur de naviguer de manière fluide et intuitive sur votre application mobile. Cette étape est en quelque sorte l'organisation des éléments graphiques et textuels afin de proposer une expérience attrayante à l'utilisateur. L'UI répond à des normes techniques (couleurs, taille de texte...).

- ❖ **Framework :**
- Flutter
- ❖ **Langages :**
- Dart

OUTILS DE DEVELOPPEMENT BACK-END

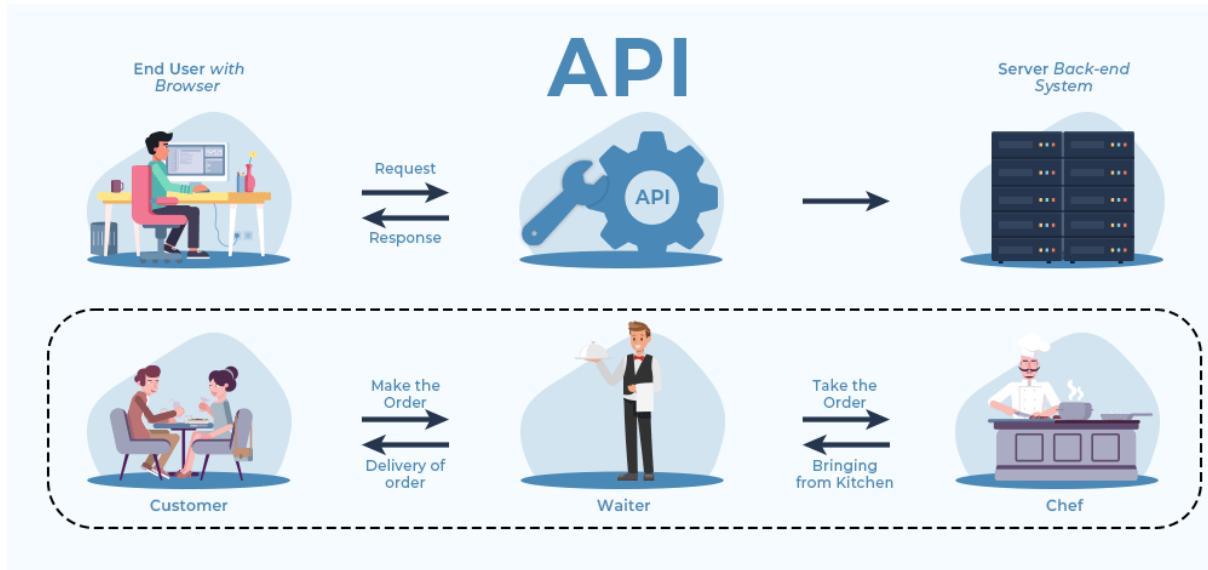
Un Back-end pour le mobile regroupe l'ensemble des supports sur lesquels une appli est présente et coordonne les actions relatives à l'administration et la possibilité de gérer cette dernière, la gestion de droit d'accès et aussi celle du stockage.

- ❖ **Framework :**
- Flask
 - Fastapi
- ❖ **Langages :**
- Python
-

ASPECT INTERCOMMUNICATION : API

Une API (*application programming interface* ou « interface de programmation d'application ») est une interface logicielle qui permet de « connecter » un logiciel ou un service à un autre logiciel ou service afin d'échanger des données et des fonctionnalités

Figure 37 : représentation d'un api



- ❖ **Stratégie des gestion d'API :**
 - Interne : micro service
 - Partenaire
- ❖ **Architecture d'API :**
 - REST
 - Orienté évènement

ASPECT TRAITEMENT ET ANALYSE DE DONNEES

Le traitement des données correspond à l'ensemble des opérations techniques qui permettent d'analyser les informations recueillies pour répondre aux objectifs de l'étude.

- ❖ **Framework et Packages :**
 - Tensorflow
 - Pytorch
 - Numpy
 - Opencv
 - Matplotlib
 - Nltk
 - Sklearn
- ❖ **Langages :**
 - Python
 - C
- ❖ **Gestion de base de données :**
 - MySQL
 - DB Browser for Sqlite

VI.4.2. GRAPHIQUE ET DATASET

Cette base de données comprend des enregistrements ECG à long terme de 334 sujets (118 hommes, âgés de 226 à 71 ans, et 4 femmes, âgées de 54 à 63 ans) atteints d'insuffisance cardiaque congestive sévère (classe NYHA 3-4). Ce groupe de sujets faisait partie d'un groupe d'étude plus large recevant un traitement médical conventionnel avant de recevoir l'agent inotrope oral, la milrinone. De plus amples détails sur le groupe d'étude plus large sont

disponibles dans la première référence citée ci-dessus. Un certain nombre d'études supplémentaires ont utilisé ces enregistrements; voir les références supplémentaires ci-dessous.

Les enregistrements individuels durent chacun environ 20 heures et contiennent deux signaux ECG échantillonnés chacun à 250 échantillons par seconde avec une résolution de 12 bits sur une plage de ± 10 millivolts. Les enregistrements analogiques originaux ont été réalisés au Beth Israel Hospital de Boston (maintenant le Beth Israel Deaconess Medical Center) à l'aide d'enregistreurs ECG ambulatoires avec une bande passante d'enregistrement typique d'environ 0,1 Hz à 40 Hz. Les fichiers d'annotation (avec le suffixe .ecg) ont été préparés à l'aide d'un détecteur automatisé et n'ont pas été corrigés manuellement.

Figure 38 :mimic-III data base

ADMISSIONS.csv 26.82 kB	CALLOUT.csv 13.82 kB	CAREGIVERS.csv 178.14 kB	CHARTEVENTS.csv 77.73 MB	CPTEVENTS.csv 149.02 kB	DATETIMEEVENTS.csv 1.78 MB	DIAGNOSES_ICD.csv 49 kB
DRGCODES.csv 23.12 kB	D_CPT.csv 12.72 kB	D_ICD_DIAGNOSES.csv 1.31 MB	D_ICD_PROCEDURES.csv 289.91 kB	D_ITEMS.csv 853.17 kB	D_LABITEMS.csv 37.76 kB	ICUSTAYS.csv 13.5 kB
INPUTEVENTS_CV.csv 4.74 MB	INPUTEVENTS_MV.csv 3.39 MB	LABEVENTS.csv 4.76 MB	LICENSE.txt 25.82 kB	MICROBIOLOGYEVENTS... 221.96 kB	NOTEVENTS.csv 95 B	OUTPUTEVENTS.csv 1.01 MB
PATIENTS.csv 8.61 kB	PRESCRIPTIONS.csv 1.67 MB	PROCEDUREEVENTS... 134.28 kB	PROCEDURES_ICD.csv 13.51 kB	SERVICES.csv 7.46 kB	SHA256SUMS.txt 2.19 kB	TRANSFERS.csv 47.2 kB

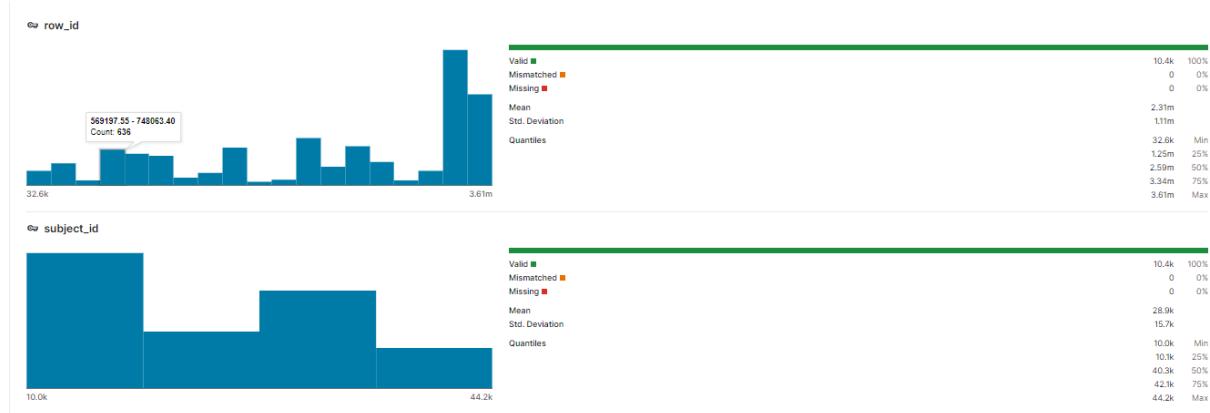
Figure 39 : data mining : jonction de la table Drug_prescription et admission

row_id	subject_id	hadm_id	icustay_id	startdate	enddate	drug_type	drug	drug_name_poe	drug_name_gene...
32.6k	3.61m	10.0k	44.2k	100k	199k	MAIN	Potassium Chloride	5%	[null]
32680	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Pneumococcal Vac Polyvalent	Pneumococcal Vac Polyvalent	Pneumococcal Vac Polyvalent
32681	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Bisacodyl	Bisacodyl	Bisacodyl
32682	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Bisacodyl	Bisacodyl (Rectal)	
32683	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Senna	Senna	Senna
32684	42458	159647		2146-07-21 00:00:00	2146-07-21 00:00:00	MAIN	Docusate Sodium (Liquid)	Docusate Sodium (Liquid)	Docusate Sodium (Liquid)
32685	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Heparin	Heparin	Heparin Sodium
32686	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Simvastatin	Simvastatin	Simvastatin
32687	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Aspirin	Aspirin	Aspirin
32688	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Donepezil	Donepezil	Donepezil
32689	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Gabapentin	Gabapentin	Gabapentin
32690	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Namenda	Namenda	*NF* Memantine
32691	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Tolterodine	Tolterodine	Tolterodine
32692	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Metoprolol Succinate XL	Metoprolol Succinate XL	Metoprolol Succinate XL
32693	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Tamsulosin	Tamsulosin	Tamsulosin
32694	42458	159647		2146-07-21 00:00:00	2146-07-22 00:00:00	MAIN	Docusate Sodium	Docusate Sodium	Docusate Sodium
59637	42275	128293	298478	2135-10-25 00:00:00	2135-10-27 00:00:00	MAIN	Sodium Chloride 0.9% Flush	Sodium Chloride 0.9% Flush	Sodium Chloride 0.9% Flush

Pour juger de l'efficacité d'un médicament par rapport aux autres le résultat de ce dernier doit être statiquement comparer aux résultats des autres sur une multitude de patients. La colonne drug, drug-type et drug-genre décrivant respectivement la base

chimique, les éléments de compositions et le classement selon le genre de prise par le patient. La première ligne est une représentation statique de l'aperçu général de données de la table.

Figure 40 : étude sur la validité de données par colonne | row_id, subject_id : authentification



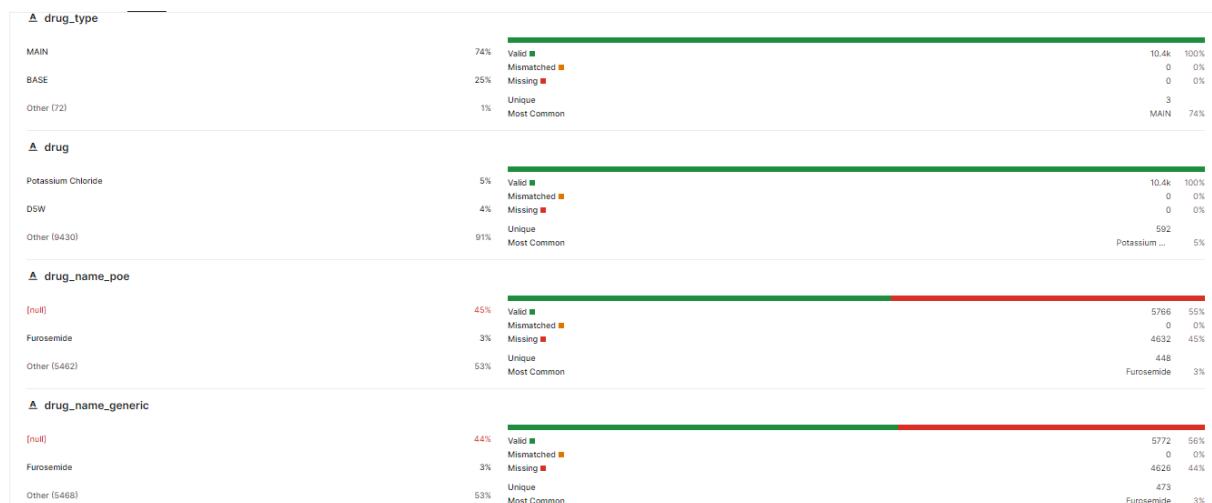
Les deux colonnes `row_id` et `subject_id` contiennent les données personnelles sur les patients, des éléments clés pour toute la suite de l'utilisation de la Platform. L'étude de la validité se fera suivant les critères suivants :

- ❖ Valid : vérifie si le nombre des lignes est strictement égales à n-1 index
- ❖ Mismatched : pour Nan
- ❖ Missing : vérification des éléments manquant

Figure 41 : étude sur la validité de données par colonne | startdate, enddate : séjour



Figure 42 : étude sur la validité de données par colonne | pour les médicaments



Pour les données pharmaceutiques, on va remarquer que la variable *valid* n'est pas au max pour drug-name-generic cela est dû que certains médicaments ne sont pas livré avec toute les informations chimiques justifier par le secret de fabrication commerciale.

Figure 43 : relation entre toutes les variables du heart attack dataset

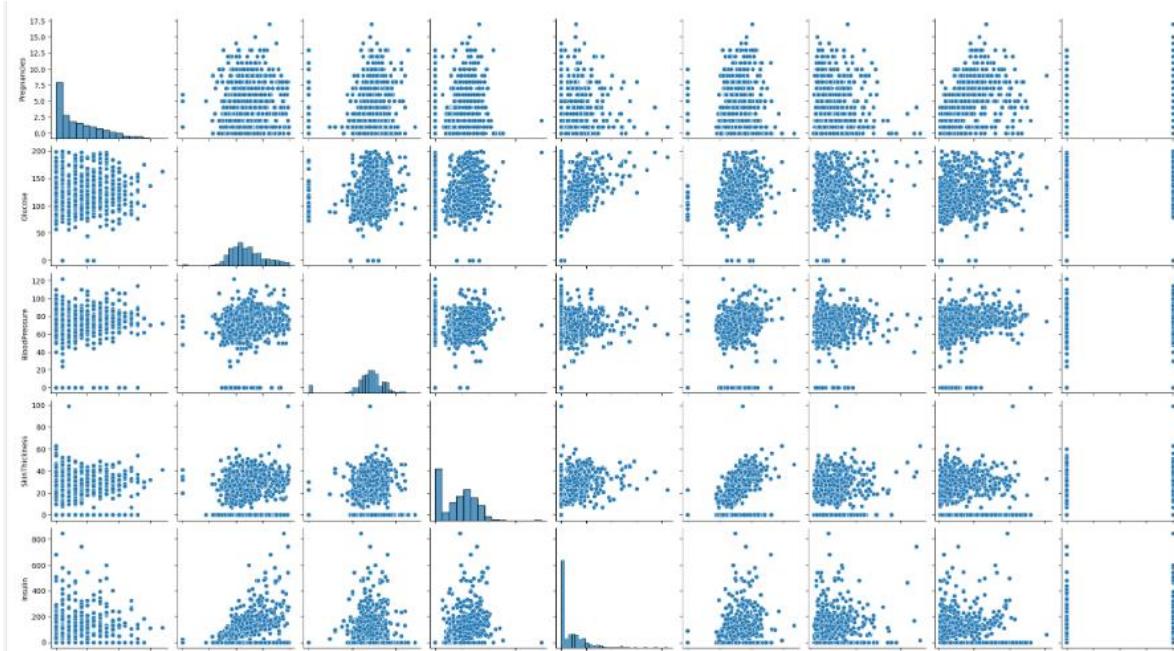
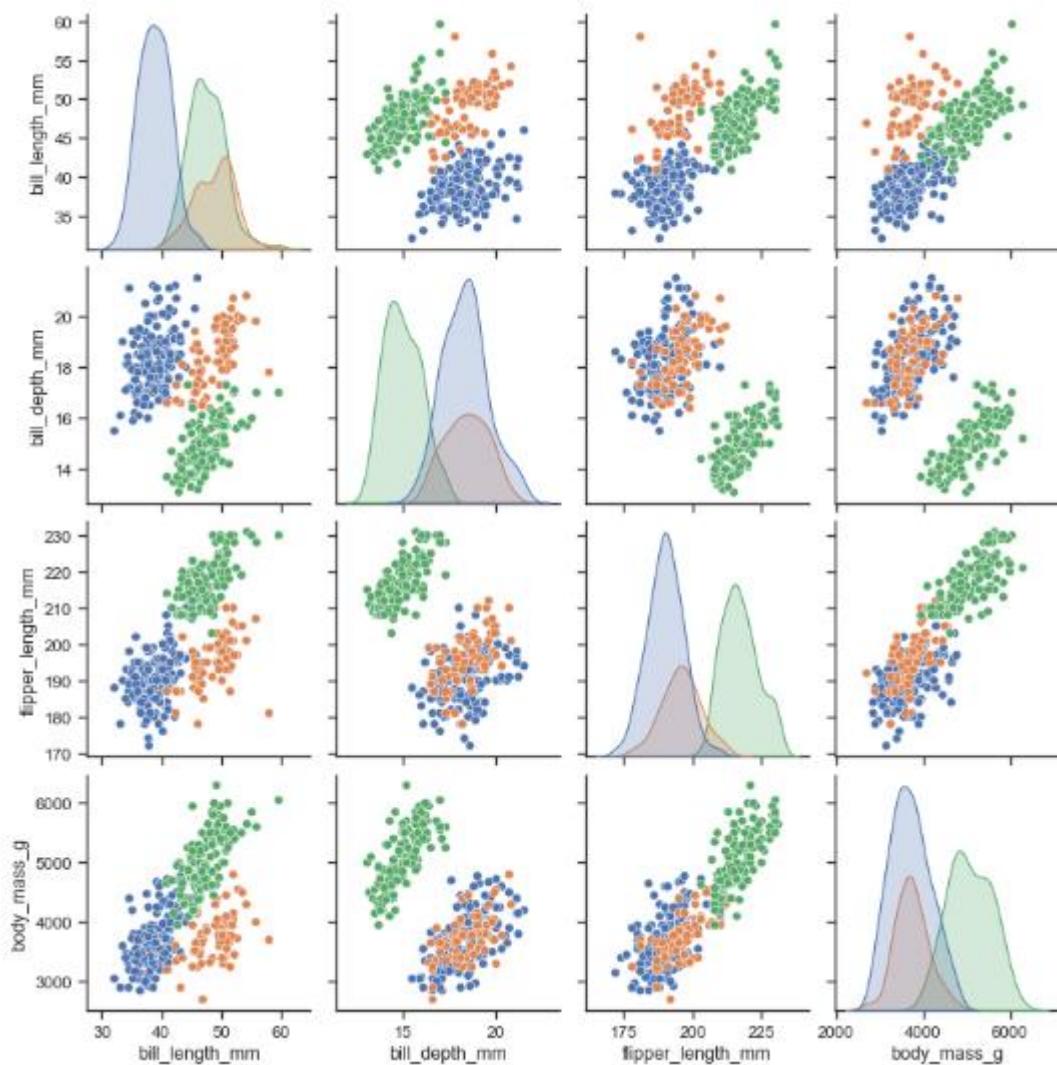


Figure 44 : représentation de notre modèle par tri des hyper paramètre par model pré entrainer



VI.4.3. PRESENTATIONS DE MODELS ET DES INTERFACES GRAPHIQUES

La première partie du programme est l'importation des bibliothèques nécessaires pour mener à bien les différentes tâches. En python après la création d'un environnement virtuel les packages doivent être installer pour la gestion de dépendances

Figure 45 : installation de packages

```

File Edit Selection View Go Run Terminal Help
File Explorer Terminal main.py speechRecognition_controller.py pneumonia_DETECT.py
EXPLORER App > models > pneumonia_DETECT.py ...
    > amen_model
    > App
        > controllers
            > __init__.py
            > speechRecognition_controller.py
        > models
            > __init__.py
            > pneumonia_DETECT.py
        > node_modules
        > src
        > static
        > templates
            > auth
            > chats
            > Desease_info
            > doctor_manager
            > home
            > pharma_manager
            > shared
                > __init__.py
                > main.py
                > package-lock.json
                > package.json
                > postcss.config.js
                > tailwind.config.js
        > Data
    > OUTLINE
    > TIMELINE
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL SQL CONSOLE
al-packages\python310\site-packages (from watson-streaming) (2.28.2)
Collecting sounddevice (from watson-streaming)
  Downloading sounddevice-0.4.6-py3-none-win_amd64.whl (199 kB)
  199.7/199.7 kB 38.5 kB/s eta 0:00:00
Collecting websocket-client (from watson-streaming)
  Downloading websocket_client-1.5.1-py3-none-any.whl (55 kB)
  55.9/55.9 kB 31.5 kB/s eta 0:00:00
Collecting fluteline (from watson-streaming)
  Downloading fluteline-0.0.8-py3-none-any.whl (5.5 kB)
Requirement already satisfied: charset-normalizer<4,>2 in c:\users\sethi\appdata\local\packages\pythonsoftwarefoundation.python.3.10_qbz5n2kfra8p\localcache\local-packages\python310\site-packages (from requests-watson-streaming) (3.1.0)
Requirement already satisfied: idna<4,>2.5 in c:\users\sethi\appdata\local\packages\pythonsoftwarefoundation.python.3.10_qbz5n2kfra8p\localcache\local-packages\python310\site-packages (from requests-watson-streaming) (3.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\sethi\appdata\local\packages\pythonsoftwarefoundation.python.3.10_qbz5n2kfra8p\localcache\local-packages\python310\site-packages (from requests-watson-streaming) (1.26.15)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\sethi\appdata\local\packages\pythonsoftwarefoundation.python.3.10_qbz5n2kfra8p\localcache\local-packages\python310\site-packages (from requests-watson-streaming) (2022.12.7)
Collecting CFFI<1.0 (from sounddevice-watson-streaming)
  Downloading cffi-1.15.1-cp310-cp310-win_amd64.whl (179 kB)
  179.1/179.1 kB 59.4 kB/s eta 0:00:00
TERMINAL pip3.10 + v ⌂ ... ^ ⌂ ...
Ln 6, Col 55 Spaces:4 UTF-8 CRLF Python 3.10.11 ('amen_model:venv') ⌂ Prettier ⌂

```

Figure 46 : entrainement pour le chatbot

```

File Edit Selection View Go Run Terminal Help
File Explorer Terminal train.py - AmenHealthCareSystème - Visual Studio Code
EXPLORER AMENHEALTHCARESYSTEME > chatbot-deployment > train.py ...
    > __pycache__
    > .venv
    > standalone-frontend
    > static
    > templates
        > app.py
        > chat.py
        > data.pth
        > intents.json
        > model.py
        > nltk_utils.py
    > train.py
    > langs
    > System_predict
        > app
        > data
        > models
        > venv
            > multiple_disease_pred.py
            > streamlit run command.txt
    > OUTLINE
    > TIMELINE
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL GITLENS SQL CONSOLE
chatbot-deployment > train.py ...
Patrick Loeb (Python Engineer), 21 months ago | 1 author (Patrick Loeb (Python Engineer))
1 import numpy as np
2 import random
3 import json
4
5 import torch
6 import torch.nn as nn
7 from torch.utils.data import Dataset, DataLoader
8
9 from nltk_utils import bag_of_words, tokenize, stem
10 from model import NeuralNet
11
12 with open('intents.json', 'r') as f:
52 unique stemmed words: [...', 'a', 'allez-v', 'amen', 'artâ'r', 'au', 'auteurn', 'avc', 'beaucoupâ', 'bien', 'blag', 'bonjour', 'bonn', 'bye', 'c'est', 'cardiaqu', 'cardiovasculair', 'chou', 'coeur', 'comment', 'crâateur', 'câdeâbro-vasculair', 'detent', 'dire', 'es', 'est', 'est-c', 'hâo', 'je', 'journâe', 'lâ', 'merci', 'nerf', 'neurologiqu', 'nomm', 'plu', 'poun', 'problâ'm', 'que', 'quelqu'un', 'qui', 'voi', 'revoir', 'rire', 'salut', 't'appel', 'tarâ', 'ton', 'tu', 'util', 'vou']
52 11
Epoch [100/1000], Loss: 1.0988
Epoch [200/1000], Loss: 0.4188
Epoch [300/1000], Loss: 0.2351
Epoch [400/1000], Loss: 0.0141
Epoch [500/1000], Loss: 0.1483
Epoch [600/1000], Loss: 0.0078
Epoch [700/1000], Loss: 0.0026
Epoch [800/1000], Loss: 0.1161
Epoch [900/1000], Loss: 0.0013
Epoch [1000/1000], Loss: 0.0021
final loss: 0.0021
training complete. file saved to data.pth
(.venv) C:\Users\sethi\Desktop\AmenHealthCareSystème\chatbot-deployment\ []
TERMINAL cmd + v ⌂ ... ^ ⌂ ...
Patrick Loeb (Python Engineer), 21 months ago Ln 6, Col 22 Spaces:4 UTF-8 CRLF Python 3.10.11 64-bit (microsoft store) ⌂ Prettier ⌂

```

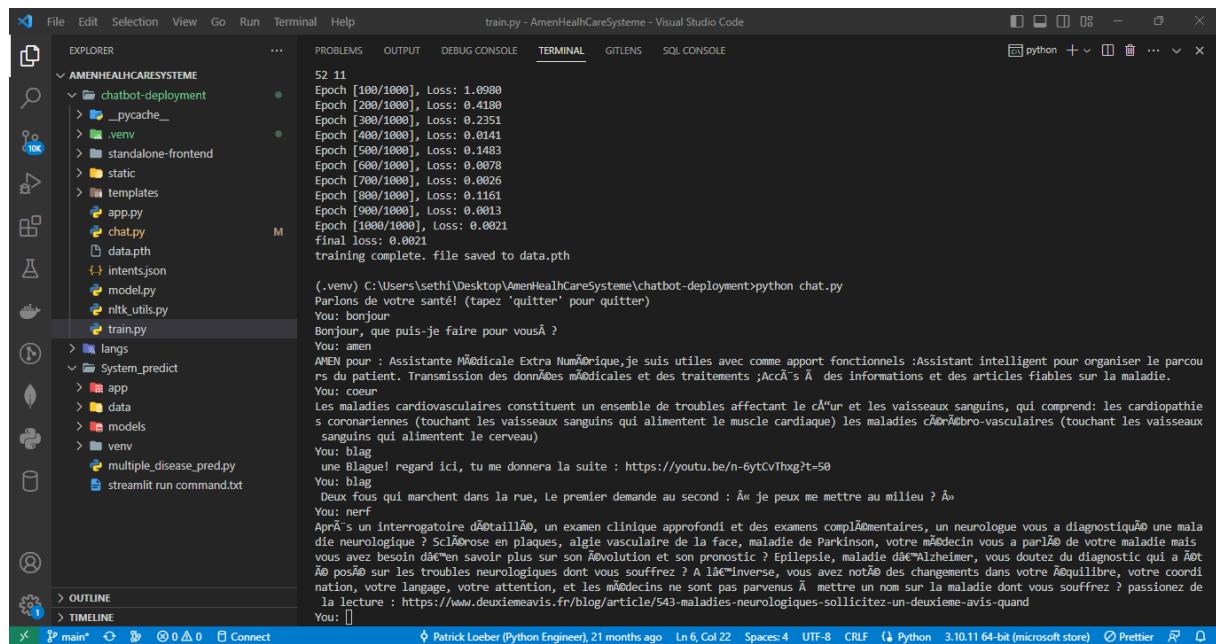
Une fois le modèle construit et sauvegardé nous passons à l'étape suivante pour ce faire nous allons additionner une partie pour l'apprentissage et une partie pour le test que nous avons préparé et ensuite nous faisons l'apprentissage avec le total de ces données, X_train et X_test les données d'apprentissage et de test, Y_train et Y_test les données d'apprentissage et de test des classes.

Nous avons défini notre modèle et l'avons compilé pour qu'il soit prêt pour un calcul efficace, il est maintenant temps d'exécuter le modèle sur quelques données nous pouvons entraîner ou ajuster notre modèle sur nos données chargées en appelant la fonction fit () sur le modèle, l'entraînement se fait sur :

Epoch : on passe par toutes les lignes du jeu de données d'entraînement (les itérations) et chaque Epoch est composée d'un ou plusieurs Batch, en fonction de la taille du Batch choisi et le modèle est adapté à de nombreuses époques.

Batch : un ou plusieurs échantillons considérés par le modèle au cours d'une Epoch avant que les poids ne soient mis à jour. Le processus de formation se déroulera pendant un nombre fixe d'itérations à travers Epoch, nous devons également définir le nombre de lignes de l'ensemble de données qui sont prises en compte avant la mise à jour des poids du modèle dans chaque Epoch appelée taille du Batch et définie à l'aide de l'argument batch_size, pour ce problème nous allons fonctionner pour un petit nombre d'Epoch (150) et utiliser une taille de Batch (128)

Figure 47 : test en console de notre chatbot



```

File Edit Selection View Go Run Terminal Help train.py - AmenHealthCareSystème - Visual Studio Code
EXPLORER PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL GITLENS SQL CONSOLE
AMENHEALTHCARESYSTEME
chatbot-deployment
__pycache__
.venv
standalone-frontend
static
templates
app.py
chat.py
data.pth
intents.json
model.py
nltk_utils.py
train.py
langs
System_predict
app
data
models
venv
multiple_disease_pred.py
streamlit run command.txt

52.11
Epoch [100/1000], Loss: 1.0998
Epoch [200/1000], Loss: 0.4198
Epoch [300/1000], Loss: 0.2351
Epoch [400/1000], Loss: 0.0141
Epoch [500/1000], Loss: 0.1483
Epoch [600/1000], Loss: 0.0078
Epoch [700/1000], Loss: 0.0026
Epoch [800/1000], Loss: 0.1161
Epoch [900/1000], Loss: 0.0013
Epoch [1000/1000], Loss: 0.0021
final loss: 0.0021
training complete. file saved to data.pth

(.venv) C:\Users\sethi\Desktop\AmenHealthCareSystème\chatbot-deployment>python chat.py
Parlons de votre santé! (tapez 'quitter' pour quitter)
You: bonjour
Bonjour, que puis-je faire pour vous?
You: amen
AMEN pour : Assistante Médicale Extra Numérique, je suis utiles avec comme apport fonctionnels :Assistant intelligent pour organiser le parcours du patient. Transmission des données médicales et des traitements ;Accès à des informations et des articles fiables sur la maladie.
You: cœur
Les maladies cardiovasculaires constituent un ensemble de troubles affectant le cœur et les vaisseaux sanguins, qui comprend: les cardiopathies coronariennes (touchant les vaisseaux sanguins qui alimentent le muscle cardiaque) les maladies cardio-vasculaires (touchant les vaisseaux sanguins qui alimentent le cerveau)
You: blag
une Blague! regard ici, tu me donnera la suite : https://youtu.be/n-6ytCvThxg?t=50
You: blag
Deux fous qui marchent dans la rue, le premier demande au second : « je peux me mettre au milieu ? »
You: nerf
Après un interrogatoire détaillé, un examen clinique approfondi et des examens complémentaires, un neurologue vous a diagnostiqué une maladie neurologique ? Si l'érosion en plaques, algie vasculaire de la face, maladie de Parkinson, votre médecin vous a parlé de votre maladie mais vous avez besoin d'en savoir plus sur son évolution et son pronostic ? Epilepsie, maladie d'Alzheimer, vous doutez du diagnostic qui a été posé sur les troubles neurologiques dont vous souffrez ? À l'inverse, vous avez noté des changements dans votre équilibre, votre coordination, votre langage, votre attention, et les médecins ne sont pas parvenus à mettre un nom sur la maladie dont vous souffrez ? passionnez de la lecture : https://www.deuxiemeavis.fr/blog/article/543-maladies-neurologiques-solicitez-un-deuxieme-avis-quand
You: []

Patrick Loebel (Python Engineer), 21 months ago Ln 6, Col 22 Spaces: 4 UTF-8 Python 3.10.11 64-bit (microsoft store) Prettier R Q

```

Afin de montrer les résultats obtenus pour notre modèle, Nous illustrons dans ce qui suit les résultats en termes de précision nous venons de démarrer le programme en console.

Figure 48 : Détection des maladie basée sur ML diabets

Diabetes Prediction using ML

Number of Pregnancies	Glucose Level	Blood Pressure value
Skin Thickness value	Insulin Level	BMI value
Diabetes Pedigree Function value	Age of the Person	
Diabetes Test Result		

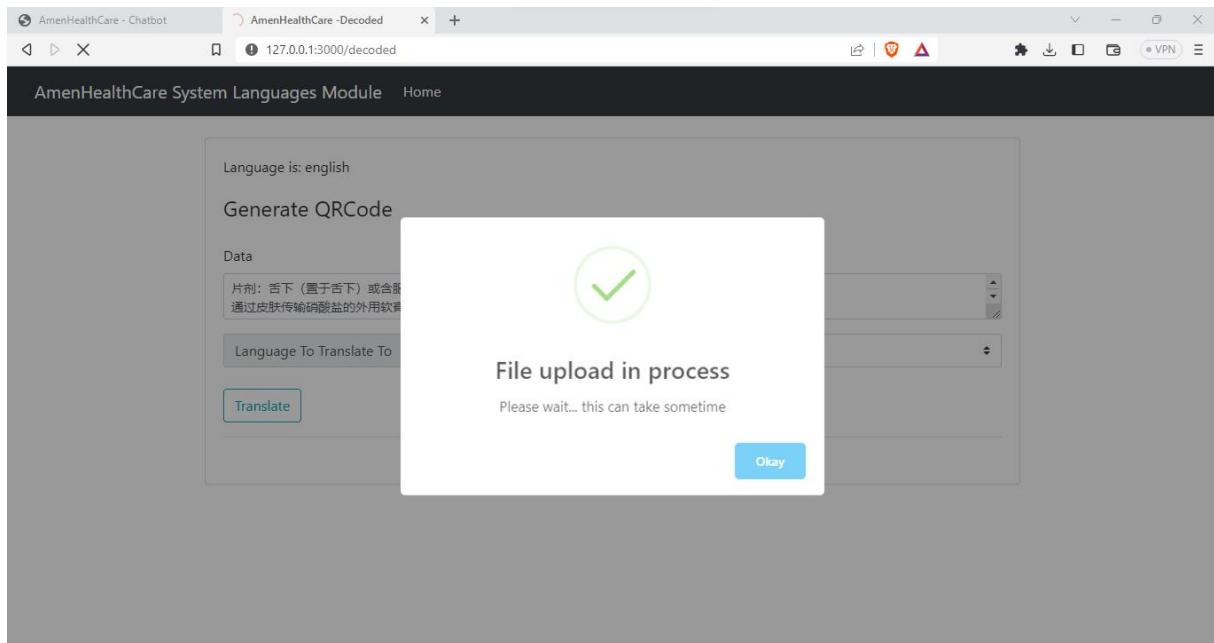
Made with Streamlit

Figure 49 : Détection des maladie basée sur ML Heart Disease Prediction

Heart Disease Prediction using ML

Age	Sex	Chest Pain types
Resting Blood Pressure	Serum Cholestral in mg/dl	Fasting Blood Sugar > 120 mg/dl
Resting ECG results	Maximum Heart Rate achieved	Exercise Induced Angina
ST depression induced by exercise	Slope of the peak exercise ST segment	Major vessels colored by flourosopy
thal: 0 = normal; 1 = fixed defect; 2 = reversible defect		
Heart Disease Test Result		

Figure 50 : optical character image to Text to audio



CONCLUSION ET PERSPECTIVE

Ce mémoire s'est intéressée au problème de classification mono- et multi-label dans un contexte médical lié au suivi des maladies cardio-vasculaires (MCV). L'objectif de ce mémoire était de proposer des approches intelligentes qui anticipent l'évolution des MCV chez des patients dans leur lieu de vie quotidien sans imposer des contraintes gênantes. Ces approches se basent sur des mesures de signaux physiologiques en temps réel pour les analyser et en sortir une décision. Dans ce mémoire, nous avons opté pour l'utilisation de deux formalismes différents pour répondre à ce problème. Le premier adopte la notion de classification mono-label qui vise la prédiction d'un épisode de fibrillation atriale (FA) basée sur la théorie des fonctions de croyance, et le second adopte la notion multi-label pour pouvoir effectuer un diagnostic multi-pathologique. Le défi scientifique à relever dans ce mémoire était la proposition d'approches de classification permettant de répondre aux contraintes de données manquantes par moment, hétérogènes, imprécises et indirectement liées au phénomène observé.

Le manuscrit commence par une introduction des approches technologiques dont l'intelligence artificielle, le système de santé, le Bigdata ainsi que le chatbot et leurs capital intérêt appliqués aux MCV démontrant l'ampleur de leur impact sur la santé et mettant en évidence les enjeux des travaux de recherche tout en introduisant le matériel utilisé dans ces travaux qui demeure principalement la base de données médicales MIMIC III. Ensuite, le manuscrit introduit la théorie des fonctions de croyance, étant l'outil technique de base de la majorité des travaux développés. Quatre contributions sont présentées par la suite. La première porte sur une approche de classification visant à prédire un épisode de la FA. Cette méthode est basée sur la théorie des fonctions de croyance, qui est une théorie permettant de modéliser l'information d'une manière différente de la théorie de probabilités en tenant en compte des spécificités de l'application relatives au manque d'information.

La mise en œuvre de la méthode développée a donné, après optimisation des paramètres et sur un ensemble de patients test et un groupe de contrôle, une précision de 79.3%, une sensibilité de 74.2%, une spécificité de 83.7% ainsi qu'une portée de prédiction d'environ 2 heures et 10 minutes avant l'occurrence de l'épisode. On a apporté par la suite des développements améliorant les performances de cette méthode en proposant une nouvelle modélisation de l'information et une option de rejet supplémentaire. Le manuscrit introduit aussi une approche originale pour la fusion des données se traduisant par une méthode de combinaison plus flexible que la règle de YAGER. On a montré qu'on peut atteindre avec ces améliorations des précisions allant jusqu'à 95.8% pour un taux de rejet d'environ 38.21%. La méthode est évaluée sur la base de données MIMIC III.

Par ailleurs, on a introduit la notion de classification multi-label avant de l'appliquer au diagnostic multi-pathologique. Les signaux physiologiques de la base de données médicale MIMIC III ont été utilisés pour faire le diagnostic des éventuelles complications cardiaques relatives à 55 maladies cardio-vasculaires. On a aussi introduit une approche multi-label basée sur la théorie des fonctions de croyance qu'on a validée grâce aux données extraites de la base médicale MIMIC III.

La méthode a été comparée à des méthodes multi-label classiques basées sur une transformation de plusieurs méthodes de classification mono-label. Il s'est avéré que la méthode proposée est meilleure que toutes les autres méthodes en utilisant plusieurs critères d'évaluation. Ceci montre la capacité de la méthode proposée à traiter les données

manquantes. Enfin, on a proposé une méthode originale appelée COLEX pour la classification multi-label puis on l'a appliquée à notre problème de diagnostic. Cette méthode s'inscrit dans le cadre des méthodes de transformation, comme elle transforme le problème de classification multi-label en plusieurs sous-problèmes mono-label. Elle est une généralisation de la méthode RAkEL qui se base sur l'optimisation du choix des k-labelsets en considérant les corrélations existantes entre les labels. En effet, la méthode permet un réglage automatique de la répartition des k-labelsets sur les n modèles de la méthode RAkEL. Les résultats obtenus sur des données du monde réel ont montré l'efficacité de notre méthode par rapport à la méthode RAkEL classique et à d'autres méthodes de la littérature.

Des perspectives sont envisageables pour la suite des travaux de recherche sur ce sujet :

- ❖ D'abord, les approches proposées extraient des caractéristiques statistiques simples des signaux physiologiques. Dans des travaux futurs, il est envisageable d'extraire des caractéristiques en utilisant d'autres approches linéaires ou non linéaires, telles que celles basées sur l'analyse par composantes principales.
- ❖ Les caractéristiques extraites de la base d'apprentissage sont modélisées individuellement par des distributions de probabilité normales. Il serait intéressant de proposer des distributions de probabilité plus adéquates, adaptées à chaque caractéristique. On peut par exemple se servir de l'estimation à noyau des fonctions de densité ou de concevoir une sélection automatique parmi plusieurs distributions de probabilité standards bien connues pour chaque caractéristique indépendamment. Il est intéressant aussi d'envisager une modélisation multidimensionnelle par signal, en considérant toutes les caractéristiques d'un même signal comme un vecteur. Cela permettra de profiter de la dépendance entre les caractéristiques et préservera la robustesse des méthodes aux signaux manquants.
- ❖ Une autre perspective consiste à adapter l'approche de la théorie des fonctions de croyance à la classification multi-classe permettant ainsi de résoudre des problèmes de classification multi-label en appliquant une méthode de transformation en plusieurs problèmes de classification multi-classe. Cette extension doit prendre en considération la conception d'une modélisation adéquate pour chaque sous-ensemble de classes.
- ❖ D'autre part, la théorie des fonctions de croyance permet d'attribuer des poids à des sous-ensembles de l'ensemble des labels possibles. Cela laisse présager que cette théorie pourrait être convenable à la conception d'une méthode de classification multi-label par adaptation, agissant directement sur les données multi-label.
- ❖ Il est aussi possible d'intégrer dans le modèle de nouvelles sources d'information qui s'appuient sur des données statistiques de la maladie (prévalence, incidence, symptômes connus, etc.) ou catégoriques propres au patient (séjours précédents, fiche de soins, antécédents médicaux, etc.).
- ❖ De nos jours, grâce à la domotique et l'affluence des objets connectés dans la vie quotidienne, il est possible de collecter de grands volumes de données liées d'une

manière directe ou indirecte à l'état de santé d'un sujet. En présence de ces données, il est intéressant d'utiliser l'apprentissage profond (deep Learning) à plusieurs niveaux comme au niveau de l'extraction de caractéristiques, le remplissage des données manquantes ou la classification elle-même. En fonction du volume de données, il est aussi possible que plusieurs de ces tâches soient accomplies simultanément par un réseau de neurones utilisant l'apprentissage profond. Il serait alors intéressant d'étudier la transformation des signaux en images cohérentes pour tirer profit de ces réseaux.

- ❖ Les approches proposées permettent le suivi d'un patient quelconque grâce à un classifier pré-entraîné. Il est judicieux de calibrer le système aux informations du patient, en exploitant le transfert d'apprentissage.
- ❖ L'apprentissage proposé utilise la base de données MIMIC III, formée de signaux collectés en unités de soins intensifs. Cela nous a permis de valider nos approches grâce à des signaux réels. Toutefois, ces signaux ne sont pas représentatifs du cadre applicatif auquel est dédié le système, à savoir la surveillance à domicile. Une collaboration avec une équipe de cardiologues est envisagée pour l'accès à des données plus réalistes.
- ❖ Enfin, les développements menés à ce jour se situent au niveau méthodologique. D'un point de vue technologique, d'autres perspectives visent à développer un système de surveillance à domicile respectant les contraintes pratiques d'acceptabilité, éthiques et économiques. Il est nécessaire que le produit soit accepté et validé par la communauté visée.

BIBLIOGRAPHIE

- A. Freeman, S. T.-M.-A.-K. (2016). The role of socioeconomic status in depression : results from the courage (aging survey in europe). *BMC public health*, 1098.
- actualiteinformatique. (2023, Mars 24). *actualiteinformatique*. Récupéré sur actualiteinformatique: <https://actualiteinformatique.fr/tag/machine-learning>
- Alexandre. (2017). *La guerre des intelligences*. paris.
- Artefacto. (2023, Mars 22). Récupéré sur artefacto-ar.com: <https://www.artefacto-ar.com/realite-virtuelle/>
- Bali, S. (2018). Barriers to development of telemedicine in developing countries. In *Telehealth*. IntechOpen.
- Baumberg, P. A. (2006). alcohol in europe. ondon : *Institute of alcohol studies*, 73-75.
- Blonde, L. K. (2018). Interpretation and impact of real-world clinical data for the practicing clinician. *Advances in therapy* 35, 1763-1774.
- C. J. Murray, T. V. (2012). "Disability-adjusted life years (dalys) for 291 diseases and injuries in 21 regions, 1990–2010 : a systematic analysis for the global burden of disease study 2010.". *The lancet*, 2197–2223.
- C. Shi, X. K. (2011). Multi-label ensemble learning. in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 223-239.
- Carré, G. C. (2008). « Technologies pour le soin, l'autonomie et le lien social des personnes âgées : quoi de neuf ? ». *Gérontologie Société*.
- clarte. (2023, Mars 22). Récupéré sur clarte-lab: <https://www.clarte-lab.fr/domaines-de-competences/realite-virtuelle>
- CLAUDE, M. (2014). *Le parcours de soins du patient : visible et partagé*. . Paris: Mémoire de l'école des hautes études.
- Cogito. (2019). *l'intelligence humaine associée à l'informatique cognitive*.
- D. Coggon, D. B. (2009). *Epidemiology for the Uninitiated*. . ohn Wiley & Sons.
- D. Sullivan. (2017). *What Is Arrhythmia ?* Récupéré sur What Is Arrhythmia ?: <https://www.healthline.com/health/arrhythmia>
- dataanalytics. (2023, Mars 25). *dataanalyticspost*. Récupéré sur dataanalyticspost: <https://dataanalyticspost.com/category/innovation-en-action/>
- datascientest. (2023, Mars 22). *datascientest*. Récupéré sur <https://datascientest.com/big-data-tout-savoir>
- David, K. (2021). *RÉSEAUX DE NEURONES (Techniques et concepts associés)*Formation sur les outils mathématiques de l'IA.

- Dendane, M. A. (2017). *Conception et réalisation d'un capteurs passif biométrique implantable et d'un lecteur assurant l'activation et la communication sans fil.* université du Québec.
- developpez.com. (2023, Avril 05). Récupéré sur developpez.com:
<https://sabrique.developpez.com/uml/tutoriel/unifiedProcess/>
- Eroukhmanoff J, B. (2018). *La médecine au défis de l'intelligence artificielle. médecine et politique . publoc.*
- e-santé, I. (2019). Livre blanc: chatbot du lab e-santé. *le futur de la santé sera-t-il conversationnel?*, 1-44.
- esechos.fr. (2023, Mars 27). esechos.fr. Récupéré sur esechos.fr:
<https://www.esechos.fr/2018/05/un-chatbot-pour-reperer-les-troubles-de-lautisme-991463>
- Fiorini, M. (2018). *Médecins et patients dans le monde des data, des algorithmes, et de l'intelligence artificielle.* Ariis (Alliance pour la recherche et l'innovation des industries de santé).
- G. Tsoumakas, I. K. (2010). Random k-labelsets for multilabel classification. *IEEE Transactions on Knowledge and Data Engineering*, 1079–1089.
- GUILIANO. (2009). *10 Règles pour tracer le parcours du patient* (éd. 1ere).. Paris: France santé.
- I. Guyon, S. G. (2008). *Feature extraction : foundations and applications.* Springer.
- I.-M. Lee, E. J. (2012). Effect of physical inactivity on major non-communicable diseases worldwide : an analysis of burden of disease and life expectancy. *The Lancet*, 2019-2029.
- IBM. (2023, Avril 05). Récupéré sur www.ibm.com:
<https://www.ibm.com/docs/fr/icos/12.9.0?topic=overview-modeling-languages>
- Injibar, M. M. (2021). *Classification évidentielle mono- et multi-label :application à la détection de maladies cardio-vasculaires.* Beyouth: Hal Open science.
- ionos. (2023, Avril 05). Récupéré sur www.ionos.f: [https://www.ionos.fr/digitalguide/sites-internet/developpement-web/uml-un-langage-de-modélisation-pour-la-programmation-orientée-objet/#:~:text=Le%20langage%20UML%20\(Unified%20Modeling,un%20logiciel%20ou%20un%20syst%C3%A8me.](https://www.ionos.fr/digitalguide/sites-internet/developpement-web/uml-un-langage-de-modélisation-pour-la-programmation-orientée-objet/#:~:text=Le%20langage%20UML%20(Unified%20Modeling,un%20logiciel%20ou%20un%20syst%C3%A8me.)
- J. Alcalá-Fdez, A. F. (2011). Keel data-mining software tool : data set repository, integration of algorithms and experimental data-mining software tool : data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic & Soft Computing*.
- Jamal, K. J. (2015). *the setting in place of patient control and monitoring system in the room of a hospital.* Butarhe: Catholic University of Rwanda.

- Karboub, K. (2022). *Contribution à l'amélioration des performances des services médicaux urgents appliquant l'IoT et l'intelligence artificielle*. Rabbat: université de Lorraine.
- Klinger, E. (2006). *Apports de la réalité virtuelle à la prise en charge de troubles cognitifs et comportementaux*. paris: Télécom ParisTech.
- Kunihiko, F. (1980). A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Neocognitron*, 193-202.
- lab-e-santé. (2020). *le futur de la santé sera-t-il conversionnel?* lab e-santé.
- Laura DI ROLLO, M. G. (2019). *Débat : L'intelligence artificielle en santé*. Lyon: Lyon Catholic University.
- M. Durand, A. S. (2020). *Applications médicales de l'intelligence artificielle : opportunités & challenges*. paris: Progrès en Urologie – FMC.
- M. H. Forouzanfar, P. L. (2017). Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm hg. *Jama*, 165–182.
- M. Hamer, G. O. (2017). Physical inactivity and the economic and health burdens due to cardiovascular disease : exercise as medicine," in *Exercise for Cardiovascular Disease Prevention and Treatment*. Springer, 3-18.
- M. Stahre, J. R. (2014). Peer reviewed : Contribution of excessive alcohol consumption to deaths and years of potential life lost in the united states. *Preventing chronic disease*.
- M.-L. Zhang, Y.-K. L.-Y. (2018). Binary relevance for multi-label learning : an overview. *Frontiers of Computer Science*, 191-202.
- Maes. (2018). *Nouvelle Organisations et Architectures Hospitalières*. paris: Ministère de la santé.
- Marvin Minsky, S. P. (1969). *Perceptrons : An Introduction to Computational Geometry*. MIT Press.
- McCarthy, J. (1995). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (Vol. IV). *AI Magazin*.
- Metrics, I. (2020). *Evaluation : the global burden of disease : generating evidence, guiding policy*. Seattle: WA : IHME Seattle.
- MIT-LCP. (2023, Avril 05). *mimic.mit.edu*. Récupéré sur *mimic.mit.edu*: <https://mimic.mit.edu/docs/gettingstarted/>
- Motivation, E. (Réalisateur). (2020). *la prise de décision* [Film].
- NACUA, R. (2009). *Conception et développement d'un système ambulatoire pour la mesure de l'activité du système nerveux Autonomes pour la surveillance de personnes âgées*. Université Joseph-Fourier - Grenoble I.
- netapp. (2023, Mars 22). *netapp-ia*. Récupéré sur *netapp:com*: <https://www.netapp.com/fr/artificial-intelligence/what-is-artificial-intelligence/>

- OMS. (2023). Des systèmes de santé renforcés sauvent plus de vies. OMS.
- Oracle. (2023, Mars 22). *chatbot-oracle*. Récupéré sur oracle:
<https://www.oracle.com/fr/chatbots/what-is-a-chatbot/#:~:text=Au%20niveau%20le%20plus%20fondamental,communiquaient%20avec%20une%20personne%20r%C3%A9elle.>
- oracle. (2023, Mars 22). *oracle*. Récupéré sur Oracle inc: <https://www.oracle.com/fr/big-data/what-is-big-data/>
- OURAMDANE, N. (2008). Vers un syst`eme d'assistance `a l'interaction 3D pour le travail et le t`el`etravail collaboratif dans les environnements de r'ealit`e virtuelle et augment'ee. UNIVERSITE D'EVRY-VAL D'ESSONNE.
- physionet. (2023, Avril 05). Récupéré sur physionet: <https://physionet.org/>
- Pierre-louis, B. (2009). *Traité d'économie et de gestion de la santé*. . France: Editions de santé.
- Rjiba, S. (2022). *Assistance Virtuelle pour la prevention des risque cardio vasculaire* . cannon .
- Roques, P. (2017). *UML3 par la pratique*. Eyrolles.
- Rosenblatt, F. (1958). A Probabilistic Model for Information Storage and Organization in The Brain. *Psychological Review*, 65-386.
- S. J. Hoffman, C. T. (2015). Overview of systematic reviews on the health-related effects of government tobacco control policies. *BMC public health*, 744.
- S. Lewington, R. C. (2003). Age-specific relevance of usual blood pressure to vascular mortality : a meta-analysis of individual data for one million adults in 61 prospective studies. . *Hellenic Journal of Cardiology*, 1903–1913.
- S. Yusuf, S. R. (2014). cardiovascular risk and events in 17 low-, middle-, and high-income countries. *New England Journal of Medicine*, 818–827.
- Sage advices. (2023, Mars 22). Récupéré sur sage.com: <https://www.sage.com/fr-fr/blog/glossaire/intelligence-artificielle-ia-definition/>
- science, F. (2023, Mars 22). *futura-science-medecine-e-santé-def*. Récupéré sur futura-science: <https://www.futura-sciences.com/sante/definitions/medecine-e-sante-15728/>
- Shilo, S. H. (2020). Axes of a revolution: challenges and promises of big data in healthcare. *Nature medicine* 26, 29-38.
- Statista. (2023, Mars 26). *Healthcare data volume globally 2020 forecast* .
- Stephan, H. (2018). *L'intelligence artificielle, "la pire ou meilleure chose arrivée à l'humanité"*.
- Subramanian, I. S. (2020). Multi-omics data integration, interpretation, and its application. *Bioinformatics and biology insights* 14.
- Sullivan, D. (2018). , *Acute Myocardial Infarction*. Récupéré sur healthline.com: <https://www.healthline.com/health/acute-myocardial-infarction>.

- Sullivan, D. (2018). *Stable Angina*. Récupéré sur www.healthline.com:
<https://www.healthline.com/health/stable-angina>.
- Sullivan, D. (2020). *Heart Failure*. Récupéré sur .healthline.com:
<https://www.healthline.com/health/Heart Failure>
- techno-science. (2023, Mars 25). *techno-science*. Récupéré sur techno-science:
<https://www.techno-science.net/definition/3690.html>
- V. Miller, A. M. (2017). Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (pure) : a prospective cohort study. *The Lancet*, 2037–2049.
- vie-publique-fiche thématique. (2023, Mars 22). Récupéré sur vie publique: <https://www.vie-publique.fr/fiches/37853-definition-et-acteurs-du-systeme-de-sante-francais>
- Williams, D. E. (1986). Learning InternalRepresentations by Error Propagation, Parallel Distributed Processing. *Explorations in the Microstructure of Cognition*, vol1, 318-362.
Récupéré sur [https://www.bibsonomy.org/bibtex/](https://www.bibsonomy.org/bibtex)
- X. Xie, E. A. (2016). Effects of intensive blood pressure lowering on cardiovascular and renal outcomes : updated systematic review and meta-analysis. *the lancet*, 435–443.
- Yann Lecun, L. B. (1998). Gradient-based learning applied to document recognition.
Proceedings of the IEEE, 2278-2324.
- Zorha, B. S. (2019). *Système d'information Hospitalier: Admission et planification des bloc opératoires*.