République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Chadli Bendjedid
Faculté des Sciences et de la Technologie
Département d'Informatique

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة الشاذلي بن جديد
كلية العلوم و التكنولوجيا
قسم الاعلام الالي

جامعة الشاذلي بن جديد
UNIVERSITE CHADLI BENDJEDID

MEMOIRE

Présenté par

AIT HAMOU LIZA

Pour l'obtention de diplôme de

MASTER

Filière : Informatique

Spécialité : Systèmes Informatiques Intelligents

Thème

Apprentissage automatique pour la prédiction des maladies cardiaques

Soutenu le : / / 2024 **Devant le Jury composé de :**

Qualité	Nom et Prénom	Grade	Université
Président	Mlle.Ziani	MCB	Chadli Bendjedid El-Tarf
Rapporteur	**Mme. MILOUD AOUIDATE-TACHOUCHE AMAL**	**MCB**	**Chadli Bendjedid El-Tarf**
Examinateur	Mr. Messaoudi	MAB	Chadli Bendjedid El-Tarf

Année Universitaire : 2023/2024

Remerciements

Je tiens à remercier tout d'abord **le bon Dieu** de m'avoir donné la foi et la volonté d'aller jusqu'au bout pour terminer ce travail de recherche.

Je remercie également mon encadreur **Mme. TACHOUCHE AMAL**, qui a su me comprendre et m'a offert, sans hésiter, son aide, son temps et ses conseils qui m'ont été précieux.

Je remercie également **les membres du jury** d'avoir accepté d'examiner ce travail.

Et toutes les personnes qui nous ont soutenus, et apportés leur aide de près ou de loin à sa réalisation.

Dédicace

Je dédie ce travail À **ma chère mère et à mon cher père** qui n'ont jamais cessé de me soutenir et m'encourager dans mes études.

À **ma seule sœur.**

Et toutes les personnes qui nous ont soutenus, et apportés leur aide de près ou de loin à sa réalisation.

La table des matières

Remerciements 1

Dédicace 2

Table des matières 3

Liste des figures 8

Liste des tableaux 9

Liste des abréviation 10

Résumé 11

Abstract 12

ملخص 13

Introduction générale 14

1. Contexte du projet et problématique 14

2. Motivations 15

3. Objectifs 15

4. Contenu du mémoire 15

Chapitre 1 : Généralités sur les maladies cardiaques 17

Introduction 18

1.Définition 18

2. Les types de maladies cardiaques 18

a. Les maladies coronariennes 18

b. Les maladies cérébrovasculaires 18

c. Les maladies artérielles périphériques 18

d. La cardiopathie rhumatismale 19

e. La cardiopathie congénitale 19

f. La thrombose veineuse profonde et embolie 19

3. Les facteurs de risque des maladies cardiovasculaires 19

a. L’hypertension artérielle.....19
b. Le diabète 20
c. La dyslipidémie 20
d. Le tabagisme.....20
e. L'alimentation malsaine.....20
f. La sédentarité 20
g. L'embonpoint et l'obésité 20
4. Symptômes des maladies cardiaques.....20
5.Les recommandations pour les maladies cardio-vasculaires.....21
6.Le diagnostic 22
6.1. ECG (électrocardiogramme).....22
6.2. Les analyses de sang.....23
6.3. L’échocardiographie (écho).....23
6.4. L’angiographie coronarienne et cathétérisme cardiaque.....24
6.5. La tomodensitométrie cardiaque (CT).....25
6.6. La scintigraphie de perfusion myocardique (MPS).....26
6.7. L'imagerie par résonance magnétique cardiaque (IRM cardiaque).....26
7. L'électrocardiogramme (ECG) 27
7.1.Définition.....27
7.2. Indications 28
7.3.Les différents type D’ECG.....28
7.4. Les ondes et les intervalles de l’ECG.....29
Conclusion.....30
Chapitre 02 : L'apprentissage automatique 31
Introduction 32
1.Définition de l'apprentissage automatique 32
2.Les différents types d’apprentissage automatique.....32

2.1. L'apprentissage supervisé 32
2.2. L'apprentissage non supervisé 33
2.3. L'apprentissage semi-supervisé 33
2.4. L'apprentissage par renforcement 33
3. Quelques algorithmes de l'apprentissage automatique 33
3.1. Arbre de décision (DT) 33
3.1.1. Les avantages de l'arbres de décision.....34
3.1.2. Les inconvénients 34
3.2. Forêt aléatoire (RF) 34
3.2.1. Définition.....34
3.2.2. Les avantages.....35
3.2.3. Les inconvénients 35
3.3. Les K plus proches voisins (KNN).....35
3.3.1. Définition.....35
3.3.2. Quelques règles sur le choix de k 35
3.3.3. Les avantage 36
3.3.4. Les inconvénients 36
4. La logique floue.....36
4.1. Définition.....36
4.2. Intérêt de la logique floue.....37
4.3. La structure générale de la logique floue.....37
4.3.1. L'interface de fuzzification.....38
4.3.2. La base de connaissances 38
4.3.3. Le moteur d'inférence floue 38
4.3.4. L'interface de défuzzification.....39
4.4. Les avantage de la logique flou 39
4.5. Les inconvénients de la logique flou 40

5. La logique floue VS la logique classique 40
6. La logique floue dans la prédiction des maladies cardiaque 40
Conclusion 42
Chapitre 03 : Conception et réalisation 43
Introduction 44
1. L'environnement de développement 44
1.1.La platforme de dévelopement 44
a. Google colab 44
b. PyCharm 44
1.2. Le langage de programmation 45
a. Python 45
1.3 Les bibliothèques de Python 46
a. NumPy 46
b. Pandas 46
c. Scikit-learn 46
d. Matplotlib 46
2. Le système proposé 47
2.1. La base de données 48
2.1.1. Description de la base de données 48
2.1.2. Description des attributs de la base de données 48
2.2. L'analyse exploratoire des données 50
2.3. Prétraitement de données 51
2.4. Répartiotion des données 52
2.5. Modèle proposé 53
2.5.1. La logique floue 53
a. La fuzzification 53
b. Inférence floue 60

c. La défuzzification 63
2.5.2. K plus proches voisins (KNN) 64
2.5.3. Hybidration de La logique floue et KNN 65
2.6. Évaluation des résultats et discussion.....66
Conclusion 69
Conclusion générale 70
Bibliographies.....71

Liste des figures

Numéro de figure	Intitulé	Page
Figure 1	Les types des maladies cardiaques.	19
Figure 2	L'appareil de ECG.	22
Figure 3	Les analyses de sang.	23
Figure 4	L'appareil d'échocardiographie (écho).	24
Figure 5	La coronarographie.	25
Figure 6	La tomodensitométrie cardiaque.	26
Figure 7	L'IRM.	27
Figure 8	Les ondes et intervalles de l'ECG.	30
Figure 9	Le logo du PyCharm.	45
Figure 10	Le logo du Python.	45
Figure 11	L'architecture du système proposé.	47
Figure 12	L'extrait de la base de données.	48
Figure 13	Les informations de la base des données.	50
Figure 14	La visualisation de la classe cible.	51
Figure 15	Les valeurs manquantes	52
Figure 16	La répartition des données de train/test.	52
Figure 17	La fuzzification de l 'âge.	53
Figure 18	La fuzzification de sexe.	54
Figure 19	La fuzzification de cp.	54
Figure 20	La fuzzification de trestbps.	55
Figure 21	La fuzzification de chol.	55
Figure 22	La fuzzification de fbs.	56
Figure 23	La fuzzification de restecg.	56
Figure 24	La fuzzification de thalach.	57
Figure 25	La fuzzification de exang.	57
Figure 26	La fuzzification de oldpeak.	58
Figure 27	La fuzzification de slope.	58
Figure 28	La fuzzification de ca.	59
Figure 29	La fuzzification de Thal.	59
Figure 30	La fuzzification de Target.	60
Figure 31	Le nombre optimal de k-voisins de KNN.	65
Figure 32	Le taux de précision de sain et malade.	66

Liste des Tableaux

Numéro de tableau	Intitulé	Page
Tableau 1	Les résultats d'évaluation des performances.	67
Tableau 2	La comparaison des résultats.	68

Liste des abréviations

Abréviation	Sens
MCV	Maladie Cardiovasculaire.
ECG	Électrocardiogramme.
Echo	Échocardiographie.
CT	Tomodensitométrie Cardiaque.
LDL	Low density lipoproteins.
HDL	High density lipoprotein.
MPS	Myocardial Perfusion Scintigraphy.
IRM	Imagerie par Résonance Magnétique.
ML	Machine Learning.
IA	Intelligence artificielle.
DT	Decision Tree.
RF	Random Forest.
KNN	K- Nearest Neighbour.
NumPy	Numerical Python.
CP	Chest pain type.
Trestbps	Resting Blood Pressure.
Chol	Cholesterol.
Fbs	Fasting Blood Sugar.
Restecg:	Resting Electrocardiographic Results.
Thalach	Maximum Heart Rate Achieved.
Exang	Exercise-Induced Angina.
Oldpeak	ST Depression Induced by Exercise.
Slope	Segment ST d'exercice de pointe.
CA	Number of Major Vessels Colored by Fluoroscopy.
THAL	Thallium Stress Test Result.

Résumé

Les maladies cardiaques constituent une menace sérieuse pour la vie humaine, elles doivent donc être prédites et traitées efficacement, là où l'intelligence artificielle a joué un rôle important dans le diagnostic de ces maladies grâce à l'une de ses techniques les plus importantes, l'apprentissage automatique.

Le but de notre travail est de créer un système de prédiction des maladies cardiaques en utilisant l'algorithme d'apprentissage supervisé K plus proches voisins (KNN), la logique floue, ainsi que la combinaison de la logique floue avec KNN, pour trouver le meilleur modèle pour un résultat maximal.

D'après notre étude, nous avons constaté que l'algorithme combinant la logique floue avec l'algorithme des k-plus proches voisins (KNN) fournit un bon compromis entre performance et taille des données de classification pour la prédiction des maladies cardiaques.

Cependant l'algorithme KNN présente le meilleur taux de précision par rapport a l'ensemble de données utilisé.

Mots clés : Apprentissage automatique, Maladies cardiaques, Maladies cardio-vasculaires, KNN, La logique floue.

Abstract

Heart diseases pose a serious threat to human life so they need to be predicted and treated effectively, where artificial intelligence has played an important role in diagnosing these diseases through one of its most important techniques, machine learning.

The goal of our work is to create a heart disease prediction system using the K nearest neighbors (KNN) supervised learning algorithm, fuzzy logic, as well as the combination of fuzzy logic with KNN, to find the best model for maximum result.

From our study, we proved that the algorithm combining fuzzy logic with the k-nearest neighbors (KNN) algorithm offers a good compromise between performance and data classification size for heart disease prediction.

However, the KNN algorithm provides the best accuracy rate compared to the dataset used.

Keywords: Machine learning, Heart diseases, Cardiovascular disease, KNN, Fuzzy logic.

ملخص:

تشكل أمراض القلب تهديدًا خطيرًا لحياة الإنسان لذا يجب التنبؤ بها وعلاجها بشكل فعال، ولقد لعب الذكاء الاصطناعي دورًا مهمًا في تشخيص هذه الأمراض من خلال إحدى أهم تقنياته، وهي التعلم الآلي.

الهدف من عملنا هو إنشاء نظام للتنبؤ بأمراض القلب باستخدام خوارزمية التعلم الخاضعة للإشراف K أقرب جيران (KNN)، والمنطق الضبابي ، وكذلك دمج المنطق الضبابي مع KNN، للعثور على أفضل نموذج لتحقيق أقصى قدر من النتائج.

من خلال دراستنا، وجدنا أن الخوارزمية التي تجمع بين المنطق الضبابي وخوارزمية K -أقرب الجيران (KNN) توفر توازنًا جيدًا بين الأداء وحجم بيانات التصنيف للتنبؤ بأمراض القلب.

ومع ذلك، فإن خوارزمية KNN تُظهر أفضل معدل دقة بالنسبة لمجموعة البيانات المستخدمة.

الكلمات المفتاحية: التعلم الآلي، أمراض القلب، الأمراض القلبية الوعائية، K أقرب الجيران (KNN)، ، المنطق الضبابي.

Introduction générale

Les maladies cardiaques sont considérées comme l'une des maladies les plus dangereuses qui menacent directement l'organe vital qu'est le cœur. Le diagnostic de ces maladies dépend des tests médicaux et de divers indicateurs, ce qui rend leur identification complexe et difficile en raison du grand nombre de tests nécessaires. Sachant que ces diagnostics sont parmi les plus exigeants et les plus coûteux, notamment avec la grande quantité des données qui seront analysées.

Ces dernières années, le domaine médical a connu un grand développement grâce aux technologies de l'information, notamment avec l'émergence de ce qu'on appelle l'intelligence artificielle (IA), qui regroupe des technologies visant à simuler la logique humaine et à s'en rapprocher. Par exemple, la contribution de l'intelligence artificielle à la détection des maladies cardiaques peut être constatée à travers l'une de ses branches les plus importantes, à savoir l'apprentissage automatique.

Des études ont montré que l'utilisation de méthodes d'apprentissage automatique peut être plus efficace que les modèles statistiques traditionnels pour prédire le risque de récidive de la maladie chez les patients ayant déjà souffert d'une maladie cardiaque.

1. Contexte du projet et problématique

Le processus de diagnostic et de prévision des risques de maladies cardiaques basé sur la méthode traditionnelle, c'est-à-dire s'appuyant sur l'intuition, les connaissances et l'expérience du médecin, est devenu insuffisant en raison de la grande quantité de données, de la grande complexité des problèmes médicaux, des capacités limitées du médecin. capacité à résoudre certains problèmes médicaux...etc. Tous ces facteurs provoquent des erreurs de diagnostic dans les maladies cardiaques, il est donc nécessaire d'utiliser de nouvelles techniques d'analyse pour créer des modèles facilitant la prise de décision.

2. Motivations

Les progrès de l'intelligence artificielle créent de nouvelles opportunités pour améliorer la prise de décision médicale. Cela est dû à la capacité des algorithmes d'IA à analyser de grandes quantités de données qui échappent souvent à l'œil humain. Cette capacité d'analyse approfondie peut aider les médecins à établir des diagnostics plus précis, à prédire les résultats du traitement et à développer des stratégies de traitement plus efficaces.

L'apprentissage automatique est l'une des disciplines fondamentales de l'intelligence artificielle (IA). Il comprend un ensemble de techniques et d'algorithmes qui ont démontré leurs efficacités dans le domaine de la prédiction des maladies cardiaques.

3. Objectifs

Notre objectif principal dans ce projet est de créer un système de prédiction des maladies cardiaques en utilisant l'algorithme d'apprentissage supervisé K plus proches voisins (KNN) et la logique floue, ainsi que la combinaison de ces deux techniques. Ces méthodes seront utilisées pour prédire les maladies cardiaques à partir des données des patients, afin de trouver le meilleur modèle pour un résultat maximale.

4. Contenu du mémoire

Pour atteindre notre objectif, notre mémoire est organisé comme suit :

Le premier chapitre intitulé : « Généralités sur les maladies cardiaques ». Dans ce chapitre nous présentons une introduction générale sur les maladies cardiaques, en mettant l'accent sur ses types, les facteurs de risque des maladies cardiovasculaires, leurs symptômes, les recommandations pour prévenir de ces maladies. Ainsi que les techniques de diagnostic utilisées.

Le deuxième chapitre intitulé : « l'apprentissage automatique ». Ce chapitre est consacré à l'apprentissage automatique, ses types et ses différents algorithmes, en mentionnant les avantages et les inconvénients de chaque algorithme. Ensuite, nous présenterons la logique floue.

Le troisième chapitre intitulé : « conception et réalisation ». Nous avons présenté la conception et la réalisation d'un système de prédiction des maladies cardiaques et nous évaluons les résultats de notre système proposé.

Chapitre 01 : Généralités sur les maladies cardiaques

Introduction

Actuellement, les maladies cardiovasculaires sont de plus en plus répandues dans le monde, ce qui a contribué à une augmentation significative de la mortalité. Pour lutter contre cette crise, il est nécessaire de renforcer le diagnostic.

Dans ce chapitre, nous aborderons le concept de maladie cardiaque, ses types, les facteurs de risque des maladies cardiovasculaires, leurs symptômes, les recommandations pour prévenir de ces maladies. Ainsi que les techniques de diagnostic utilisées. Nous terminerons le chapitre en nous concentrant sur l'électrocardiogramme (ECG).

1. Définition des maladies cardiaques

Les maladies cardiaques ou les maladies cardiovasculaires (MCV) sont des affections qui affectent le cœur et les vaisseaux sanguins. Elles regroupent un ensemble de problèmes de santé liés à la circulation sanguine, tels que l'athérosclérose, les maladies coronariennes, les accidents vasculaires cérébraux, les maladies valvulaires et les troubles du rythme cardiaque.

Ces affections peuvent entraîner des symptômes tels que des douleurs thoraciques, des essoufflements, des palpitations et une fatigue excessive. Les maladies cardiovasculaires sont une préoccupation majeure en matière de santé publique et nécessitent des mesures de prévention, de détection précoce et de prise en charge appropriées. [W1]

2. Les types de maladies cardiaques

Il existe différents types de maladies cardiovasculaires, notamment [W2] :

a. **Les maladies coronariennes :** Représente des affections des vaisseaux sanguins qui fournissent du sang au muscle cardiaque.
b. **Les maladies cérébrovasculaires :** Ce sont des affections des vaisseaux sanguins qui fournissent du sang au cerveau.
c. **Les maladies artérielles périphériques :** Elles affectent les vaisseaux sanguins qui irriguent les bras et les jambes. Il s'agit d'une affection résultant d'un rétrécissement ou d'un blocage des artères périphériques, entraînant une réduction du flux sanguin vers les extrémités.

d. La cardiopathie rhumatismale : Elle englobe l'ensemble des dommages au muscle cardiaque et aux valves cardiaques causés par une inflammation résultant d'une infection bactérienne appelée rhumatisme articulaire aigu, généralement causée par des bactéries streptococciques.

e. La cardiopathie congénitale : C'est des anomalies structurelles qui apparaissent à la naissance et affectent la croissance et le fonctionnement normaux du cœur. Ils peuvent inclure des anomalies valvulaires, des anomalies septales ou des problèmes avec les vaisseaux sanguins du cœur.

f. La thrombose veineuse profonde et embolie pulmonaire : Elle est caractérisée par la formation de caillots sanguins dans les veines profondes des jambes, ce qui peut entraîner une obstruction partielle ou totale des vaisseaux sanguins.

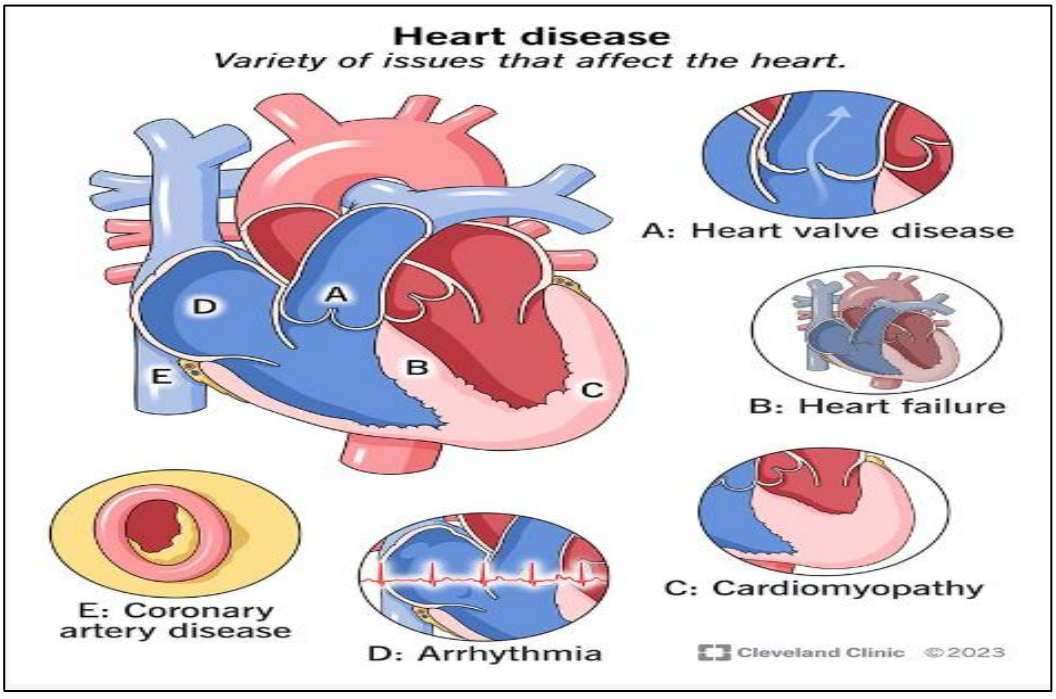

Figure 1 : Les types des maladies cardiaques**. [W3]**

3. Les facteurs de risque des maladies cardiovasculaires

Les maladies cardiovasculaires sont influencées par divers facteurs de risque. Certains des facteurs de risque les plus courants sont répertoriés ci-dessous : [1]

a. L'hypertension artérielle : L'hypertension artérielle (ou tension artérielle élevée), exerce une pression supplémentaire sur les parois artérielles, ce qui peut endommager les vaisseaux sanguins et augmenter le risque de maladie cardiaque.

b. **Le diabète :** Le diabète, en particulier le diabète de type 2, est souvent associé à d'autres facteurs de risque de maladie cardiaque. Une glycémie élevée peut endommager les vaisseaux sanguins et les nerfs du cœur, augmentant ainsi le risque de maladie cardiaque.
c. **La dyslipidémie :** La dyslipidémie fait référence à des taux anormaux de lipides (cholestérol et triglycérides) dans le sang. Des niveaux élevés de cholestérol LDL (mauvais cholestérol) et de triglycérides, ainsi que des niveaux bas de cholestérol HDL (bon cholestérol), peuvent contribuer à l'accumulation de plaque dans les artères et augmenter le risque de maladies cardiaques.
d. **Le tabagisme :** Le tabagisme est un facteur de risque majeur de maladie cardiaque. Les substances chimiques présentes dans la fumée de tabac endommagent les parois des artères, augmentent la formation de caillots sanguins et réduisent la quantité d'oxygène disponible pour le cœur.
e. **L'alimentation malsaine :** Une alimentation riche en gras saturés, en cholestérol, en sel et en sucres ajoutés peut contribuer au développement de maladies cardiaques. La consommation excessive de gras saturés peut augmenter les niveaux de cholestérol LDL, tandis qu'une consommation élevée de sel peut augmenter la tension artérielle.
f. **La sédentarité :** Un mode de vie sédentaire, caractérisé par une activité physique insuffisante, est associé à un risque accru de maladie cardiaque.
g. **L'embonpoint et l'obésité :** L'excès de poids et l'obésité sont liés à un risque accru de maladies cardiaques. L'excès de graisse corporelle peut augmenter la pression artérielle, les taux de cholestérol et les risques de diabète, contribuant ainsi au développement de maladies cardiaques.

4.Symptômes des maladies cardiaques

Les symptômes couramment associés aux maladies cardiaques [2] :

- Oppression cardiaque, pression et douleur.
- Douleurs dans la cage thoracique, les bras, le cou, la mâchoire et le dos.
- Difficulté à respirer.
- Étourdissements ou sensation de vertige.
- Agitation.
- Changement de couleur de la peau, notamment le visage qui devient brun.

Cependant, ces symptômes ne suffisent pas à confirmer la présence d'une maladie cardiaque. Des tests et des examens médicaux sont nécessaires pour obtenir des informations supplémentaires permettant de détecter la maladie. Ces tests fournissent des attributs essentiels qui nous aident à établir un diagnostic précis. [W2]

5.Les recommandations pour les maladies cardio-vasculaires

Les recommandations pour la prévention des maladies cardiovasculaires et le maintien d'une bonne santé cardiovasculaire sont essentiels. Voici quelques recommandations clés [W4] :

- Adopter une alimentation saine, privilégiant une consommation élevée en fruits, légumes et céréales complètes, tout en limitant les gras saturés, le sel et le sucre ajouté.
- Pratiquer régulièrement une activité physique, en visant au moins 150 minutes par semaine d'exercice d'intensité modérée à vigoureuse.
- Maintenir un poids santé en suivant un régime équilibré et en intégrant l'exercice physique dans votre routine.
- Éviter le tabagisme, car il constitue un facteur de risque majeur pour les maladies cardiovasculaires. Chercher des stratégies d'arrêt adaptées aux besoins en demandant un soutien médical si nécessaire.
- Contrôler sa pression artérielle en adoptant un mode de vie sain, comprenant une alimentation équilibrée, une activité physique régulière et, si nécessaire, la prise de médicaments prescrits par un professionnel de la santé.
- Gérer le stress en pratiquant des techniques de relaxation, telles que la méditation, la respiration profonde ou le yoga, et en trouvant des moyens sains de faire face au stress quotidien.
- Suivre les recommandations médicales, prendre régulièrement les médicaments prescrits et effectuer les examens de dépistage recommandés pour détecter toute anomalie cardiovasculaire le plus tôt possible.

6. Le diagnostic

Le diagnostic des maladies cardiovasculaires peut être réalisé à l'aide de différentes méthodes. Voici quelques méthodes couramment utilisées :

6.1. ECG (électrocardiogramme)

L'électrocardiographie (ECG) est une technique diagnostique couramment utilisée en cardiologie pour évaluer l'activité électrique du cœur. C'est une procédure non invasive, rapide et indolore qui permet d'amplifier et d'enregistrer les impulsions électriques générées par le muscle cardiaque.

L'ECG se base sur le placement d'électrodes sur la peau du patient, généralement sur les bras, les jambes et le thorax, qui captent les signaux électriques du cœur à chaque battement. Ces électrodes mesurent l'amplitude et la direction de ces courants électriques cardiaques, permettant ainsi de créer un tracé graphique appelé électrocardiogramme.

L'ECG est utilisé dans de nombreux contextes cliniques, que ce soit en cas de suspicion de maladie cardiaque, lors d'examens de routine chez des patients d'âge moyen ou chez les personnes âgées, ou encore lors de l'évaluation préopératoire avant une intervention chirurgicale. **[W5]**

Figure 2 : L'appareil de ECG **[3]**

6.2. Les analyses de sang

Les analyses de sang sont des tests de laboratoire importants pour évaluer la santé globale et détecter des affections spécifiques, y compris les maladies cardiaques. Ils mesurent les marqueurs cardiaques tels que la troponine et la créatine kinase, qui révèlent des lésions ou un dysfonctionnement du muscle cardiaque. La troponine, en particulier, est un marqueur spécifique au muscle cardiaque dont les niveaux augmentent après un infarctus du myocarde. Ces tests sanguins fournissent des informations essentielles pour le diagnostic et le suivi des maladies cardiovasculaires. [4]

Figure 3 : Les analyses de sang[W6]

6.3. L'échocardiographie (écho)

L'échocardiographie, également connue sous le nom d'écho, est une technique d'imagerie diagnostique qui utilise des ondes ultrasonores pour fournir une représentation visuelle de l'anatomie cardiaque. Elle peut être réalisée soit par échographie transthoracique (ETT), avec un transducteur externe placé sur la paroi thoracique, soit par échographie transœsophagienne (ETO), qui implique l'insertion du transducteur dans l'œsophage du patient pour obtenir des informations plus détaillées.

L'échocardiographie permet d'obtenir des informations sur la taille et la forme des cavités cardiaques, les vitesses de flux sanguin, la fonction systolique et diastolique, la

contractilité, les anomalies de la motilité des parois, la fraction d'éjection, la fonction valvulaire et la présence de thrombus dans les cavités cardiaques.

Elle est considérée comme la méthode de référence pour le diagnostic de l'insuffisance cardiaque et devrait être réévaluée après l'ajustement des médicaments, puis au moins tous les deux ans par la suite. L'échocardiographie permet également de déterminer le type d'insuffisance cardiaque en fonction de la fraction d'éjection, d'identifier les anomalies régionales de la motilité des parois et la dyssynchronie des parois, ainsi que d'évaluer la fonction valvulaire. [5]

Figure 4 : L'appareil d'échocardiographie (écho) **[W7]**

6.4. L'angiographie coronarienne et cathétérisme cardiaque

L'angiographie coronarienne est une méthode d'imagerie invasive dans laquelle un colorant de contraste est injecté dans les artères coronaires par le biais d'un cathéter. Ce colorant permet d'obtenir des images détaillées des vaisseaux sanguins du cœur, ce qui permet de détecter d'éventuelles obstructions ou rétrécissements des artères coronaires. Cette procédure, également connue sous le nom de coronarographie, est couramment utilisée pour diagnostiquer les maladies coronariennes telles que les sténoses ou les blocages des artères coronaires.

Le cathétérisme cardiaque est une procédure dans laquelle un cathéter mince et flexible est inséré dans une artère ou une veine, généralement dans l'aine, le poignet ou le cou, et guidé jusqu'au cœur. Cette procédure permet de mesurer les pressions à l'intérieur du cœur, d'évaluer la fonction cardiaque et de recueillir des échantillons de sang pour analyse. Le cathétérisme cardiaque peut également être utilisé pour effectuer des interventions telles que l'angioplastie coronaire, qui consiste à dilater une sténose coronaire, ou la pose d'un stent pour maintenir l'ouverture d'une artère coronaire. [4]

Figure 5 : La coronarographie **[W8]**

6.5. La tomodensitométrie cardiaque (CT)

La tomodensitométrie cardiaque (CT) utilise la technologie CT pour obtenir des images détaillées du cœur. Elle permet de détecter des anomalies anatomiques, d'évaluer l'état des vaisseaux coronaires et de fournir des informations sur la présence de plaque calcifiée. Cette procédure est utile pour évaluer la perméabilité des greffes après une chirurgie de pontage coronarien et pour diagnostiquer les maladies coronariennes. Elle peut également guider les décisions de prise en charge et les investigations suivantes. [4]

Figure 6 : La tomodensitométrie cardiaque **[W9]**

6.6. La scintigraphie de perfusion myocardique (MPS)

La scintigraphie de perfusion myocardique (MPS), également appelée scintigraphie au sestamibi ou au thallium, est une technique de médecine nucléaire non invasive. Elle permet d'évaluer la perfusion du muscle cardiaque au repos et sous stress à l'aide d'un traceur radioactif. Cette procédure peut être réalisée après un exercice physique ou en utilisant des médicaments pour augmenter le flux sanguin vers le cœur.

La MPS aide à diagnostiquer la sévérité de la maladie coronarienne et fournit des informations sur la prise en charge, telle que la nécessité d'une angiographie ou d'une revascularisation des artères coronaires. Elle peut également être utilisée pour identifier les patients présentant une ischémie récurrente après une intervention de revascularisation. [4]

6.7. L'imagerie par résonance magnétique cardiaque (IRM cardiaque)

L'imagerie par résonance magnétique cardiaque (IRM cardiaque) utilise des champs magnétiques et des radiofréquences pour produire des images 3D haute résolution du cœur. Il fournit des informations précises sur le volume cardiaque, la fonction systolique, les cicatrices tissulaires et la fraction d'éjection.

L'IRM cardiaque peut également identifier et caractériser un infarctus du myocarde. Il peut identifier les anomalies de mouvement de la paroi cardiaque, les problèmes valvulaires, la fibrose myocardique et aider à diagnostiquer diverses maladies cardiaques telles que la cardiomyopathie amyloïde, la myocardite et la sarcoïdose cardiaque. [4]

Figure 7 : L'IRM **[W10]**

7. L'électrocardiogramme (ECG)

7.1. Définition

Un électrocardiogramme (ECG) est une technique qui analyse l'activité électrique du cœur, présentée sous forme de trace graphique. Chaque fois que le cœur se contracte, une impulsion électrique, ou « onde », traverse le muscle cardiaque, provoquant l'expulsion du sang du cœur. Un électrocardiogramme utilise des électrodes placées sur la peau pour mesurer les changements de potentiel électrique produits par le muscle cardiaque. Ces changements sont enregistrés et représentés graphiquement pour permettre une analyse précise de l'activité électrique du cœur.

Ce test non invasif et indolore est essentiel au diagnostic et à la surveillance des maladies cardiaques. Il permet aux médecins de déterminer si l'activité électrique observée est normale ou irrégulière, ainsi que de détecter et d'analyser des anomalies qui pourraient ne pas apparaître lors d'une consultation médicale. [6]

7.2. Indications

L'électrocardiogramme (ECG) est un outil clinique non invasif et essentiel qui permet d'identifier diverses conditions cardiaques et problèmes de santé. Il est utilisé pour détecter :

- Les anomalies du rythme cardiaque, y compris les arythmies et les troubles du rythme.
- Les problèmes au niveau du système de conduction cardiaque.
- L'ischémie myocardique, indiquant une diminution de l'apport sanguin au muscle cardiaque.
- La péricardite, une inflammation de la membrane entourant le cœur.
- Les embolies pulmonaires, des caillots sanguins dans les artères pulmonaires.
- L'hypertension artérielle.
- Les perturbations métaboliques affectant le fonctionnement cardiaque.
- Le suivi de l'efficacité de certains traitements, notamment les médicaments antiarythmiques.
- Ces indications montrent l'importance de l'ECG dans le diagnostic et la prise en charge des affections cardiaques et des troubles connexes. [7]

7.3. Les différents types D'ECG

Il existe plusieurs types d'ECG, chacun adapté à des circonstances particulières [3] :

- **ECG au repos :** Cette procédure dure généralement entre 5 et 10 minutes et se déroule alors que le patient est allongé en position de repos, sans bouger. Elle vise à enregistrer l'activité cardiaque dans un état de calme.
- **ECG d'effort :** D'une durée de 30 à 45 minutes, cet examen est réalisé pendant que le patient s'engage dans un exercice physique, généralement sur un vélo ergométrique ou un tapis roulant. Ce type d'ECG permet d'évaluer la réponse du cœur sous contrainte physique.
- **ECG de longue durée :** Connu sous le nom d'ECG de Holter, il enregistre l'activité cardiaque de manière continue sur une période de 24 heures ou plus. Pour ce faire, trois électrodes ou plus sont fixées sur le thorax du patient et reliées à un enregistreur portable qui collecte les données. Les informations recueillies sont ensuite analysées en laboratoire pour détecter d'éventuelles anomalies cardiaques qui ne se manifestent pas durant les enregistrements plus courts.

Chaque type d'ECG fournit des informations précieuses pour le diagnostic et le suivi des conditions cardiaques, permettant aux professionnels de santé de recommander des traitements appropriés en fonction des résultats obtenus.

7.4. Les ondes et les intervalles de l'ECG

L'électrocardiogramme (ECG) capture les séquences de dépolarisation et de repolarisation des muscles auriculaires et ventriculaires. Cette activité est tracée sous forme d'ondulations ou de déflexions par rapport à une ligne de potentiel zéro, désignée comme la ligne isoélectrique ou ligne de base. Les déflexions au-dessus de cette ligne sont considérées positives, tandis que celles situées en dessous sont négatives. Un cycle cardiaque sur l'ECG est caractérisé par plusieurs ondes et différents intervalles qui les séparent : [8]

- **Onde P :** Indique la dépolarisation auriculaire, présentant une faible amplitude et une forme arrondie. En rythme sinusal, les ondes P précèdent systématiquement les complexes ventriculaires et sont exclusivement positives.
- **Intervalle PR (ou PQ) :** Ce délai, qui s'étend du début de l'onde P au début du complexe QRS, représente le temps nécessaire pour la transmission de l'impulsion électrique des oreillettes vers les ventricules.
- **Segment PR :** Un court segment isoélectrique qui isole l'onde P du complexe ventriculaire suivant.
- **Complexe QRS :** Il reflète la dépolarisation ventriculaire à travers une série de déflexions rapides :
  - ✓ **Onde Q :** Typiquement de faible amplitude, ne dépassant pas un tiers de l'amplitude totale du complexe QRS.
  - ✓ **Onde R :** Définie comme la première déflexion positive du complexe, qu'elle soit ou non précédée par une onde Q.
  - ✓ **Onde S:** Une onde négative suivant immédiatement l'onde R.
- **Segment ST :** Correspond au commencement de la repolarisation ventriculaire, et se trace de manière horizontale à partir de la ligne de base.
- **Onde T :** Signale la repolarisation ventriculaire. Son orientation est généralement conforme à celle du complexe QRS et elle se caractérise par une asymétrie notable avec une montée graduelle, un sommet arrondi et une descente abrupte.

- **Intervalle QT :** Mesuré depuis le début du complexe QRS jusqu'à la fin de l'onde T, cet intervalle englobe les périodes complètes de dépolarisation et de repolarisation cardiaque.

Les amplitudes normales pour les ondes P, Q, S, et T sont environ 0.1 mV, tandis que l'onde R peut parfois atteindre ou excéder 1 mV. La fréquence cardiaque est déterminée par la mesure de l'intervalle entre les complexes RR.

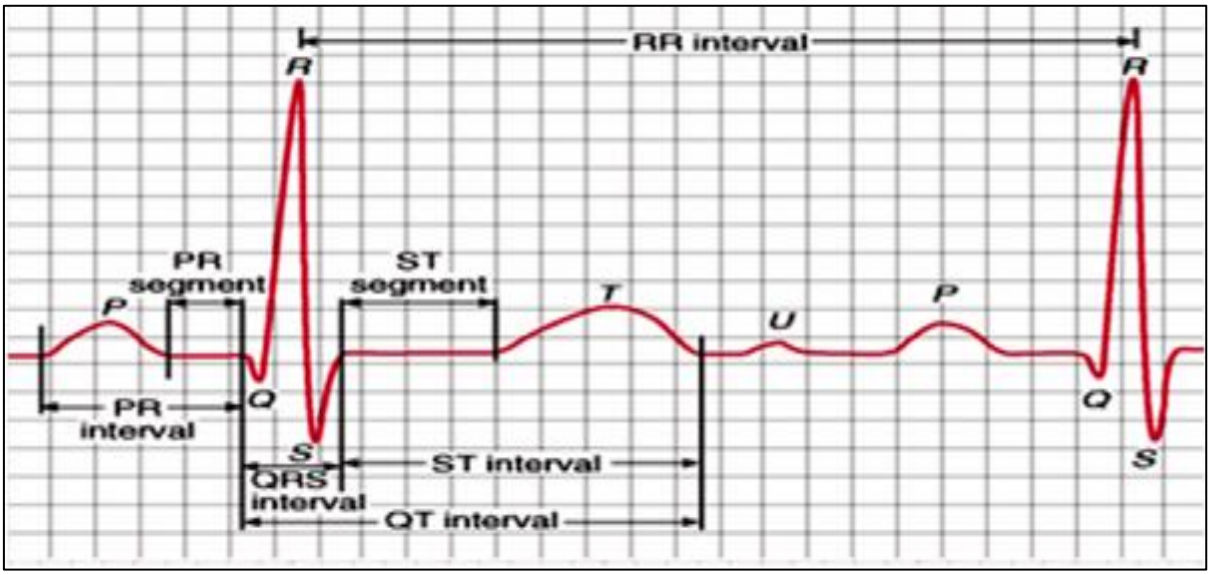

Figure 8 : Les ondes et intervalles de l'ECG **[W11]**

Conclusion

Dans ce chapitre, nous avons essayé de faire une description assez détaillée des maladies cardiaque. Cela a permis de mieux comprendre les techniques de diagnostic cardiaque qui utilisent des outils et des appareils médicaux tels que l'échocardiographie, l'image stéthoscopique, l'ECG et l'IRM.

Dans le chapitre suivant, nous nous intéresserons au domaine de l'apprentissage automatique et à ses différentes méthodes.

Chapitre 02 : L'apprentissage automatique

Introduction

L'apprentissage automatique est un domaine de l'intelligence artificielle qui permet aux machines d'apprendre et d'évoluer grâce à un processus d'apprentissage basé sur les données. Il permet à la machine de découvrir des modèles et des relations dans les données pour effectuer des tâches complexes.

Dans ce chapitre, nous explorerons l'apprentissage automatique, ses types et ses différents algorithmes, en mentionnant les avantages et les inconvénients de chaque algorithme. Ensuite, nous présenterons la logique floue.

1. Définition de l'apprentissage automatique

L'apprentissage automatique, également connu sous le nom de "machine Learning" en anglais, est une branche de l'intelligence artificielle qui se concentre sur le développement de techniques et d'algorithmes permettant aux systèmes informatiques d'apprendre à partir de données et d'améliorer leurs performances sans être explicitement programmés.

L'objectif principal de l'apprentissage automatique est de permettre aux machines de détecter des modèles, des relations ou des structures cachées dans les données, et d'utiliser ces informations pour prendre des décisions, faire des prédictions ou accomplir des tâches spécifiques. Il repose sur l'analyse statistique et algorithmique des données, ainsi que sur l'utilisation de modèles et d'approches mathématiques pour représenter et traiter l'information. [9]

2. Les différents types d'apprentissage automatique

L'apprentissage automatique (ou machine Learning) comprend différentes approches permettant aux machines d'apprendre à partir des données et d'améliorer leurs performantes avec le temps. Voici les principaux types d'apprentissage automatique [9] :

2.1. L'apprentissage supervisé :

Est un type d'apprentissage automatique dans lequel des modèles sont préalablement définis et les systèmes apprennent à partir de paires d'entrée et de sortie fournies, c'est-à-dire que les données d'entrée et les sorties souhaitées sont étiquetées. Avec suffisamment de données et de connaissances, on peut aider la machine à établir des liens avec l'apprentissage

supervisé en utilisant des données d'échantillons étiquetées et les sorties correctes correspondantes.

2.2. L'apprentissage non supervisé :

L'apprentissage non supervisé est une méthode d'apprentissage automatique où l'intelligence artificielle apprend sans référence à des valeurs cibles prédéfinies. Le modèle cherche à identifier des schémas dans des données d'entrée non étiquetées. Contrairement à l'apprentissage supervisé, il n'y a pas de données étiquetées pour guider l'apprentissage. On utilise divers algorithmes pour permettre à la machine de détecter des modèles en examinant les données. Cette méthode offre la possibilité à la machine de découvrir des structures ou des classifications que les humains pourraient ne pas percevoir.

2.3. L'apprentissage semi-supervisé :

Est une méthode qui mélange les avantages de deux types d'apprentissage : supervisé et non supervisé. Cela permet à la machine d'apprendre avec un petit ensemble de données et d'utiliser ensuite ce qu'elle a appris pour comprendre de nouvelles informations sans avoir besoin de beaucoup de données supplémentaires.

2.4. L'apprentissage par renforcement :

Est un processus où le modèle interagit avec un environnement dynamique et reçoit des retours sous forme de récompenses et de punitions. Il apprend par des interactions positives et négatives avec son environnement. Contrairement aux autres méthodes, ce type d'apprentissage implique une itération continue où le modèle s'améliore progressivement pour produire des résultats de haute qualité.

3. Quelques algorithmes de l'apprentissage automatique

Voici quelques-uns des principaux d'algorithmes d'apprentissage automatique :

3.1. Arbre de décision (DT) :

Un arbre de décision est un schéma en forme d'arbre utilisé pour prendre des décisions en fonction des caractéristiques des données. Chaque nœud de l'arbre pose une question sur

une caractéristique, chaque branche représente une réponse possible à cette question, et chaque feuille donne une conclusion ou une prédiction basée sur les réponses données. [10]

3.1.1. Les avantages de l'arbres de décision

Les arbres de décision sont largement utilisés dans les domaines de l'apprentissage automatique et de l'analyse de données en raison de leurs nombreux avantages. Voici quelques-uns des principaux avantages des arbres de décision : [11]

- Simple et facile à comprendre et à interpréter
- Les arbres peuvent être visualisés.
- Nécessite peu de préparation des données.
- Adaptabilité à différents types de problèmes.

3.1.2. Les inconvénients

Les inconvénients d'algorithme arbres de décision sont : [11]

- Instabilité : De petits changements dans les données peuvent complètement modifier l'arbre, rendant les prédictions moins fiables.
- Difficulté à exprimer des idées complexes : Certains concepts peuvent être difficiles à représenter clairement avec un arbre de décision.
- Faibles performances avec des données continues.

3.2. Forêt aléatoire (RF)

3.2.1. Définition

La Forêt aléatoire (RF) est un modèle d'apprentissage automatique qui combine plusieurs arbres de décision pour des tâches de classification et de régression. Chaque arbre est construit en utilisant un sous-ensemble aléatoire des caractéristiques disponibles.

La RF utilise ensuite ces arbres pour améliorer la précision des prédictions en prenant en compte la moyenne de leurs résultats. C'est un algorithme robuste qui offre une meilleure précision que les arbres de décision individuels et qui peut résoudre des problèmes de classification sans nécessiter de normalisation des données. [12]

3.2.2. Les avantages

Les avantages d'algorithme RF sont [13] :

- La Forêt aléatoire produit des prédictions précises et compréhensibles.
- Elle est capable de traiter efficacement de grands ensembles de données.
- La Forêt aléatoire offre une précision supérieure dans la prédiction des résultats par rapport à l'algorithme de l'Arbre de décision.

3.2.3. Les inconvénients

Les inconvénients d'algorithme RF sont [13] :

- Temps de calcul plus long.
- Complexité de visualisation.
- Difficulté de compréhension

3.3. Les K plus proches voisins (KNN)

3.3.1. Définition

Le KNN (k plus proches voisins ou K-Nearest Neighbors) est un algorithme d'apprentissage supervisé largement utilisé qui attribue une classe à un nouvel exemple en se basant sur les classes des exemples voisins les plus proches dans l'espace des caractéristiques, en utilisant généralement la distance euclidienne. Il appartient à la catégorie des algorithmes non paramétriques, ce qui signifie qu'il ne fait aucune hypothèse sur la distribution des données sous-jacentes.

Concrètement, lorsque le KNN est confronté à un nouvel exemple, il calcule la distance entre cet exemple et tous les exemples du jeu de données d'entraînement, puis sélectionne les 'K' exemples les plus proches. La classe majoritaire parmi ces 'K' voisins est ensuite attribuée au nouvel exemple. L'algorithme est sensible au choix de la valeur de 'K', qui détermine le nombre de voisins à considérer pour la classification. [12]

3.3.2. Quelques règles sur le choix de k

Voici quelques règles concernant le choix de k dans l'algorithme KNN [14] :

- Le paramètre k doit être choisi par l'utilisateur, où k est un nombre entier positif. En classification binaire
- Si vous faites une classification binaire (deux classes), préférez choisir un k impair pour éviter les égalités
- Le choix optimal de k dépend du jeu de données spécifique. En règle générale, des valeurs élevées de k réduisent l'impact du bruit sur la classification et donc le risque de sur-apprentissage, mais peuvent également rendre les frontières entre les classes moins distinctes.

3.3.3. Les avantage

Le KNN présente plusieurs avantages principaux [15] :

- Sa simplicité le rend facile à comprendre et à mettre en œuvre.
- Son efficacité est souvent compétitive dans de nombreux domaines de classification.
- Son fonctionnement intuitif le rend accessible même pour les débutants.
- Il est robuste aux données d'entraînement bruitées, ce qui signifie qu'il peut bien fonctionner même en présence de données imprécises ou incomplètes.
- Il est robuste aux données d'entraînement contenant du bruit et performant lorsque les données d'entraînement sont de grande taille.

3.3.4. Les inconvénients

Les inconvénients d'algorithme KNN sont [15] :

- Le KNN peut être lent lorsque vous avez beaucoup de données d'entraînement.
- Il est difficile de déterminer quel type de distance ou quelle caractéristique utiliser pour obtenir les meilleurs résultats.
- Le calcul des distances entre chaque instance de requête et tous les échantillons d'entraînement peut être coûteux en termes de temps de calcul.

4.La logique floue

4.1. Définition

La logique floue, une branche de la théorie des ensembles initiée dans les années 1960 par Lotfi Zadeh, offre un cadre mathématique pour représenter et traiter l'incertitude et

l'imprécision dans les systèmes complexes. Contrairement aux modèles traditionnels basés sur des ensembles définis de manière précise, la logique floue permet de quantifier le degré d'appartenance d'un élément à un ensemble à l'aide de nombres réels dans l'intervalle [0,1]. Une valeur de 0 indique la non-appartenance, une valeur de 1 indique l'appartenance totale, et les valeurs entre 0 et 1 représentent une appartenance graduelle ou partielle.

Cette approche trouve une application particulièrement pertinente dans la modélisation des systèmes humains, où les caractéristiques sont souvent décrites en termes qualitatifs plutôt que quantitatifs. La logique floue se distingue également de la théorie de la probabilité en mettant l'accent sur la représentation de la qualité ou des caractéristiques des objets ou des événements, plutôt que sur la quantification du risque associé à leur occurrence. Ainsi, la logique floue constitue un outil essentiel pour traiter l'incertitude inhérente à de nombreux domaines, offrant une approche souple et adaptative à la modélisation des systèmes complexes. [16]

4.2. Intérêt de la logique floue

Les intérêts de la logique floue sont [17] :

- Elle suggère de modéliser le comportement humain face au système plutôt que de modéliser directement le système lui-même, ce qui correspond mieux à la réalité des interactions humaines avec les systèmes complexes.
- La logique floue reconnaît que nous ne comprenons pas toujours les situations de manière claire et précise.
- Elle nous aide à résoudre des problèmes même lorsque nous avons des informations approximatives plutôt que des données précises.
- Plus un système est complexe, moins il est possible de faire des affirmations précises sur son comportement, ce qui rend la logique floue encore plus utile.
- Elle utilise des termes généraux pour décrire le fonctionnement des variables, ce qui correspond mieux à la manière dont nous percevons et comprenons le monde.

4.3. La structure générale de la logique floue :

La logique floue comprend quatre éléments principaux :

- Interface de fuzzification.

- Base de connaissances.
- Moteur d'inférence floue.
- Interface de défuzzification.

4.3.1. L'interface de fuzzification

L'interface de fuzzification est la première étape dans la mise en œuvre d'un système de contrôle basé sur la logique floue. Pendant cette phase, les données provenant du système sont d'abord normalisées, puis transformées en termes linguistiques. Étant donné que les entrées dans un système flou sont généralement mesurées à l'aide de capteurs analogiques, il est nécessaire de les convertir en valeurs numériques avant de les soumettre à la fuzzification.

Cette adaptation des entrées garantit qu'elles appartiennent à l'univers de discours spécifié. Enfin, la fuzzification génère les degrés d'appartenance à chaque terme linguistique défini par sa fonction d'appartenance. [18]

4.3.2. La base de connaissances

L'interface de Base de Connaissances constitue une phase cruciale dans la conception des systèmes experts. Elle se compose de deux éléments essentiels : la base de données et la base des règles floues.

La base de données comprend la définition des ensembles flous, les facteurs d'échelle pour normaliser les ensembles de référence, ainsi que la partition de l'espace flou d'entrée et de sortie.

D'autre part, la base des règles floues réunit toutes les règles de type "Si-Alors" décrivant, en termes linguistiques basés sur l'expertise d'un spécialiste, le comportement dynamique du système. [19]

4.3.3. Le moteur d'inférence floue :

Le Moteur d'Inférence Floue est un composant crucial des systèmes basés sur la logique floue, utilisé pour prendre des décisions en exploitant le raisonnement approximatif à partir des faits observés dans la base des règles floues. Dans le processus d'inférence, les opérateurs logiques ET et OU sont utilisés, l'opérateur ET étant appliqué aux variables à l'intérieur d'une règle spécifique et l'opérateur OU liant différentes règles entre elles. Plusieurs méthodes sont

utilisées pour réaliser ces opérations d'inférence, telles que la méthode max-min (Mamdani), la méthode max-produit (Larsen), et la méthode somme-produit (Sugeno), qui déterminent les valeurs de sortie en fonction des valeurs d'entrée et des règles floues définies dans le système. [19]

4.3.4. L'interface de défuzzification

L'interface de défuzzification est une étape essentielle dans les systèmes de contrôle flou. Elle intervient après le processus d'inférence pour convertir les sorties floues en valeurs numériques significatives. Cette opération est nécessaire car les entrées du processus nécessitent des valeurs non floues pour être interprétées correctement. La défuzzification consiste à extraire une valeur numérique précise à partir de l'ensemble flou de sortie du régulateur.

Parmi les méthodes de défuzzification les plus couramment utilisées, on trouve :

- La méthode du centre de gravité (DEF1), qui détermine la valeur de sortie en calculant le centre de gravité de l'ensemble flou.
- La méthode du plus grand maximum (DEF2), qui renvoie la plus grande valeur avec le plus grand degré d'appartenance.
- La méthode du plus petit maximum (DEF3), qui renvoie la plus petite valeur avec le plus grand degré d'appartenance.
- La méthode de la moyenne des maxima (DEF4), qui calcule la moyenne des valeurs ayant le degré d'appartenance maximal.

Ces méthodes permettent de transformer les sorties floues en valeurs numériques précises, facilitant ainsi l'interprétation et l'utilisation des résultats dans les systèmes de contrôle flou. [19]

4.4. Les avantage de la logique floue

Les avantages de la logique floue sont [20] :

- Ils permettent une intégration directe des informations floues et linguistiques, fournies par des experts humains, dans le système flou.

- Il n'est pas nécessaire de créer un modèle mathématique précis du système à réguler, ce qui simplifie le processus de conception.
- Les systèmes flous sont des approximations universelles, ce qui signifie qu'ils sont assez généraux pour générer une gamme variée d'actions.
- La logique floue est accessible à la compréhension même pour ceux qui ne sont pas des spécialistes, car elle reflète la stratégie de raisonnement humain.

4.5. Les inconvénients de la logique floue

Les inconvénients de la logique floue sont [20] :

- Adaptation difficile aux changements des données
- Absence de théorie générale qui caractérise rigoureusement la stabilité, la robustesse.

5. La logique floue VS la logique classique

La logique classique et la logique floue sont deux approches distinctes pour représenter et traiter les connaissances et les données. La logique classique utilise une valeur binaire pour déterminer si une proposition est vraie ou fausse, tandis que la logique floue introduit des degrés d'appartenance pour représenter l'imprécision et l'incertitude. Par exemple, lorsque l'on considère la température d'une pièce, la logique classique divise les valeurs en deux catégories, telles que "moins de 15 degrés" et "15 degrés ou plus", sans tenir compte des nuances entre les deux.

En revanche, la logique floue permet de représenter des concepts tels que "plus ou moins froid" ou "plus ou moins chaud", reflétant ainsi la perception humaine qui ne se limite pas à des distinctions binaires précises. Cette approche flexible de la logique floue est particulièrement utile dans les situations réelles où les informations sont souvent imprécises ou incertaines, permettant ainsi une modélisation et une prise de décision plus adaptées à la complexité du monde réel. [21]

6.La logique floue dans la prédiction des maladies cardiaques

La littérature montre que les systèmes basés sur la logique floue sont fréquemment utilisés dans le diagnostic des maladies, où divers facteurs influencent le processus de prise de décision et conduisent à des divergences dans les opinions des praticiens. Il a déjà été

démontré l'application de la logique floue et son efficacité dans le diagnostic médical de la spondylarthrite ankylosante, de l'anémie, de la dengue, de la thyroïde, de la maladie d'Alzheimer, de la tension artérielle, du diabète et de la santé mentale, dans les systèmes de recommandation d'aliments et de médicaments assistés par une ontologie pour un patient, etc. pour identifier divers troubles [25, 26, 27, 28, 29, 30, 31,32]. De même, de nombreuses recherches sont en cours pour créer un système basé sur la logique floue, précis, économique et efficace pour le diagnostic des maladies cardiaques, car l'organisation mondiale de la santé (OMS) a signalé que les maladies cardiovasculaires (MCV) sont désormais la principale cause de décès dans le monde [25, 33,34,35,36]. Il est intéressant de noter que le risque de mortalité peut être réduit par un diagnostic précoce des maladies cardiaques. Cependant, des algorithmes adaptatifs sont nécessaires pour une prédiction plus précoce. Certains d'entre eux sont présentés dans [37, 38].

Gadekallu, Wiharto, Tan, et un bon groupe de chercheurs dans [39, 40, 41, 42, 43, 44,45, 46, 47,48] ont travaillé sur la prédiction des maladies cardiaques. Un autre groupe de chercheurs dans [49, 50, 51, 52, 53, 54, 55,56] se sont concentré sur le diagnostic des maladies cardiaques. Avec un nombre différent d'attributs d'entrée et de sortie, les auteurs de [57, 58, 59, 60, 61, 62, 63, 64,65] présentent différents niveaux de précision. La précision obtenue par Terrada et al est de 63,24 % [64] et Krishnaiah et son groupe ont pu obtenir une précision de 94,05 % [61] sur la base de diverses fonctionnalités d'entrée et de sortie implémentées dans les systèmes d'inférence Mamdani. L'algorithme proposé par Adeli et Neshat [57] affiche une précision de 94 % en utilisant 44 règles dans le système expert flou avec onze entrées et une sortie. La précision est un peu plus élevée (94,05 %) dans [61] à mesure que le travail introduit l'algorithme d'arbre de décision avec le système expert flou. Ces dernières années, certains travaux de recherche [36, 66, 67, 68,69] ont présenté une amélioration supplémentaire de la précision grâce à une approche hybride d'algorithmes flous et de réseaux neuronaux, mais l'hybridation rend le système global compliqué et gourmand en calculs.

D'autres solutions ont été proposées pour la prédiction des maladies cardiaques en utilisant différentes techniques. Anooj [70] a proposé un système d'aide à la décision clinique pour le diagnostic des maladies cardiaques, où des règles floues pondérées ont été utilisées pour construire le système en utilisant le système d'inférence floue Mamdani. Cette méthode a atteint une précision de 62,3 %.

Bashir et coll. [71] ont présenté une méthode de classification d'ensemble pour la prédiction des maladies cardiaques, où cinq classificateurs hétérogènes tels que Naive Bayes, arbre de décision basé sur Gini-index, arbre de décision basé sur le gain d'informations, un algorithme d'apprentissage basé sur la mémoire et une machine vectorielle ont été utilisée. Cette méthode a atteint une précision de 88,52 %. Pal et coll. [72] ont proposé un système expert flou pour le dépistage des maladies coronariennes (MAC), où les règles ont été formulés par les médecins et une approche de système expert flou a été adoptée pour faire face à l'incertitude présente dans le domaine médical.

Khatibi et Montazer [73] ont proposé un moteur d'inférence hybride à preuves floues pour l'évaluation du risque de maladie coronarienne. Ils ont atteint un taux de précision de 91,58 %.

Conclusion

L'apprentissage automatique est un terme populaire de nos jours en raison des progrès technologiques qu'il a apportés dans tous les domaines. C'est pourquoi nous avons commencé dans ce chapitre par définir ce qu'est l'apprentissage automatique et ses différents types et algorithmes. Ensuite, nous avons exploré la logique floue.

Dans le prochain chapitre, nous nous concentrerons sur la conception et la mise en œuvre d'un système permettant de prédiction les maladies cardiaques de manière efficace.

Chapitre 03 : Conception et réalisation

Introduction

Dans ce chapitre, nous commençons par présenter l'environnement de développement ainsi que les outils utilisés dans le cadre de ce projet. Nous décrivons ensuite les différentes étapes typiquement impliquées dans la création d'un système de prédiction des maladies cardiaques.

Enfin nous présentons et comparons les résultats obtenus pour chaque modèle (la logique floue, KNN et la logique floue combinée eu KNN).

1. L'environnement de développement

1.1. La plateforme de développement

a. Google colab :

Google Colab ou « Colaboratory » est un produit de recherche de Google. Google colab permet à quiconque d'écrire et d'exécuter du code Python via un navigateur et est particulièrement adapté à l'apprentissage automatique, à l'analyse de données et à l'apprentissage. Cela évite à la personne de devoir installer les logiciels nécessaires sur son ordinateur. Il est notamment possible de travailler avec d'autres services Google comme Google Drive. [22]

b. PyCharm :

PyCharm est un environnement de développement intégré (IDE) Python (développé par JetBrains) qui fournit aux développeurs Python une large gamme d'outils essentiels étroitement intégrés pour créer un environnement pratique pour un développement Python, Web et scientifique efficace des données. Il s'agit d'un logiciel multiplateforme qui fonctionne sous Windows, Mac OS X et GNU/Linux. Il a une version professionnelle et une version communautaire, dans notre projet, nous avons utilisé la version communautaire 2023.2.3[W12]

Figure 9 : Le logo du PyCharm **[W13]**

1.2. Le langage de programmation

a. Python :

Python est un langage de programmation puissant et facile à apprendre. Il contient des structures de données de haut niveau et permet une programmation orientée objet en utilisant une approche simple mais efficace. En raison de sa syntaxe élégante et de la nature dynamique de son écriture et de son interprétation, Python est un langage idéal pour la création de scripts et le développement rapide d'applications dans de nombreux domaines et sur la plupart des plateformes. Il est très populaire, notamment dans les domaines de la science des données et de l'apprentissage automatique.

Python se démarque en proposant un grand nombre de bibliothèques de très haute qualité qui facilitent l'apprentissage et améliorent ses fonctionnalités. [W14]

Figure 10 : Le logo du Python **[W15]**

1.3 Les bibliothèques de Python

a. NumPy

NumPy, signifiant « Python Numérique » ou « Numerical Python », est une bibliothèque largement utilisée en Python conçue pour les calculs numériques. Il a été développé en 2005 par Travis Oliphant. NumPy est un projet open source qui peut être utilisé librement pour manipuler des tableaux multidimensionnels, ce qui est essentiel pour diverses applications scientifiques et techniques. Il comprend également des fonctions puissantes pour l'algèbre linéaire, les transformées de Fourier et les opérations matricielles. [W16]

b. Pandas

Pandas est une bibliothèque de langage de programmations Python open source conçue pour la manipulation et l'analyse de données. Elle offre spécifiquement des structures de données ainsi que des fonctionnalités pour manipuler des tableaux numériques et des séries temporelles. [23]

c. Scikit-learn

Scikit-learn est la bibliothèque la plus utile et la plus puissante pour l'apprentissage automatique en Python. Elle fournit une gamme d'outils efficaces pour l'apprentissage automatique et la modélisation statistique, y compris la classification, la régression et le clustering, via une interface cohérente en Python. [24]

D. Matplotlib

Matplotlib est une bibliothèque de langage de programmation python open source destinée à tracer et visualiser des données sous formes de graphiques. [W17]

2.Le système proposée

Nous présentons à présent l'architecture du système créé, et cela à travers les différentes étapes impliquées dans la construction du système de prédiction de maladies cardiaques.

Cette architecture comprend la collecte de données, l'analyse exploratoire de ces dernières, leur prétraitement, et leur répartition en un ensemble de test et un ensemble d'entrainement. Ensuite le système passe à l'entrainement des modèles (logique floue, KNN et modèle hybride), et fini par évaluer cet entrainement.

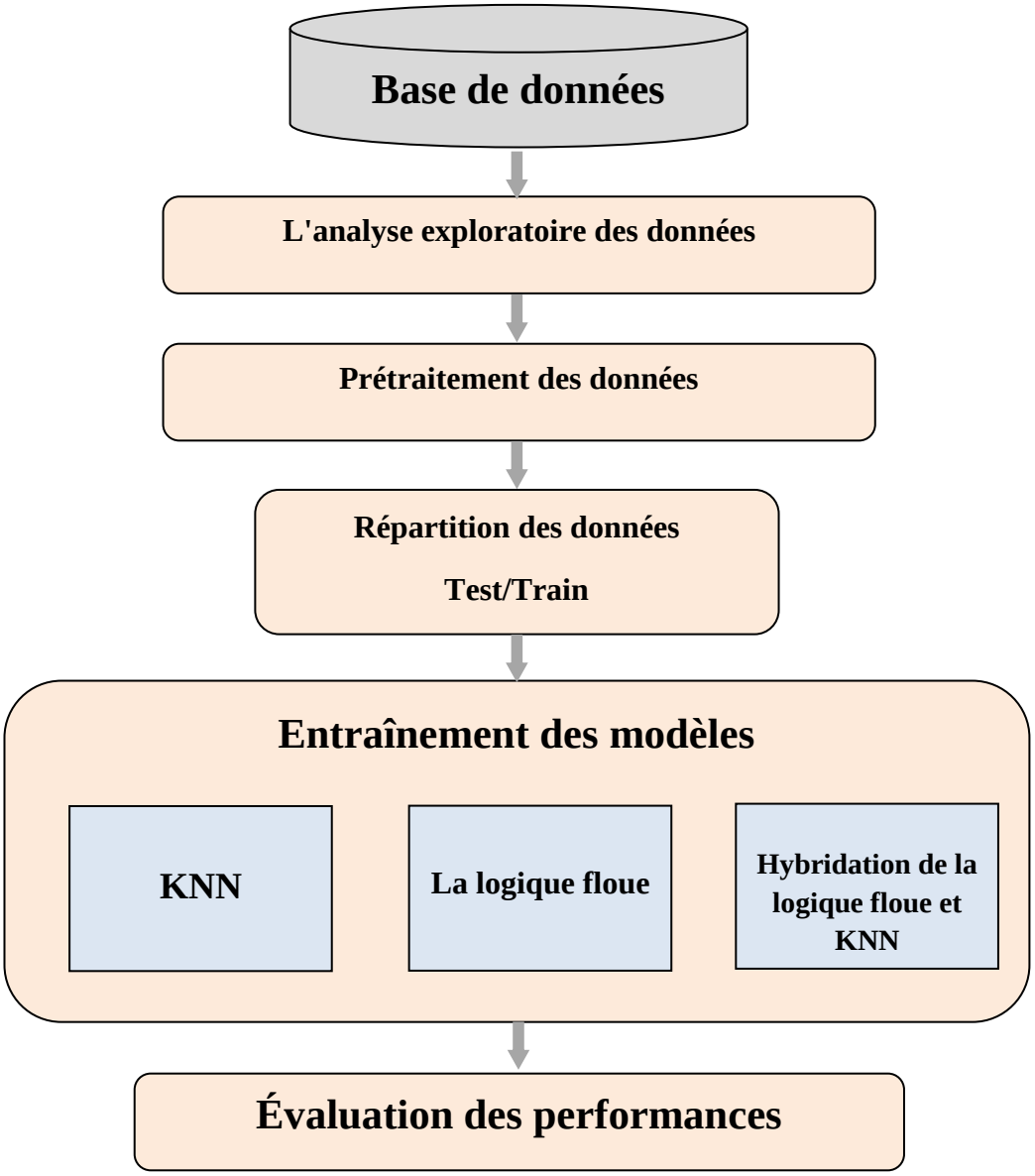

Figure 11 : L'architecture du système proposé

2.1. La base de données

Pour cette étude, nous avons fait appel à une des bases de données les plus pertinentes portant sur les maladies cardiaques. Chaque instance de ladite base de données comprends 14 attributs cardiaques. Cet ensemble appelé " Heart " et est disponible sur le site web Kaggle.

2.1.1. Description de la base de données

L'ensemble de données utilisé est un ensemble englobant plusieurs cas correspondant à des diagnostiques faits portants sur une suspicion de maladie cardiaque. Cet ensemble a été acquis par le biais du site web Kaggle.

Il est composé du fichier "heart.csv" contenant le détail des attributs et des variables. Nous avons utilisé les 14 attributs, y compris la caractéristique "Target" des 356 patients. L'attribut "Target" indique la présence ou l'absence de maladies cardiaques chez les patients. Il est représenté par une valeur entière dans l'ensemble {0,1}, ou 0 signifie l'absence de maladies cardiaques et 1 indique leur présence.

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
58	1	0	100	234	0	1	156	0	0.1	2	1	3	0
47	1	0	110	275	0	0	118	1	1	1	1	2	0
52	1	0	125	212	0	1	168	0	1	2	2	3	0
58	1	0	146	218	0	1	105	0	2	1	1	3	0
57	1	1	124	261	0	1	141	0	0.3	2	0	3	0
58	0	1	136	319	1	0	152	0	0	2	2	2	0
61	1	0	138	166	0	0	125	1	3.6	1	1	2	0
42	1	0	136	315	0	1	125	1	1.8	1	0	1	0

Figure 12 : L’extrait de la base de données

2.1.2. Description des attributs de la base de données

L'ensemble de données comprend 356 données individuelles. Il y a 14 colonnes dans l'ensemble de données, qui sont décrites ci-dessous.

- **Âge :** affiche l'âge de l'individu (29 à 68 ans).
- **Sexe :** affiche le sexe de l'individu selon le format suivant :

  1 = homme

0 = femme

- **Type de douleur thoracique (Cp) :** affiche le type de douleur thoracique ressentie par l'individu en utilisant le format suivant :

  0 = angine typique

  1= angine atypique

  2= douleur non angineuse

  3 = asymptotique
- **Pression artérielle au repos (trestbps) :** affiche la valeur de la pression artérielle au repos d'un individu en mmHg (unité) 94 à 152.
- **Cholestérol sérique (Chol) :** indique le taux de cholestérol sérique, mesuré en mg/dl (166 à 564 mg/dl).
- **Glycémie à jeun (fbs) :** compare la valeur de glycémie à jeun d'un individu à 120 mg/dl.
- Si glycémie à jeun > 120 mg/dl alors : 1 (vrai) ; Sinon : 0 (faux)
- **ECG de repos(restecg) :** affiche les résultats électrocardiographiques au repos (0 à 2) :

  0=normal

  1=anomalie

  2=hypertrophie
- **Fréquence cardiaque maximale atteinte (thalach):** affiche la fréquence cardiaque maximale atteinte par un individu (105 à 202 battements par minute).
- **Angor induit par l'exercice (exang)** : indiquant la présence ou l'absence d'angine de poitrine induite par l'exercice (1 = oui ;0 = non).
- **Dépression ST induite par l'exercice par rapport au repos (oldpeak) :** ST se rapporte aux positions sur le graphique ECG.
- **Segment ST d'exercice de pointe (slope) :** indiquant la pente du segment ST lors de l'exercice.
- **Nombre de vaisseaux principaux colorés par la fluoroscopie (ca) :** affiche le nombre de grands vaisseaux (0 à 3).
- **Thalassémie (Thal) :** affiche un trouble sanguin appelé thalassémie (1 à 3).

  1=Normale

  2=Défaut fixe

  3=Défaut réversible

  Diagnostic de maladie cardiaque (Target) : Indique si l'individu souffre ou non d'une maladie cardiaque : (0 = l'absence de maladies ;1= présent de maladies).

2.2. L'analyse exploratoire des données

L'analyse d'exploratoire des données est définie comme l'exploration visuelle des données, afin d'obtenir et de comprendre des informations et des caractéristiques approfondies et claires concernant toutes les données et les variables.

- **La taille :** Il se compose de 356 lignes et 14 colonnes y compris le target.
- **Les colonnes :** Les colonnes de l'ensemble de données sont : ['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach','exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'].
- **Le type :** L'ensemble de données et de type : Object.
- **Cellules manquantes :** 0 cellule.
- **La taille :** La taille de l'ensemble de données 39.1KB.

```
Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach',
       'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'],
      dtype='object')
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 356 entries, 0 to 355
Data columns (total 14 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   age       356 non-null    int64
 1   sex       356 non-null    int64
 2   cp        356 non-null    int64
 3   trestbps  356 non-null    int64
 4   chol      356 non-null    int64
 5   fbs       356 non-null    int64
 6   restecg   356 non-null    int64
 7   thalach   356 non-null    int64
 8   exang     356 non-null    int64
 9   oldpeak   356 non-null    float64
 10  slope     356 non-null    int64
 11  ca        356 non-null    int64
 12  thal      356 non-null    int64
 13  target    356 non-null    int64
dtypes: float64(1), int64(13)
memory usage: 39.1 KB
```

Figure 13 : Les Informations de la base des données

a. Visualisation de la classe cible : la figure 14 montre un graphe qui représente la cible, pour la classe 0, correspondant aux individus non malades, représente 73.03 % du total. La classe 1, correspondant aux individus malades, représente 26.97 %.

Nous pouvons émettre qu'avec de telles proportions, si un algorithme, entraine grâce a cet ensemble, arrive à classer correctement (taux de précision élevé) des patients malades, ledit algorithme permettra de classer avec une meilleure certitude s'il est alimenté par plus de données par le praticien de sante.

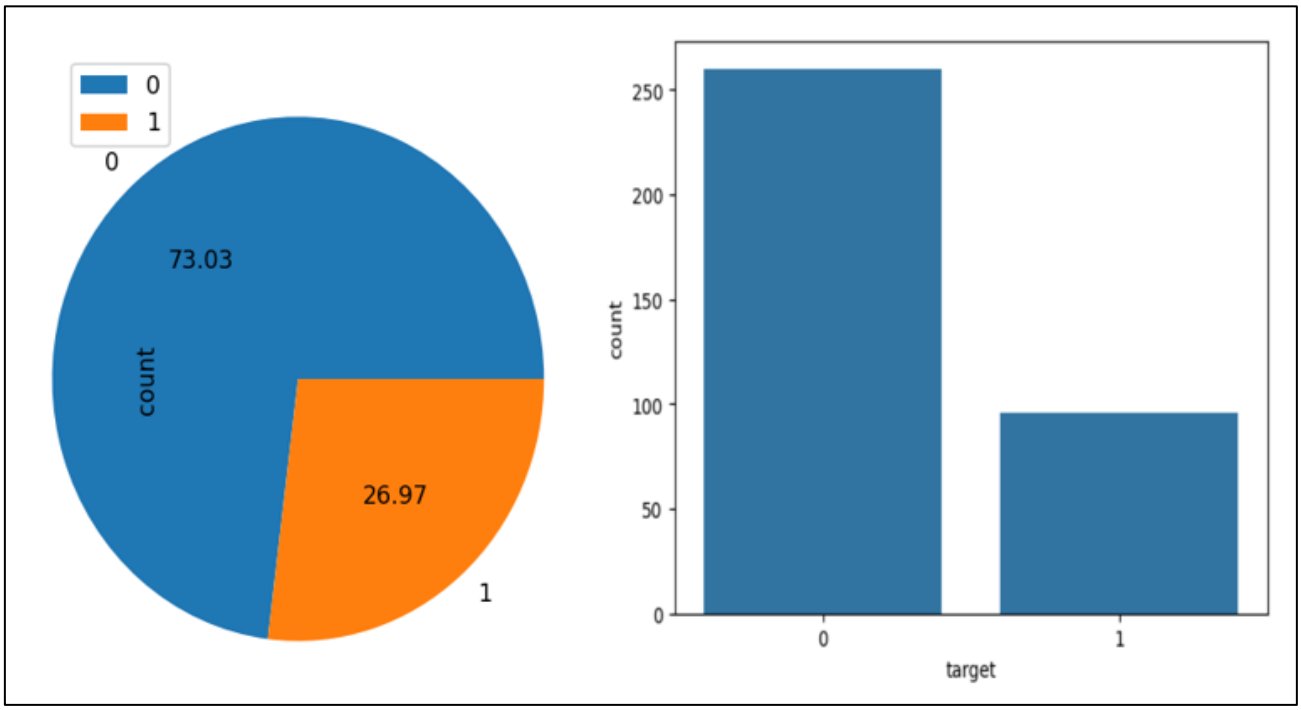

Figure 14 : La visualisation de la classe cible

2.3. Prétraitement des données

L'étapes de prétraitement des données est en effet cruciale dans le domaine d'apprentissage automatique (ML). Son objectif principal est de convertir des ensembles de données brutes dans un format adapté à l'application d'algorithmes d'apprentissage automatique.

Les données brutes peuvent contenir des valeurs manquantes, ce qui peut entraîner des problèmes lors de la formation des modèles d'apprentissage automatique. Pour notre base de données (Heart.csv), nous avons constaté qu'elle ne contient aucune valeur manquante, comme mentionné dans la figure ci-dessous.

```
age         0
sex         0
cp          0
trestbps    0
chol        0
fbs         0
restecg     0
thalach     0
exang       0
oldpeak     0
slope       0
ca          0
thal        0
target      0
dtype: int64
```

Figure 15 : Les valeurs manquantes

2.4. Répartition des données

La répartition ou la division des données est une étape qui consiste à séparer l'ensemble des données en ensembles d'entraînement et de test. Pour ce projet, nous avons choisi d'allouer 80 % des données à l'entraînement et 20 % pour le test.

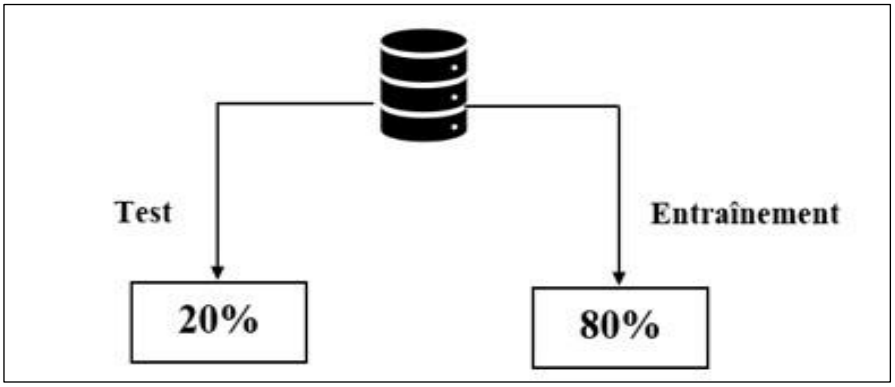

Figure 16 : La répartition des données de train/test

2.5. Modèle proposé

Le modèle proposé pour la prédiction des maladies cardiaques implique l'utilisation de divers algorithmes d'exploration de données. Dans ce projet, trois modèles ont été utilisés : la logique floue, le KNN et l'hybridation de la logique floue et KNN.

2.5.1. La logique floue

Pour la prédiction des maladies cardiaques, on utilise le modèle de logique floue qui nous a permis, après tests, d'obtenir une bonne précision. La logique floue comprend trois étapes principales : la fuzzification, l'inférence floue et la défuzzification.

a. La fuzzification

La fuzzification est la première étape du processus de raisonnement flou, où les valeurs d'entrée réelles sont converties en degrés d'appartenance à des ensembles flous. Par exemple, l'âge d'une personne peut être traduit en termes flous tels que "jeune", "moyen", "vieux", "très vieux". Ces catégories sont définies par des règles qui attribuent un degré d'appartenance à chaque catégorie en fonction de l'âge. Cette méthode de conversion est également appliquée à d'autres attributs pour permettre le traitement par des systèmes basés sur la logique floue.

Pour cela, les fonctions d'appartenance (membership) des ensembles requis sont représentées dans les figures suivantes :

- **Fuzzification de l 'âge**

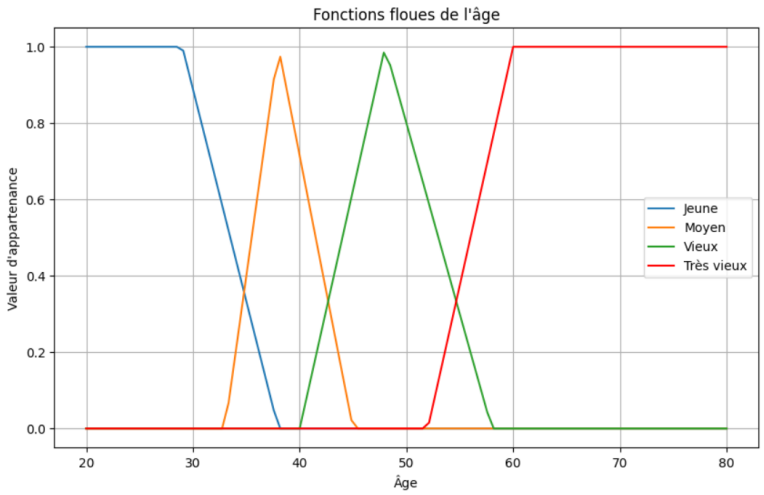

Figure 17 : La fuzzification de l 'âge

Le graphique illustre les fonctions d'appartenance floues pour différentes catégories d'âge, divisées en quatre groupes : "jeune", "moyen", "vieux", et "très vieux". Chaque catégorie est représentée par une courbe distincte indiquant le degré d'appartenance à un groupe d'âge spécifique en fonction de l'âge d'une personne.

- **Fuzzification de sexe**

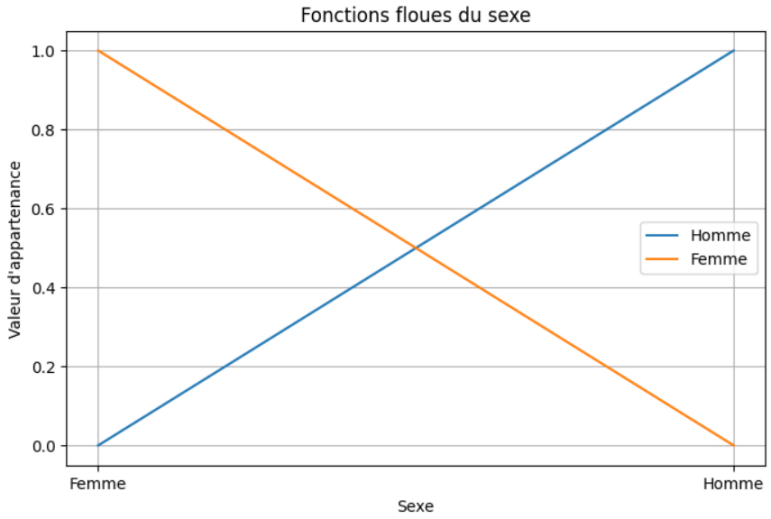

Figure 18 : La fuzzification de sexe

Cette image représente les fonctions d'appartenance pour la variable "sexe". Il y a deux ensembles flous : "Homme" et "Femme". Une personne de sexe masculin aura un degré d'appartenance de 1 pour l'ensemble "Homme" et de 0 pour "Femme".

- **Fuzzification de fonctions de douleur thoracique floues (cp)**

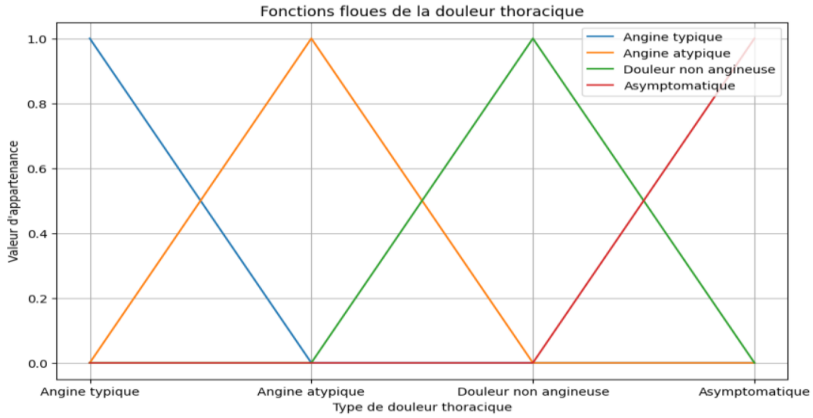

Figure 19 : La fuzzification de cp

Cette image montre les fonctions d'appartenance pour la variable "douleur thoracique". Il y a quatre ensembles flous : "angine typique", "angine atypique", "douleur non angineuse" et "asymptomatique".

- **Fuzzification de pression artérielle au repos (trestbps)**

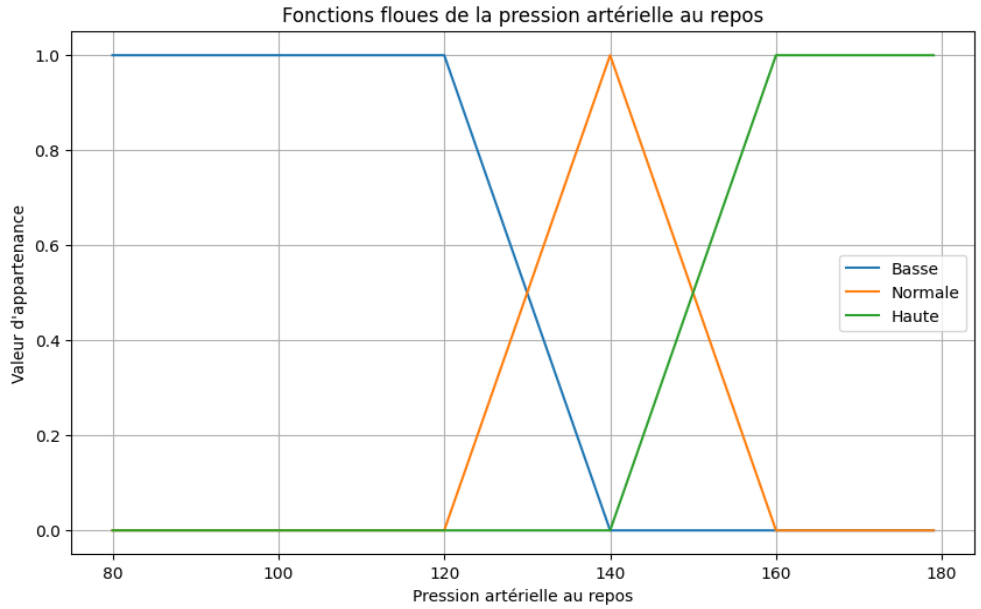

Figure 20 : La fuzzification de trestbps

Cette image représente les fonctions d'appartenance pour la variable "tension artérielle au repos". Il y a trois ensembles flous : "basse", "normale" et "haute".

- **Fuzzification de cholestérol sérique (Chol)**

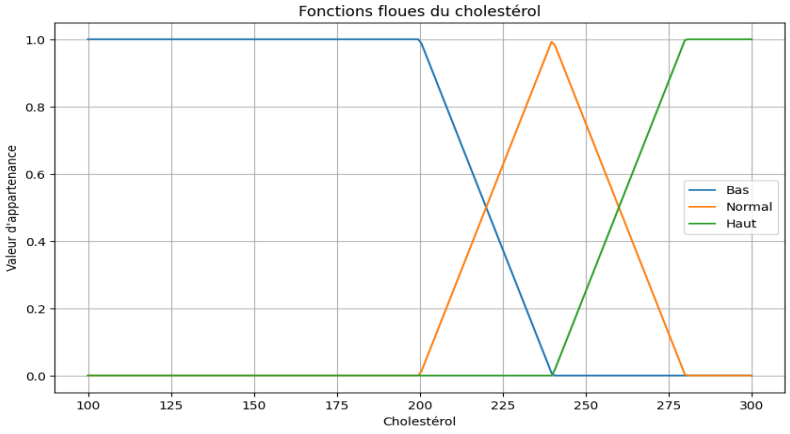

Figure 21 : La fuzzification de chol

Cette image montre les fonctions d'appartenance pour la variable "cholestérol". Il y a trois ensembles flous : "bas", "normal" et "élevé".

- **Fuzzification de glycémie à jeun (fbs)**

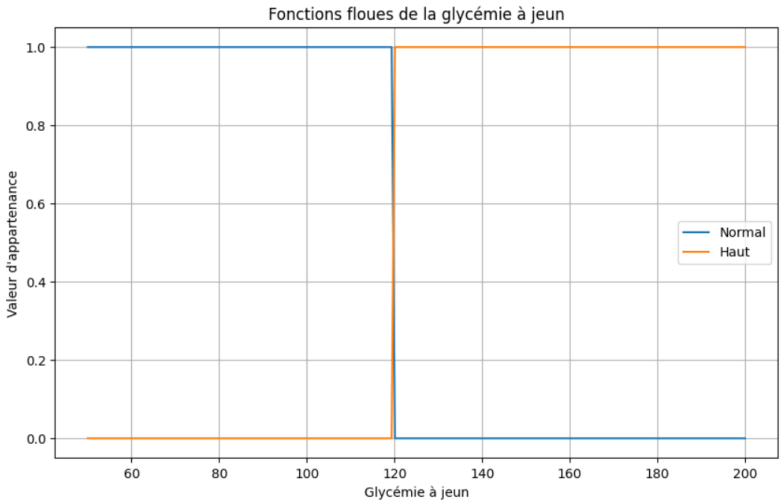

Figure 22 : La fuzzification de fbs

Cette image montre les fonctions d'appartenance floues pour les niveaux de glycémie à jeun, avec deux ensembles flous "Normal" et "Élevé".

- **Fuzzification de ECG de repos(restecg)**

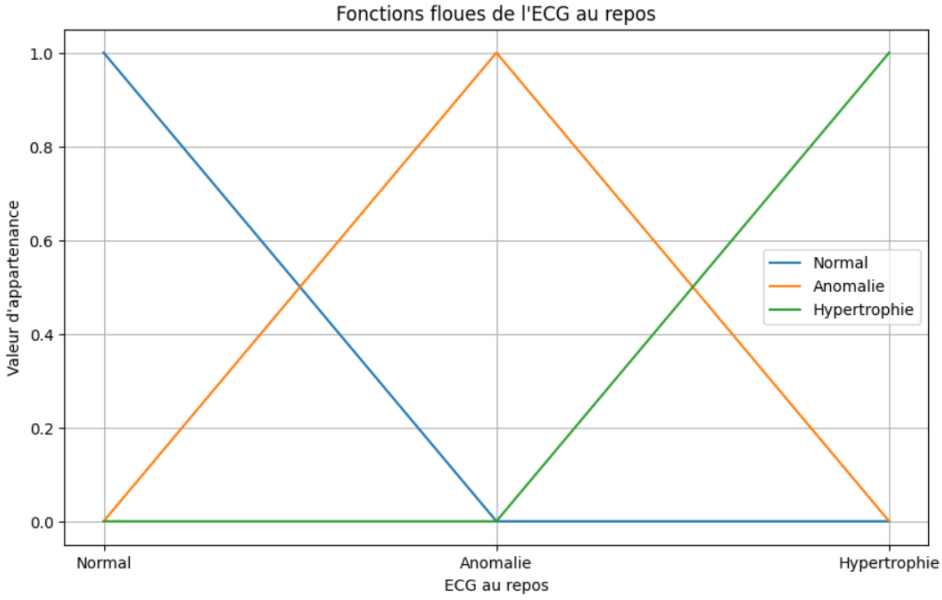

Figure 23 : La fuzzification de restecg

Cette image représente des fonctions d'appartenance floues pour les résultats d'un ECG au repos (électrocardiogramme), avec trois ensembles flous "Normal", "Anomalie" et "Hypertrophie".

- **Fuzzification de fréquence cardiaque maximale atteinte (thalach)**

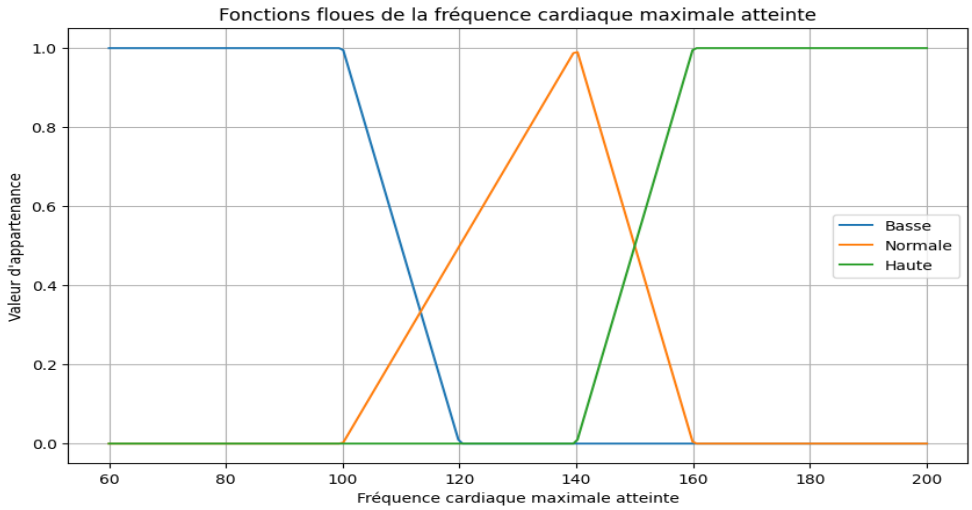

Figure 24 : La fuzzification de thalach

Cette image montre les fonctions d'appartenance floues pour la fréquence cardiaque maximale atteinte, avec trois ensembles flous "Basse", "Normale" et "Élevée".

- **Fuzzification d'angor induit par l'exercice (exang)**

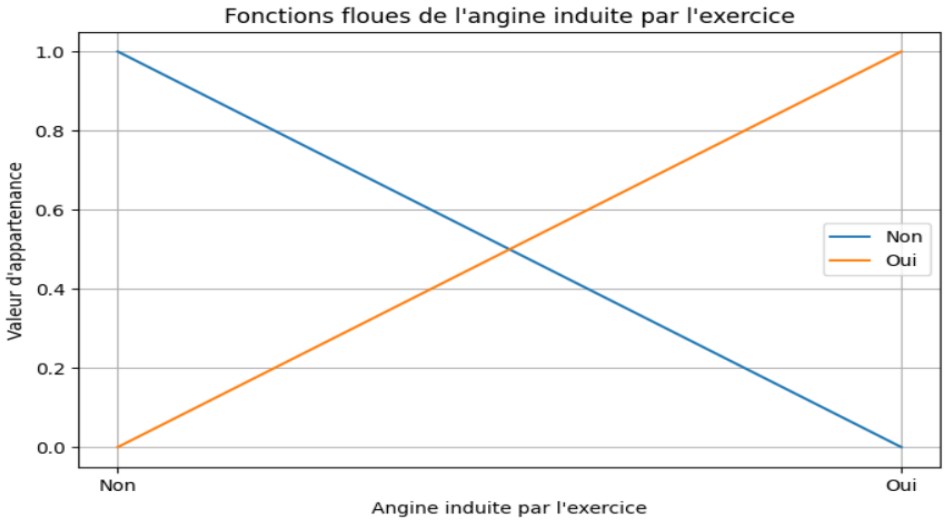

Figure 25 : La fuzzification de exang

Cette image montre des fonctions d'appartenance floues pour l'angine de poitrine induite par l'exercice, avec deux ensembles flous "Non" et "Oui".

- **Fuzzification de dépression ST induite par l'exercice par rapport au repos (oldpeak)**

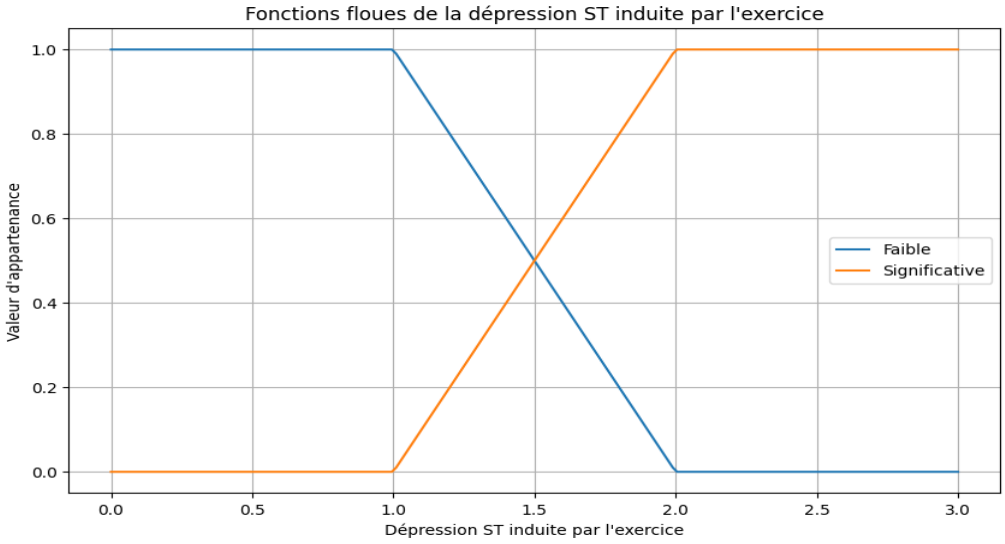

Figure 26 : La fuzzification de oldpeak

Cette image montre les fonctions d'appartenance floues pour la dépression du segment ST induite par l'exercice, avec deux ensembles flous "Faible" et "Significative".

- **Fuzzification de segment ST d'exercice de pointe (slope)**

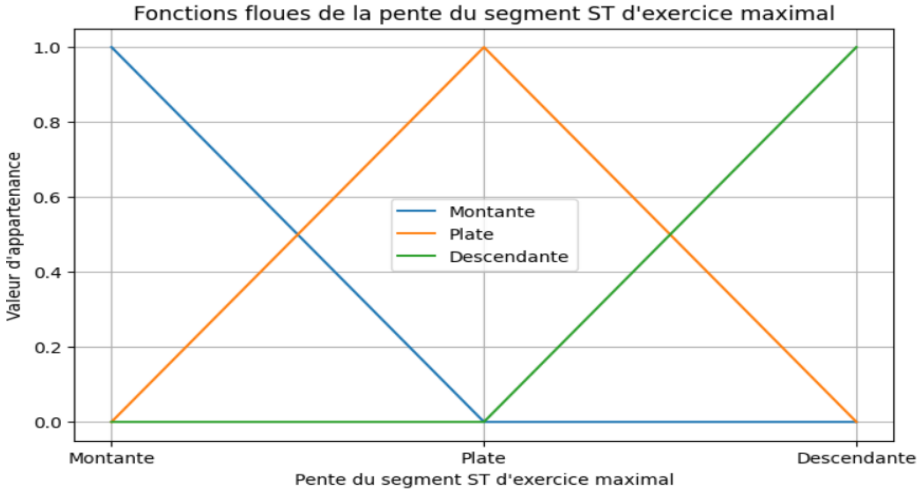

Figure 27 : La fuzzification de slope

Cette image montre les fonctions d'appartenance floues pour classer Segment ST d'exercice de pointe (slope) en trois catégories : Ascendante, Plate, et Descendante. Les valeurs d'appartenance floue varient de 0 à 1, indiquant le degré d'appartenance d'une valeur de slope donnée à chaque catégorie.

- **Fuzzification de nombre de vaisseaux principaux colorés par la fluoroscopie (ca)**

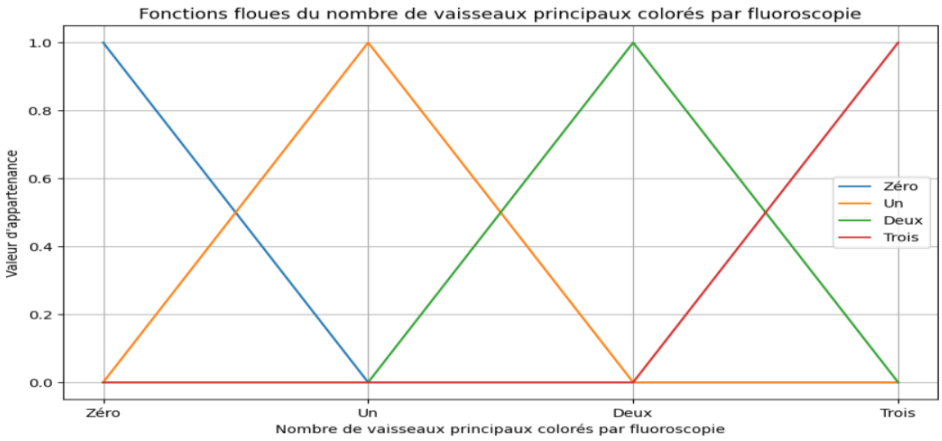

Figure 28 : La fuzzification de ca

Cette image représente les fonctions d'appartenance floues pour classer le nombre de vaisseaux principaux colorés par fluoroscopie en quatre catégories : Zéro, Un, Deux, et Trois ou plus. Encore une fois, les valeurs d'appartenance représentent le degré d'appartenance à chaque catégorie en fonction du nombre observé de vaisseaux.

- **Fuzzification de thalassémie (Thal)**

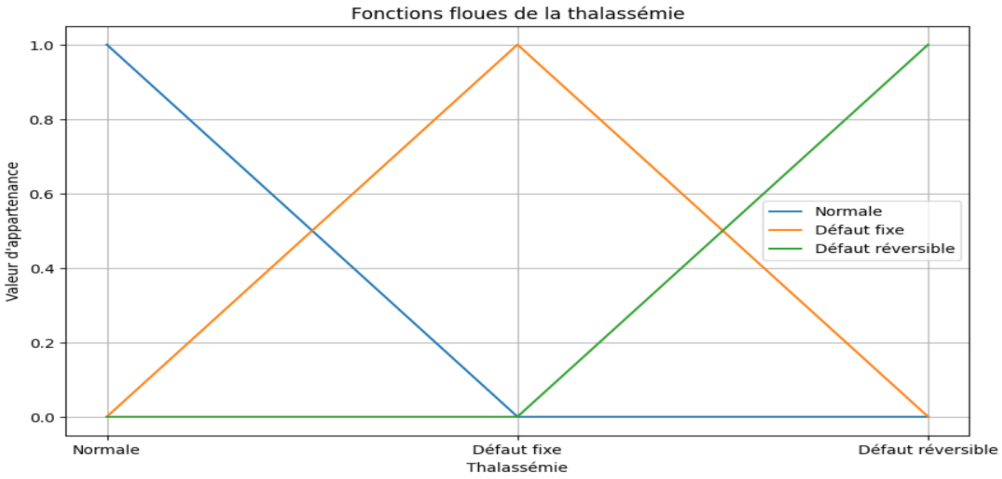

Figure 29 : La fuzzification de thal

Cette image illustre les fonctions d'appartenance floues pour classer la thalassémie (un trouble sanguin) en trois catégories : Normal, Défaut Fixe, et Défaut Réversible. Les valeurs d'appartenance indiquent le degré d'appartenance de l'état du patient à chaque catégorie.

- **Fuzzification de target**

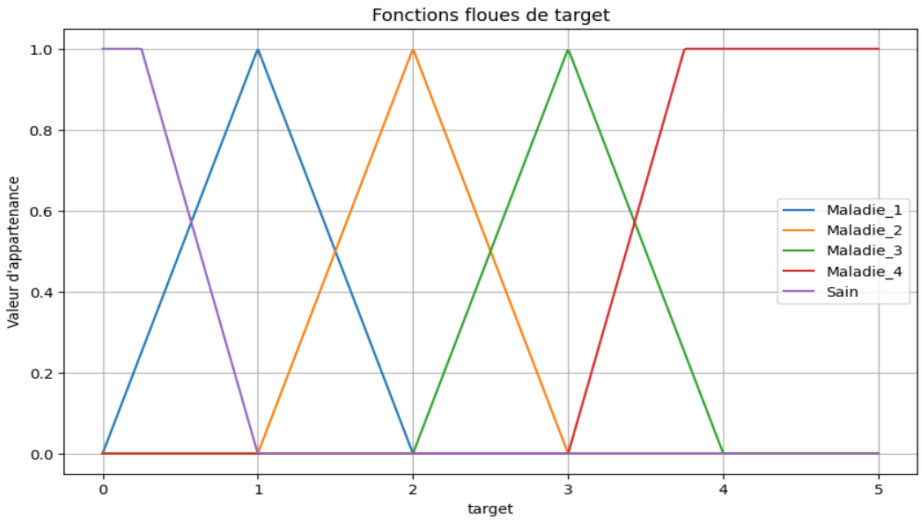

Figure 30 : La fuzzification de target

Cette image montre les fonctions de sortie floues, qui représentent le degré d'appartenance pour différentes catégories de sortie, telles que Maladie_1, Maladie_2, Maladie_3, Maladie_4, et Sain. Ces fonctions sont utilisées pour déterminer la classification de sortie finale en fonction des valeurs d'appartenance floue en entrée et des règles floues définies.

b. Inférence floue

Après une étude de modèles de diagnostique suivis par les praticiens cardiologue, et après avoir établis les différents catégories et facteur définissants les diverses possibilités de maladie, et enfin après avoir étudié les relations entre les facteurs permettant de définir un patient malade d'un patient sain, nous avons créé les règles floues suivants:

- Si (age est très vieux) ET (cp est angine atypique) ET (le sexe est homme), alors la santé est malade_4.
- Si (thalach est élevé) ET (l'âge est vieux) ET (trestbps est élevée), alors la santé est malade_4.

- Si (le sexe est femme) ET (thalach est normal) ET (fbs est élevé), alors la santé est malade_2.
- Si (cp est non angineuse) ET (trestbps est élevée) ET (l'âge est moyen), alors la santé est malade_3.
- Si (cp est angine typique) ET (thalach est normal) ET (le chol est normal), alors la santé est malade_2.
- Si (fbs est élevé) ET (l'âge est moyen) ET (le sexe est femme), alors la santé est malade_3.
- Si (fbs est normal) ET (trestbps est élevée) ET (cp est angine atypique), alors la santé est malade_2.
- Si (cp est asymptomatique) ET (l'âge est très vieux) ET (le sexe est homme), alors la santé est malade_1.
- Si (trestbps est élevée) ET (thalach est bas) ET (l'âge est vieux), alors la santé est malade_1.
- Si (trestbps est normale) ET (chol est normal) ET (fbs est normal), alors santé est saine.
- Si (restecg est normal) ET (thal est normal) ET (le sexe est femme), alors santé est saine.
- Si (l'âge est jeune) ET (exang est absente) ET (oldpeak est petit), alors santé est saine.
- Si (thalach est élevé) ET (slope est ascendante) ET (ca est zéro), alors santé est saine.
- Si (cp est angine typique) ET (le sexe est homme) ET (l'âge est jeune), alors santé est saine.
- Si (ca est trois ou plus) ET (l'âge est vieux) ET (le sexe est homme), alors la santé est malade_4.
- Si (thal est défaut réversible) ET (restecg est hypertrophie) ET (trestbps est élevée), alors la santé est malade_4.
- Si (chol est élevé) ET (fbs est élevé) ET (thal est normal), alors la santé est la malade_3.
- Si (restecg est normal) ET (thal est normal) ET (thalach est élevé), alors la santé est saine.
- Si (restecg est normal) ET (thal est normal) ET (thalach est élevé), alors la santé est saine.
- Si (le sexe est femme) ET (l'âge est jeune) ET (cp est non anginale), alors la santé est saine.
- Si (trestbps est normale) ET (restecg est normal) ET (exang est non), alors la santé est saine.

- Si (chol est élevé) ET (fbs est élevé) ET (thal est défaut réversible), alors la santé est malade_4.
- Si (oldpeak est significatif) ET (slope est descendante) ET (ca est deux), alors la santé est malade_3.
- Si (thalach est bas) ET (l'âge est très vieux) ET (le sexe est femme), alors la santé est malade_2.
- Si (le sexe est homme) ET (trestbps est élevée) ET (cp est asymptomatique), alors la santé est malade_4.
- Si (restecg est anormalité) ET (thal est défaut fixe) ET (trestbps est basse), alors la santé est saine.
- Si (fbs est normal) ET (slope est ascendante) ET (thal est normal), alors la santé est saine.
- Si (ca est un) ET (thal est un défaut fixe) ET (fbs est élevé), alors la santé est malade_3.
- Si (slope est petit) ET (l'âge est jeune) ET (thalach est normal), alors la santé est malade_2.
- Si (trestbps est normale) ET (exang est oui) ET (l'âge est vieux), alors la santé est en bonne santé.
- Si (chol est élevé) ET (le sexe est homme) ET (oldpeak est significatif), alors la santé est malade_4.
- Si (thal est défaut réversible) ET (cp est angine typique) ET (restecg est hypertrophie), alors la santé est malade_2.
- Si (fbs est normal) ET (thal est normal) ET (trestbps est basse), alors la santé est saine.
- Si (exang est oui) ET (oldpeak est significatif) ET (cp est asymptomatique), alors la santé est malade_1.
- Si (chol est élevé) ET (trestbps est élevée) ET (l'âge est très vieux), alors la santé est malade_2.
- Si (thalach est bas) ET (slope est descendante) ET (fbs est élevé), alors la santé est malade_3.
- Si (ca est trois) ET (thal est un défaut réversible) ET (oldpeak est significatif), alors la santé est malade_4.
- Si (restecg est normal) ET (fbs est normal) ET (chol est normal), alors la santé est saine.
- Si (cp est asymptomatique) ET (le sexe est homme) ET (trestbps est élevée), alors la santé est malade_1.

- Si (thalach est normal) ET (l'âge est moyen) ET (chol est élevé), la santé est malade_2.
- Si (restecg est normal) ET (l'âge est jeune) ET (thalach est élevé), alors la santé est saine.
- Si (cp est asymptomatique) ET (exang est oui) ET (oldpeak est significatif), alors la santé est malade_1.
- Si (thalach est bas) ET (slope est descendante) ET (ca est deux), alors la santé est malade_2.
- Si (fbs est élevé) ET (trestbps est élevée) ET (le sexe est homme), alors la santé est malade_3.
- Si (l'âge est très vieux) ET (thal est un défaut réversible) ET (chol est élevé), alors santé est malade_4.
- Si (cp est non angineuse) ET (exang est oui) ET (thalach est bas), alors la santé est malade_1.
- Si (chol est élevé) ET (trestbps est élevée) OU (le sexe est homme), alors la santé est malade_2.
- Si (fbs est élevé) ET (l'âge est vieux) ET (ca est deux), alors la santé est malade_3.
- Si (thal est un défaut réversible) ET (restecg est une hypertrophie) ET (oldpeak est significatif), alors la santé est malade_4.
- Si (l'âge est très vieux) ET (slope est descendante) ET (ca est trois ou plus), alors santé est malade_1.
- Si (le sexe est féminin) ET (cp est angine atypique) ET (restecg est une anomalie), alors la santé est malade_2.
- Si (thalach est normal) ET (chol est normal) ET (trestbps est élevée), alors la santé est malade_3.
- Si (cp est typique de l'angine de poitrine) ET (thal est un défaut fixe) ET (fbs est élevé), alors la santé est malade_4.
- Si (restecg est normal) ET (thalach est élevé) ET (slope est ascendante), alors santé est saine
- Si (chol est normal) ET (l'âge est jeune) ET (fbs est normal), alors la santé est saine.

c. La défuzzification

La défuzzification c'est la dernière étape qui convertit les résultats flous de l'inférence en une valeur numérique réelle qui représente le risque global de maladie cardiaque. La méthode de défuzzification que nous avons utilisée pour le travail proposé est la méthode du centroïde

(centre de gravité) qui identifie le point d'équilibre de la distribution floue résultante pour déterminer le risque final.

2.5.2. K plus proches voisins (KNN)

L'algorithme des k-plus proches voisins (KNN) est un algorithme de classification qui repose sur le principe de recherche des k voisins les plus proches. Dans le contexte de la prédiction des maladies cardiaques, les caractéristiques du patient telles que l'âge, le sexe, cp et d'autres attributs sont utilisées pour déterminer les voisins les plus proches. Une fois les voisins les plus proches identifiés, le modèle KNN attribue une étiquette au point d'intérêt en fonction de la classe majoritaire parmi ses voisins. Par exemple, si une majorité de k voisins sont étiquetés « malades », le modèle prédit que le patient est malade.

L'algorithme KNN se base sur deux paramètres principaux, qui sont la distance et le nombre de voisins. Nous avons choisi, en fonction de la nature du problème traite, de faire appel a la distance Euclidienne.

La distance euclidienne mesure la distance directe entre deux points en ligne droite. La distance euclidienne d est donnée par la formule (Où x et y sont les coordonnées des deux points) :

$$\boldsymbol{d(x,y) = \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2}}$$

Nous avons fait appel à cette distance car son concept correspond particulièrement à notre problème, ou la distance entre deux points sera la somme des différences entre deux valeurs de chaque attribut. Cette somme devra tendre vers la valeur nulle afin que les points soient identiques. Plus cette distance est petite plus les points sont rapprochés, que ce soit graphiquement ou au sens propre.

Pour définir le nombre de voisins "K" adéquat nous avons procédé par test, nous avons évalué la précision offerte par le KNN en faisant évoluer le K jusqu'à ce que la précision stagne.

La figure ci-dessous présente les résultats de cette évolution, qui désigne k=1 et k=3 comme meilleure valeur à choisir.

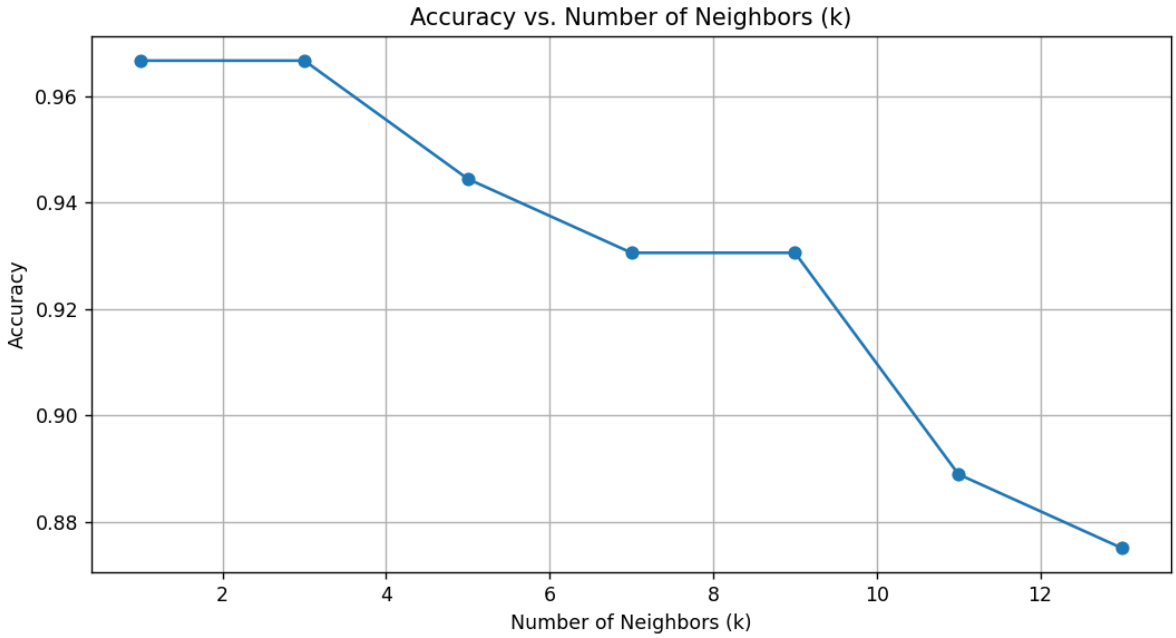

Figure 31 : Le nombre optimal de k-voisins de KNN

Malgré les résultats impressionnants de cet algorithme en matière de classification, le temps de réponse qu'il offre reste très important et devient exponentiel par rapport à la taille de l'ensemble d'apprentissage, qui dans notre cas peut devenir très important avec l'insertion de plus de données cliniques.

2.5.3. Hybridation de la logique floue et KNN

Pour la prédiction des maladies cardiaques, nous utilisons la combinaison de la logique floue avec l'algorithme k-plus proches voisins (KNN). La logique floue permet de représenter les caractéristiques du patient telles que l'âge, le sexe, la douleur thoracique (cp) et d'autres attributs de manière floue.

Une fois que les caractéristiques des patients sont évaluées en termes de catégories floues, l'algorithme KNN est utilisé pour prédire si un patient est en bonne santé ou s'il est atteint d'une maladie cardiaque.

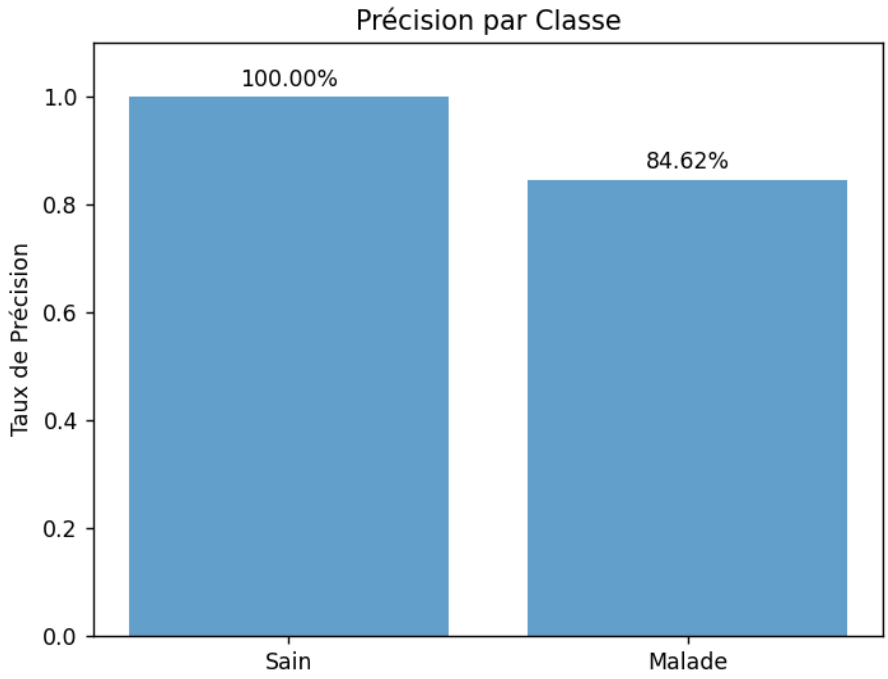

Figure 32 : le Taux de précision de sain et malade

Le graphique à barres montre la précision du modèle de classification pour les classes "Sain" et "Malade". La classe "Sain" a une précision de 100%, ce qui signifie que le modèle a correctement identifié toutes les instances saines. Pour la classe "Malade", la précision est de 84.62%, indiquant que le modèle a correctement identifié 84.62% des instances malades.

2.6. Évaluation des résultats et discussion

Après avoir entraîné les modèles proposés pour la prédiction des maladies cardiaques, nous avons obtenu les résultats suivants :

L'algorithme de logique floue basé sur les règles que nous avons précédemment décrites a permis d'atteindre une précision de 80 %, cette précision est assez acceptable pour la détection de maladies cardiovasculaires, car elle pousse le médecin à faire plus d'explorations avant de décliner la possibilité d'atteinte de maladie.

L’algorithme 1-NN standard, sans ajouts, a offert une exactitude de 96.67 %, cette exactitude est considérée comme très bonne par rapport à la prédiction d'atteinte de maladies cardiaques.

Enfin, lorsque nous avons combiné les deux concepts, la logique floue et le 1-NN, en un seul algorithme hybride, nous avons pu atteindre une précision de 93%.

Ces résultats montrent que l'algorithme KNN peut servir d'optimiseur a un modèle basé sur la logique floue. Le KNN permet d'améliorer la précision du modèle sans pour autant augmenter le temps de réponse.

Le tableau 1 présente les résultats des métriques d'évaluation des performances (exactitude, Précision, Rappel, F1-score) pour chaque algorithme :

Modèles	Exactitude	Précision	Rappel	F1_score
KNN	96.67%	97%	100%	98%
Logique floue	80%	92%	85%	88%
Hybridation de la logique floue et KNN	93%	96%	96%	96%

Tableau 1 : Les résultats d'évaluation des performances

D'après le tableau 1, nous observons que le modèle KNN se distingue par ses performances élevées dans tous les aspects mesurés.

La logique floue donne des résultats satisfaisants mais moins élevés par rapport aux deux autres modèles.

L'hybridation de la logique floue et KNN offre une précision élevée, bien que son rappel soit légèrement inférieur à celui du KNN.

Dans ce qui suit nous allons comparer les résultats que nous avons obtenus à ceux présentes par les chercheurs dont les algorithmes ont été cite dans le chapitre 02.

Il faudrait rappeler que les résultats de la littérature traités ont été obtenus en utilisant la base de données " Heart " à laquelle nous avons fait appel dans ce travail.

Test réalisé par	Année de publication	Type de méthode	Exactitude
Littérature [33]	2016	Diagnostic : Floue de type 2	73.78%
Littérature [34]	2015	Prédiction : Hybride génétique floue	86.00%
Littérature [48]	2014	Diagnostic : Modèle floue	88.79%
Littérature [56]	2010	Diagnostic : Modèle floue	94.00%
Littérature [61]	2016	Diagnostic : Modèle floue	94.05%
Littérature [64]	2019	Diagnostic : Modèle floue	91.00%
Littérature [68]	2020	Diagnostic : Modèle floue	96.60%
Littérature [35]	2021	Diagnostic : Réseaux neuronaux floue	94.55%
Littérature [65]	2017	Diagnostic : Modèle floue	97.78%
Notre modèle	2024	La logique floue	80%
Notre modèle	2024	KNN	96.67%
Notre modèle	2024	Hybride de la logique floue et KNN	93%

Tableau 2 : La comparaison des résultats [W18].

Le tableau ci-dessus montre que notre algorithme proposé se classe de manière respectable parmi les modèles basés sur la logique floue les plus récents dans la prédiction des maladies cardiaques.

Nous pouvons émettre, à partir de ces résultats, que notre modèle offre une précision de classification surpassant la plupart des algorithmes vues en chapitre 02, et se rapproche du meilleur algorithme présenté dans le même chapitre.

Conclusion

Dans ce chapitre, nous avons exploré les approches techniques utilisées pour le développement d'un système de prédiction des maladies cardiaques. Nous avons examiné les performances de différents modèles, notamment la logique floue, KNN, et la logique floue combinée avec KNN.

Conclusion générale

La prévision des maladies cardiaques à l'aide de techniques d'apprentissage automatique est un domaine prometteur de la médecine moderne. Ainsi, après une longue période de travail, nous avons enfin terminé notre mémoire de fin d'étude portant sur l'apprentissage automatique pour prédire les maladies cardiaques.

Notre travail consiste à créer un système de prédiction des maladies cardiaques en utilisant l'algorithme d'apprentissage supervisé K plus proches voisins (KNN) et la logique floue, et la combinaison de la logique floue avec KNN, et faire la comparaison des performances des algorithmes et le choix de l'algorithme le plus efficace.

D'après notre recherche, la logique floue donné une exactitude de 80 %, tandis que l'algorithme KNN a atteint une exactitude de 96.67%. Lorsque nous avons combiné la logique floue avec KNN, nous avons pu atteindre une exactitude de 93% avec un temps de réponse acceptable.

En perspective, nous proposons de combiner notre approche aux concepts d'apprentissage profond pour une meilleure analyse et prédiction précoce des maladies cardiaques. De plus, nous suggérons d'effectuer une analyse de notre étude avec d'autres ensembles de données ou des données extraites des experts médicaux afin d'inclure de nouveaux attributs pour tester les performances du système proposé.

Ce mémoire nous a permis d'explorer de nouveaux domaines et d'acquérir des connaissances dans les domaines de l'analyse de données et d'apprentissage automatique. Cela nous a aussi permis de prendre conscience de nos lacunes et de les améliorer par la suite.

Enfin, nous espérons que notre recherche inspirera d'autres chercheurs et donnera naissance à d'autres études qui pourraient contribuer à aider les patients cardiaques.

Bibliographie

A. Référence Bibliographiques

- **[1]** C. Khanji, " Évaluation de la qualité des soins et des services préventifs cardiovasculaires en première ligne ", Thèse présentée en vue de l'obtention du grade de Doctorat en Sciences pharmaceutiques Montréal, Université de Montréal, 2018 p23.
- **[2]** S. P. Dhyan Chandra Yadav, "Prediction of Heart Disease Using Feature Selection and Random Forest Ensemble Method", International Journal for Pharmaceutical Research Scholars, 2020, p56-66.
- **[3]** MOHAMMED HOCINE TEDJINI "Conception et Réalisation D'un électrocardiographe « ECG » numérique", Mémoire de projet de fin d'études Pour l'obtention du diplôme de MASTER,2014, p28.
- **[4]** M. Yousef, A. Sulieman, C. Edward, and B. Ahmed, "A Review on Diagnostic Procedures for the Cardiovascular System, " Int. J. Sci. Res., India Online ISSN: 2319-7064, 2013.
- **[5]** T. Jernberg, "Common investigations in cardiovascular disease," Hear. Educ. Rehabil. Toolkit, pp. 1–3, 2014
- **[6]** Mlle. SOLTANI Khadidja, "Etude d'un Système Intelligent Pour Suivre les Patients Atteints de Maladies Cardiaques", Mémoire de fin d'études de Master, UNIVERSITE ABDELHAMID IBN BADIS DE MOSTAGANEM,2023, p06
- **[7]** Dr George Argyriou, "Réaliser & interpréter un ECG", AppIC – article Template, ESICM Intensive Care Nursing Application,2017, p01.
- **[8]** OUALI Mohammed Assam, "Modélisation, Débruitage, Extraction des caractéristiques et Classification des signaux électrocardiogrammes ECG ", THÈSE pour obtenir le titre de Docteur en Sciences Docteur en Sciences en électronique Option : Communication, Université Batna 2 – Mostefa Ben Boulaïd ,2012, p12
- **[9]** Ahmad Abubakar, Carlos Frederico Meschini Almeida, Matheus Gemignani "Review of Artificial Intelligence-Based Failure Detection and Diagnosis Methods for Solar Photovoltaic Systems", Mdp,2019 p04.
- **[10]** S. Umadevi and K. S. J. Marseline, "A survey on data mining classification algorithms, ", International Conference on Signal Processing and Communication (ICSPC'17) ,2017, p265

- **[11]** CHERIFI Wassim, BOUDILMI Fateh "Prédiction des maladies cardiaques à l'aide des techniques d'apprentissage statistiques", Mémoire présenté pour l'obtention Du diplôme de Master Académique UNIVERSITE MOHAMED BOUDIAF - M'SILA 2022 p23
- **[12]** F. Talal, "Detecting heart failure at an early stage to reduce the high mortality rate using various machine learning algorithms ", THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE & SOCIETY, Tilburg University 2022 p12
- **[13]** M. Raihan, "Prediction on Ischemic Heart Disease using Machine Learning Approaches ", Khulna University of Engineering & Technology Khulna 9203, Bangladesh, 2019, p32
- **[14]** Eve Mathieu-Dupas, "Algorithme des k plus proches voisins pondérés et application en diagnostic", 42èmes Journées de Statistique, Marseille, France, France ,2010, p04.
- **[15]** Sadegh Bafandeh Imandoust, Mohammad, "BolandraftarApplication of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background ", S B Imandoust et al. Int. Journal of Engineering Research and Applications, Vol. 3, Issue 5, 2013, p608.
- **[16]** Kamil Murât EKSIOGLU, "MODÉLISATION DE LA DÉPENDANCE CONTEXTUELLE DES CONCEPTS FLOUS : LA STRUCTURE SFC", Thèse de doctorat en sciences appliquées Spécialité : Génie électrique, 2000, p7.
- [**17]** F. Chevrie F. Guély, "Cahier technique n° 191 La logique floue ", Schneider Electric,1998 p05.
- **[18]** Mme.zerzouri nora, "Etude comparative entre la commande par logique floue Et régulateur PI d'un MCC ", MEMOIRE DE MASTER, UNIVERSITE BADJI MOKHTAR- ANNABA,2019, p17
- **[19]** BOURAIOU Ahmed, "DIAGNOSTIC DE DEFAUTS PAR LOGIQUE FLOUE"MEMOIRE Présenté pour l'obtention du diplôme de Magister, UNIVERSITE BADJI MOKHTAR ANNABA,2010, p62
- **[20]** SAFI Adila, "APPROCHE NEURONALE POUR LA CLASSIFICATION MORPHOLOGIQUE DES SIGNAUX ECG ", MEMOIRE DE FIN D'ETUDES Pour l'obtention du diplôme de Master, UNIVERSITÉ IBN-KHALDOUN DE TIARET,2017, p31
- **[21]** AID Lahcene, " Modélisation et simulation du confort dans un bâtiment intelligent par le formalisme DEVS", THÈSE En vue de l'obtention du Diplôme de Doctorat en

Sciences, Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf ,2016, p53.

- **[22]** Dr. Poornima G. Naik, Dr. Girish R. Naik, Mr. M.B. Patil, « Conceptualizing Python in Google COLAB», Shashwat Publication,2021, ISBN: 978-93-93557-43-8, p01.
- **[23]** BENSACI SAHRA, KHRAMSSIA NOUR EL HOUDA, " La prédiction des Maladies Basée sur les Symptômes à l'Aide de l'apprentissage Automatique" Présenté en vue de l'obtention du diplôme Master en informatique Université de Mohamed El Bachir El Ibrahimi de Bordj Bou Arréridj,2023, p23
- **[24]** Raoued Chabou, " système de surveillance intelligent basé sur l'loT pour la prédiction des maladies cardiaques " Mémoire de fin d'étude Pour l'obtention du diplôme de MASTER, Université Echahid Cheikh Larbi Tebessi – Tébessa ,2023, p35
- **[25]** M. Maftouni, I. Turksen, M. F. Zarandi, F. Roshani, Type-2 fuzzy rule- based expert system for ankylosing spondylitis diagnosis, in: 2015 Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS) held jointly with 2015 5th World Conference on Soft Computing (WConSC), IEEE, 2015, pp. 1–5.
- **[26]** M. F. Shaik, M. M. Subashini, Anemia diagnosis by fuzzy logic using lab- view, in: 2017 International Conference on Intelligent Computing and Control (I2C2), IEEE, 2017, pp. 1–5.
- **[27]** D. Saikia, J. C. Dutta, Early diagnosis of dengue disease using fuzzy inference system, in: 2016 International Conference on Microelectronics, Computing and Communications (MicroCom), IEEE, 2016, pp. 1–6.
- **[28]** S. A. Biyouki, I. Turksen, M. F. Zarandi, Fuzzy rule-based expert system for diagnosis of thyroid disease, in: 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), IEEE, 2015, pp. 1–7.
- **[29]** El-Sappagh S., Ali F., Abuhmed T., Singh J., Alonso J. M., Automatic detection of Alzheimer's disease progression: An efficient information fusion approach with heterogeneous ensemble classifiers, Neurocomputing, 2022 Nov 1; 512:203–24.
- **[30]** El-Sappagh S, Saleh H., Sahal R., Abuhmed T., Islam S. R., Ali F., et al, Alzheimer's disease progression detection model based on an early fusion of cost-effective multimodal data, Future Generation Computer Systems. 2021 Feb 1; 115:680–99.
- **[31]** Ali F., El-Sappagh S., Islam S. R., Ali A., Attique M., Imran M., et al, An intelligent healthcare monitoring framework using wearable sensors and social networking data, Future Generation Computer Systems. 2021 Jan 1; 114:23–43.

- **[32]** Ali F., Islam S. R., Kwak D., Khan P., Ullah N., Yoo S. J., et al, Type-2 fuzzy ontology–aided recommendation systems for IoT–based healthcare. Computer Communications, 2018 Apr 1; 119:138–55.
- **[33]** Lafta H. A., Oleiwi W. K., A fuzzy petri nets system for heart disease diagnosis, Journal of Babylon University/Pure and Applied Sciences 25 (2) (2017) 317–328.
- **[34]** Sajiah A. M., Setiawan N. A., Wahyunggoro O., Interval type-2 fuzzy logic system for diagnosis coronary artery disease, Communications in Science and Technology 1 (2).
- **[35]** Santhanam T., Ephzibah E., Heart disease prediction using hybrid genetic fuzzy model, Indian Journal of Science and Technology 8 (9) (2015) 797.
- **[36]** Muhammad L., Algehyne E. A., Fuzzy based expert system for diagnosis of coronary artery disease in nigeria, Health and Technology 11 (2) (2021) 319–329. pmid:33614390
- **[37]** Kaur J., Khehra B. S., Fuzzy logic and hybrid based approaches for the risk of heart disease detection: State-of-the-art review, Journal of The Institution of Engineers (India): Series B (2021) 1–17.
- **[38]** Reddy G. T., Reddy M. P. K., Lakshmanna K., Rajput D. S., Kaluri R., Srivastava G., Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis, Evolutionary Intelligence 13 (2) (2020) 185–196.
- **[39]** Gadekallu T. R., Khare N., Cuckoo search optimized reduction and fuzzy logic classifier for heart disease and diabetes prediction, International Journal of Fuzzy System Applications (IJFSA) 6(2) (2017) 25–42.
- **[40]** Wiharto W., Kusnanto H., Herianto H., Interpretation of clinical data based on c4.5 algorithm for the diagnosis of coronary heart disease, Health- care informatics research 22 (3) (2016) 186–195. pmid:27525160
- **[41]** M. Tarawneh, O. Embarak, Hybrid approach for heart disease prediction using data mining techniques, in: International Conference on Emerging Internetworking, Data & Web Technologies, Springer, 2019, pp. 447–454.
- **[42]** Tan C. H., Tan M. S., Chang S. W., Yap K. S., Yap H. J., Wong S. Y., Genetic algorithm fuzzy logic for medical knowledge-based pattern classifi- cation, Journal of Engineering Science and Technology 13 (2018) 242–258.
- **[43]** Gadekallu T. R., Gao X.-Z., An efficient attribute reduction and fuzzy logic classifier for heart disease and diabetes prediction, Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science) 14 (1) (2021) 158–165.

- **[44]** V. S. Dehnavi, M. Shafiee, The risk prediction of heart disease by using neuro-fuzzy and improved goa, in: 2020 11th International Conference on Information and Knowledge Technology (IKT), IEEE, 2020, pp. 127–131.
- **[45]** L. P. Koyi, T. Borra, G. L. V. Prasad, A research survey on state of the art heart disease prediction systems, in: 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), IEEE, 2021, pp. 799–806.
- **[46]** Liu Y., Eckert C. M., Earl C., A review of fuzzy ahp methods for decision- making with subjective judgements, Expert Systems with Applications (2020) 113738.
- **[47]** Wadhawan S., Maini R., A systematic review on prediction techniques for cardiac disease, International Journal of Information Technologies and Systems Approach (IJITSA) 15 (1) (2022) 1–33.
- **[48]** Hameed A. Z., Ramasamy B., Shahzad M. A., Bakhsh A. A. S., Efficient hybrid algorithm based on genetic with weighted fuzzy rule for developing a decision support system in prediction of heart diseases, The Journal of Supercomputing (2021) 1–21.
- **[49]** Devi Y. N., Anto S., An evolutionary-fuzzy expert system for the diagnosis of coronary artery disease, Int. J. Adv. Res. Comput. Eng. Technol 3 (4) (2014) 1478–1484.
- **[50]** Wiharto W., Kusnanto H., Herianto H., Hybrid system of tiered multi-variate analysis and artificial neural network for coronary heart disease diagnosis, International Journal of Electrical and Computer Engineering 7 (2) (2017) 1023.
- **[51]** Priyatharshini R., Chitrakala S., A self-learning fuzzy rule-based system for risk-level assessment of coronary heart disease, IETE Journal of Research 65 (3) (2019) 288–297.
- **[52]** Song J., Ni Z., Jin F., Li P., Wu W., A new group decision-making approach based on incomplete probabilistic dual hesitant fuzzy preference relations, Complex & Intelligent Systems 7 (6) (2021) 3033–3049.
- **[53]** Akhoondi R., Hosseini R., Mazinani M., A GA approach for tuning membership functions of a fuzzy expert system for heart disease prognosis development risk, Journal of Computing and Security 4 (1) (2017) 13–23.
- **[54]** Eisa M. M., Alnaggar M. H., Hybrid rough-genetic classification model for IoT heart disease monitoring system, in: Digital Transformation Technology, Springer, 2022, pp. 437–451.
- **[55]** Nazari S., Fallah M., Kazemipoor H., Salehipour A., A fuzzy inference- fuzzy analytic hierarchy process-based clinical decision support system for diagnosis of heart diseases, Expert Systems with Applications 95 (2018) 261–271.

- **[56]** Hern, Framework for the development of data-driven Mamdani-type fuzzy clinical decision support systems.
- **[57]** Adeli A., Neshat M., A fuzzy expert system for heart disease diagnosis, in Proceedings of the international multi conference of engineers and computer scientists, Hong Kong, Vol. 1, Citeseer, 2010, pp. 28–30.
- **[58]** Khatibi V., Montazer G. A., A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment, Expert Systems with Applications 37 (12) (2010) 8536–8542.
- **[59]** Kim J.-K., Lee J.-S., Park D.-K., Lim Y.-S., Lee Y.-H., Jung E.-Y., Adaptive mining prediction model for content recommendation to coronary heart disease patients, Cluster computing 17 (3) (2014) 881–891.
- **[60]** W. M. Baihaqi, N. A. Setiawan, I. Ardiyanto, Rule extraction for fuzzy expert system to diagnose coronary artery disease, in: 2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), IEEE, 2016, pp. 136–141.
- **[61]** V. Krishnaiah, G. Narsimha, N. S. Chandra, Heart disease prediction system using data mining technique by fuzzy k-nn approach, in: Emerging ICT for Bridging the Future-Proceedings of the 49th Annual Convention of the Computer Society of India (CSI) Volume 1, Springer, 2015, pp. 371–384.
- **[62]** Mokeddem S., Atmani B., Assessment of clinical decision support systems for predicting coronary heart disease, International Journal of Operations Research and Information Systems (IJORIS) 7(3) (2016) 57–73.
- **[63]** T. Kasbe, R. S. Pippal, Design of heart disease diagnosis system using fuzzy logic, in 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), IEEE, 2017, pp. 3183–3187.
- **[64]** O. Terrada, B. Cherradi, A. Raihani, O. Bouattane, A fuzzy medical diagnostic support system for cardiovascular diseases diagnosis using risk factors, in: 2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), IEEE, 2018, pp. 1–6.
- **[65]** P. Jain, A. Kaur, A fuzzy expert system for coronary artery disease diagnosis, in Proceedings of the Third International Conference on Advanced Informatics for Computing Research, 2019, pp. 1–6.

- **[66]** Uyar K., Ilhan A., Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks, Procedia computer science 120 (2017) 588–593.
- **[67]** Iancu I., Heart disease diagnosis based on mediative fuzzy logic, Artificial intelligence in medicine 89 (2018) 51–60. pmid:29859751
- **[68]** H. Kahtan, K. Z. Zamli, W. N. A. W. A. Fatthi, A. Abdullah, M. Ab- dulleteef, N. S. Kamarulzaman, Heart disease diagnosis system using fuzzy logic, in: proceedings of the 2018 7th International Conference on Software and Computer Applications, 2018, pp. 297–301.
- **[69]** P. Krishnan, V. Rajagopalan, B. I. Morshed, Anovel severity index of heart disease from beat-wise analysis of ECG using fuzzy logic for smart-health, in: 2020 IEEE International Conference on Consumer Electronics (ICCE), IEEE, 2020, pp. 1–5.
- **[70]** P.K. Anooj, Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules. J. King Saud Univ. Comput. Inf. Sci. 24(1), 27–40 (2012)
- **[71]** S. Bashir, U. Qamar, F.H. Khan, M.Y. Javed, MV5: a clinical decision support framework for heart disease prediction using majority vote based classifier ensemble. Arab. J. Sci. Eng. 39(11), 7771–7783 (2014)
- **[72]** D. Pal, K.M. Mandana, S. Pal, D. Sarkar, C. Chakraborty, Fuzzy expert system approach for coronary artery disease screening using clinical parameters. Knowl. Based Syst. 36, 162–174 (2012)
- **[73]** V. Khatibi, G.A. Montazer, A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment. Expert Syst. Appl. 37(12), 8536–8542 (2010)

B. Référence Web

- **[W1]** https://my.clevelandclinic.org/health/diseases/21493-cardiovascular-disease (consulté le 16/03/2024 à 11)
- **[W2]** https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (consulté le 28/02/2024 à 10)
- **[W3]** https://my.clevelandclinic.org/health/diseases/24129-heart-disease (consulté le13 /05/2024 à 12)
- **[W4]** https://www.american-hospital.org/page/10-conseils-pour-un-cœur enbonnesante (consulté le 18/2/2024 à 12)

- **[W5]** https://www.msdmanuals.com/fr/accueil/troubles-cardiaques-etvasculaires/diagnostic-des-maladies-cardiovasculaires/%C3%A9lectrocardiographie (consulté le 18/2/2024 à 13)
- **[W6]** https://www.lasantesurtout.com/est-il-vrai-que-certains-medicaments-peuvent-fausser-une-analyse-de-sang-si-oui-lesquels-comment-faut-il-faire-arreter-son-traitement/ (consulté le 13 /05/2024 à 11)
- **[W7]** https://www.ccjj.fr/echographie-cardiaque (consulté le 13 /05/2024 à 11)
- **[W8]** https://www.fedecardio.org/je-m-informe/la-coronarographie/ (consulté le 13 /05/2024 à 12)
- **[W9]** https://www.ottawaheart.ca/fr/examen-intervention/tomodensitometrie-cardiaque (consulté le 13 /05/2024 à 12)
- **[W10]** https://www.ottawaheart.ca/fr/examen-intervention/imagerie-par-resonance-magnetique-irm-cardiaque (consulté le 13 /05/2024 à 12)
- **[W11]** https://www.avf-biomedical.com/blog/conseils/que-represente-le-signal-ecg-et-comment-linterpreter/ (consulté le 13/05/2024)
- **[W12]** https://www.edx.org/learn/pycharm(consulté le 05/05/2024 à 13
- **[W13]** https://intellij-support.jetbrains.com/hc/en-us/community/posts/5909943756818-PyCharm-Logo-Update (consulté le 12/05/2024 à 17)
- **[W14]**https://docs.python.org/fr/3/tutorial/ (consulté le 05/05/2024 à 14)
- **[W15]** https://www.python.org/community/logos/(consulté le 12/05/2024 à 17)
- **[W16]** https://www.w3schools.com/python/numpy/numpy_intro.as (consulté le 05/05/2024 à 16)
- **[W17]** https://www.w3schools.com/python/matplotlib_intro.asp (consulte le 5/05/2024 à18)
- **[W18]** https://doi.org/10.1371/journal.pone.0293112.t007 (consulte le 18/05/2024 à18)