

**UNIVERSITÉ SAAD DAHLEB DE BLIDA1**

**Faculté des sciences**

Département d'informatique



**MÉMOIRE DE MASTER**

**En Informatique**

Option : Ingénierie Des Logiciels

**THÈME :**

**Modélisation des Connaissances  
Incertaines dans l'Internet des Objets  
(IOT) : Contexte des Objets Connectés  
Médicaux.**

Réalisé par

Benstaali Kamel

Benmesbah Salim

Encadré par

Dr. I. RIALI

Dr. M. FAREH

M.Lahiani Nesrine

M.Cherfa Imane

6 Juillet 2022

## Remerciements

Ce travail est le fruit de la combinaison d'efforts, de patience et du courage durant toutes les années de notre parcours.

Nous remercions tout d'abord ALLAH le tout puissant qui, par sa grâce nous a permis d'arriver au bout de nos efforts en nous donnant la santé, la force, la volonté et en nous faisant entourer des merveilleuses personnes dont nous tenons à remercier.

Nous tenons à exprimer notre immense gratitude à notre encadreur Mr. RIALI ISHAK et à notre promotrice Mme. FAREH MESSAOUDA, d'avoir accepté de nous encadrer et de nous avoir guidé et épaulé ainsi qu'à toutes les personnes qui ont contribué à la réalisation de notre travail.

Nous tenons à remercier également Dr BENKHEDDA, qui a été d'une aide précieuse et nous a fait bénéficier de sa vaste expérience.

Nous remercions également tous les membres du jury de nous avoir honorés de leur présence et pour avoir accepté d'évaluer notre travail. Enfin, nous adressons nos plus sincères remerciements à nos familles : Nos parents, nos sœurs, nos frères et tous nos proches et amis, qui nous ont accompagnés, soutenu, et encouragé tout au long de la réalisation de ce travail.

موجز

تطور مجال إنترنت الأشياء بسرعة في السنوات الأخيرة، مما أدى إلى ظهور العديد من التخصصات الفرعية. أحد أشهر هذه التخصصات هو إنترنت الأشياء الطبية، أو المعروف أكثر باسم IoMT (إنترنت الأشياء الطبية)، نظراً لأهمية الصحة في حياتنا والفوائد التي تأتي مع تطور هذا المجال. ومع ذلك، فإن إحدى الصعوبات التي تمت مواجهتها هي عدم قدرة أنظمة IoMT على تمثيل حالة عدم اليقين في المعلومات الطبية. نتيجة لذلك، فإن الهدف من عملنا هو اقتراح نموذج دلالي يعالج عدم اليقين في مجال IoMT.

الأنطولوجيات الكلاسيكية، وهي طريقة لتمثيل معرفة IoMT، غير قادرة أيضاً على التعامل مع عدم اليقين، ولهذا السبب اخترنا استخدام أنطولوجيا عامضة تم إنشاؤها من خلال طريقة FODM (منهجية لتطوير الأنطولوجيات العامضة) من أجل تمثيل المعرفة الطبية غير المؤكدة في مجال IoMT، وهذا من خلال ربط الأنطولوجيا العامضة بشبكة بايزية عامضة تأخذ في الاعتبار الاحتمالات المتعلقة بالمعرفة العامضة. تجري الدراسة للمساعدة في تشخيص مرضى السكري من النوع 2.

الكلمات الرئيسية: IoMT، معرفة غير مؤكدة، أنطولوجيات عامضة، شبكات بايزية عامضة، تشخيص.

## Résumé

Le domaine de l'internet des objets a rapidement évolué au cours des dernières années, donnant naissance à une multitude de sous-disciplines. L'une des plus connues d'entre ces disciplines est l'internet des objets médicaux ou plus communément appelée IoMT (Internet of Medical Things), et ce compte tenu de l'importance de la santé dans notre vie et des avantages qui viennent avec l'évolution de ce domaine. Cependant, une des difficultés rencontrée est l'incapacité des systèmes de l'IoMT de représenter adéquatement l'incertitude des informations du domaine médical. De ce fait, l'objectif de notre travail est de proposer un modèle sémantique qui permet de traiter l'incertitude dans le domaine de l'IoMT.

Les ontologies classiques, qui sont un moyen de représenter les connaissances en IoMT, sont elles-aussi incapables de traiter l'incertitude, c'est pourquoi, nous avons opté pour l'utilisation d'une ontologie floue construite grâce à la méthode FODM (fuzzy ontology development methodology) afin de pouvoir représenter les connaissances médicales incertaines dans le domaine de l'IoMT, et cela en associant l'ontologie floue à un réseau bayésien flou qui permettra de prendre en compte les probabilités sur les connaissances floues. L'étude est réalisée pour l'aide au diagnostic des patients potentiellement atteints du diabète de type 2.

**Mots clés :** IoMT, connaissances incertaines, ontologies floues, réseaux bayésiens flous, diagnostic.

## Abstract

The field of the Internet of Things has evolved rapidly in recent years, giving rise to a multitude of sub-disciplines.

One of the best known of these disciplines is the Internet of Medical Things or more commonly known as IoMT (Internet of Medical Things), and this given the importance of health in our lives and the benefits that come with the evolution of this field. However, one of the difficulties encountered is the inability of IoMT systems to adequately represent the uncertainty of medical information. This is why the objective of our work is to propose a semantic model that makes it possible to deal with uncertainty in the field of IoMT.

Classical ontologies, which are a way to represent knowledge in IoMT, are also unable to process uncertainty, which is why we have opted for the use of a fuzzy ontology built using the FODM method (fuzzy ontology development methodology) in order to be able to represent uncertain medical knowledge in the field of IoMT, and this by associating the fuzzy ontology with a fuzzy Bayesian network that will allow to take into account the probabilities on the fuzzy knowledge. This study is carried out for the diagnostic assistance of patients potentially suffering from type 2 diabetes.

**Key words :** IoMT, uncertain knowledge, fuzzy ontologies, fuzzy bayesian networks, diagnosis

---

# Table des matières

---

<b>Table des figures</b>	<b>ix</b>
<b>Liste des tableaux</b>	<b>xi</b>
<b>Introduction générale</b>	<b>1</b>
<b>1 Représentation des connaissances dans l'internet des objets médicaux (IoMT)</b>	<b>5</b>
1.1 Introduction . . . . .	5
1.2 L'internet des objets (IoT) . . . . .	5
1.2.1 Définition . . . . .	6
1.2.2 Domaines d'application de l'IoT . . . . .	6
1.2.3 Fonctionnement de l'IoT . . . . .	7
1.2.4 Le processus de fonctionnement de l'IoT . . . . .	8
1.2.5 Avantages et inconvénients de l'IoT . . . . .	9
1.3 L'internet des objets médicaux (IoMT).....	10
1.4 Les ontologies.....	10
1.4.1 Définition .....	10
1.4.2 Rôle et intérêt des ontologies .....	11
1.4.3 Composants d'une ontologie.....	11
1.4.4 Les méthodes d'ingénierie ontologique .....	12
1.5 Quelques applications d'ontologies dans l'IoMT .....	13
1.5.1 Travail de : "N.Sharma et al. , 2021" .....	13
1.5.2 Travail de : "V.Kumar et al. , 2015" .....	13
1.5.3 Travail de : "F.Alsubaei et al. , 2019" .....	13
1.5.4 Analyse & Discussion.....	14
1.6 Limites des ontologies classiques.....	15
1.7 Connaissances incertaines .....	15
1.7.1 L'incertitude.....	15
1.7.2 Types d'incertitude .....	15

1.7.3	L'incertitude en médecine.....	16
1.8	Conclusion.....	17
<b>2</b>	<b>Approches de modélisation de la connaissance ontologique incertaine dans l'IoMT :</b>	
	<b>Étude comparative</b>	<b>19</b>
2.1	Introduction .....	19
2.2	Représentation des connaissances floues dans les ontologies .....	19
2.2.1	La logique floue .....	20
2.2.2	Les ontologies floues.....	21
2.2.3	La méthode de développement des ontologies floues .....	23
2.3	Représentation des connaissances probabilistes dans les ontologies.....	25
2.3.1	La probabilité .....	25
2.3.2	Les réseaux bayésiens .....	26
2.3.3	Ontologies probabilistes.....	30
2.4	Analyse des travaux existants et discussion .....	31
2.4.1	Travail de : "N.Selvan et al. , 2018" .....	31
2.4.2	Travail de : "R. Gorrepati et al. , 2021" .....	32
2.4.3	Travail de : "M.Lesani et al. , 2019" .....	32
2.4.4	Travail de : "Y.Benazzouz et al. , 2014" .....	33
2.4.5	Travail de : "G.Lee et al. , 2016" .....	33
2.4.6	Travail de : "S.Chabridon et al. , 2017" .....	33
2.4.7	Travail de : "L.Varghese et al. , 2021" .....	33
2.4.8	Travail de : "S.Titi et al. , 2020" .....	34
2.4.9	Travail de : "G.Lunardi et al. , 2018" .....	34
2.4.10	Tableau comparatif.....	35
2.4.11	Discussion .....	36
2.5	Conclusion.....	36
<b>3</b>	<b>Conception du système</b>	<b>39</b>
3.1	Introduction .....	39
3.2	Motivation derrière le travail .....	39
3.3	Étapes de construction du système .....	40
3.4	Préparation des données .....	41
3.5	Construction d'une ontologie floue pour le diabète de type 2 et l'emploi de l'internet des objets médicaux .....	43
3.5.1	Phase 1 : Objectif et portée de l'ontologie .....	43
3.5.2	Phase 2 : Identifier les besoins flous .....	44
3.5.3	Phase 3 : Déterminer les informations floues et précises .....	44
3.5.4	Phase 4 : Envisager la réutilisation des ontologies existantes.....	45
3.5.5	Phase 5 et 6 : Réutiliser les éléments d'ontologies floues et les corriger.	45

3.5.6	Phase 7 : Définir les éléments flous de l'ontologie.....	45
3.5.7	Phase 8 : Définir des éléments précis de l'ontologie .....	47
3.5.8	Phase 9 : Formalisation .....	52
3.5.9	Phase 10 : Validation .....	52
3.5.10	Phase 11 : Documentation.....	54
3.5.11	Instanciation de l'ontologie floue .....	54
3.6	Construction du réseau Bayésien Flou .....	56
3.6.1	Catégorisation du dataset .....	56
3.6.2	Apprentissage de la structure .....	58
3.6.3	Apprentissage des paramètres .....	58
3.6.4	Fuzzification du RB .....	59
3.7	Processus de diagnostic de l'état du patient .....	59
3.8	Conclusion.....	60
<b>4</b>	<b>Implémentation et test du système</b>	<b>61</b>
4.1	Introduction .....	61
4.2	Implémentation du système .....	61
4.2.1	Outils de programmation .....	61
4.3	Interfaces utilisateur (UI) .....	65
4.3.1	Interface l'accueil.....	65
4.3.2	Interface d'ajout de patient .....	66
4.3.3	Interface de modification de patient.....	67
4.3.4	Interface d'inférence patient .....	68
4.3.5	Interface de résultats .....	68
4.4	Test et Validation .....	69
4.4.1	Mesures d'évaluation .....	69
4.4.2	Résultats obtenus.....	70
4.5	Conclusion.....	77
	<b>Conclusion générale</b>	<b>79</b>
	<b>Bibliographie</b>	<b>8</b>



---

# Table des figures

---

1.1	Les types d'incertitude [1].....	16
2.1	(A) logique booléenne (B) logique floue.....	21
2.2	Exemple de concepts flous [2] .....	22
2.3	Flux de travail du FODM [3] .....	23
2.4	Un réseau bayésien simple contenant 4 variables aléatoires .....	27
2.5	Un exemple de l'application de l'évidence virtuelle.....	30
3.1	Processus de développement du système .....	41
3.2	Dataset utilisé .....	41
3.3	Propriétés du dataset.....	42
3.4	Fonctions d'appartenance de : "AgeYoung" et "AgeOld" .....	47
3.5	Représentation taxonomique de l'ontologie dans Protégé .....	52
3.6	Activation du raisonnement.....	53
3.7	L'ontologie en mode raisonnement activé .....	53
3.8	Dataset avec des valeurs discrètes .....	57
3.9	Structure de réseau bayésien .....	58
3.10	Le réseau bayésien après l'apprentissage des paramètres .....	59
3.11	Le processus complet de diagnostic incluant l'inférence du RBF et l'ontologie .....	60
4.1	Interface graphique de Protégé.....	62
4.2	Hierarchie des concepts .....	62

4.3	Relations entre les concepts .....	63
4.4	Attributs des concepts .....	63
4.5	Interface d'accueil .....	66
4.6	Interface d'ajout de patient.....	67
4.7	Interface de modification de patient .....	67
4.8	Interface d'inférence de patient .....	68
4.9	Interface de résultat .....	69
4.10	Accuracy obtenus par chaque classifieur sous forme de graphe .....	73
4.11	Rappels obtenus par chaque classifieur sous forme de graphe.....	74
4.12	Précisions obtenues par chaque classifieur sous forme de graphe .....	76

---

# Liste des tableaux

---

1.1	Quelques utilisations d'ontologies classiques dans l'IoMT .....	14
2.1	Tableau des fonctions d'appartenance de la variable GLYC .....	30
2.2	Tableau comparatif.....	36
3.1	Concepts, attributs flous ainsi que leurs fonctions d'appartenance.....	46
3.2	Concepts et attributs précis - partie 1 .....	48
3.3	Concepts et attributs précis - partie 2 .....	49
3.4	Relations - partie 1 .....	50
3.5	Relations - partie 2 .....	51
3.6	Table des catégorisations de chaque colonne du dataset.....	57
4.1	Matrice de confusion .....	70
4.2	Tableau de comparaison des 'accuracy' (%) obtenues par chaque classifieur.....	72
4.3	Tableau de comparaison des 'rappels' (%) obtenues par chaque classifieur.....	74
4.4	Tableau de comparaison des 'précisions' (%) obtenues par chaque classifieur.....	75



---

# Introduction générale

---

## Contexte

Au cours des dernières années, l'évolution du domaine de l'internet des objets (IoT) a connu une croissance phénoménale, permettant à des milliards d'appareils hétérogènes et intelligents à travers le monde de se connecter entre-eux.

Le développement de l'Internet des objets dans le domaine de la santé a donné naissance à une nouvelle tendance, plus connue sous le nom de **L'Internet of Medical Things ( IoMT )**. Il désigne l'interconnexion des dispositifs médicaux et leur intégration dans les réseaux de santé afin de partager et d'échanger des données médicales afin de les utiliser efficacement. En effet, plusieurs systèmes hétérogènes et applications médicales basés sur l'IoMT ont été récemment développés.

Compte tenu du nombre très élevé et toujours en croissance des Objets Connectés Médicaux (MCO), une quantité massive et hétérogène de données continue d'être générée, ces données peuvent être classées en 2 catégories :

**La première concerne les données techniques** qui représentent les propriétés et les capacités des objets médicaux connectés (MCO) déployés (par exemple l'autonomie, la consommation électrique, le niveau de batterie, etc.).

**La deuxième catégorie concerne les données spécifiques du patient** (tension artérielle, glycémie, fréquence cardiaque, etc..), utiles pour le diagnostic et le traitement de l'état du patient. Ces données présentent une hétérogénéité sémantique car elles sont obtenues dans des formats différents et à partir de MCO hétérogènes. De plus, la plupart de ces MCO sont mobiles et leurs contextes de déploiement évoluent dans le temps selon plusieurs critères, notamment temporels et géographiques. Cela implique une modification de leur description et de leurs données. Ces caractéristiques ont conduit à une conception complexe de systèmes médicaux interopérables capables de communiquer et d'échanger des données de manière compréhensible et fluide.

En effet, pour garantir l'interopérabilité entre les systèmes IoMT , une représentation sémantique des MCO et de leurs données est nécessaire pour définir un modèle intégré et partageable. Dans cette perspective, **l'ontologie** est connue comme étant une solution prometteuse et efficace

pour présenter explicitement les connaissances IoMT et leurs relations.

## Problématique

L'incertitude est l'un des principaux problèmes rencontrés lors de la construction des systèmes d'Objets Médicaux Connectés. En effet, les principales caractéristiques de l'IoT qui influencent l'incertitude incluent l'hétérogénéité des appareils, un grand nombre d'appareils utilisés signifie une grande différence dans leurs capacités de calcul et de communication, sans oublier que l'incertitude est omniprésente dans le domaine médical que ce soit dans l'incomplétude des informations qu'utilisent les médecins afin de diagnostiquer leurs patients où ils sont obligés de prendre des décisions, ou que ce soit l'incertitude de ces informations.

De plus, un examen approfondi des études existantes dans le domaine de l'IoMT nous a permis de tirer quelques conclusions relatives aux limites de ce domaine, celles-ci peuvent se résumer dans les points suivants :

- Les approches suggérées dans le domaine de l'IoMT ignorent la représentation de connaissances incertaines.
- Les approches suggérées dans le domaine de l'IoMT ne fournissent pas de mécanismes pour déduire de nouvelles connaissances basées sur une sémantique incertaine.

Ceci nous amène à nous poser certaines questions cruciales qui sont les suivantes :

1. Comment modéliser les connaissances incertaines des systèmes de l'IoMT ?
2. Comment répondre adéquatement aux requêtes de décisions traitant de connaissances incertaines dans les systèmes de l'IoMT ?
3. Quelles sont les étapes à suivre, et outils à utiliser pour développer un système d'aide au diagnostic dans le contexte de l'IoMT ?

## Objectif du travail

Le but de ce travail est de proposer un nouveau modèle sémantique qui prend en compte les connaissances médicales incertaines dans le contexte de l'IoMT.

Comme évoqué précédemment, les ontologies sont une solution efficace afin de représenter les connaissances dans le domaine de l'IoMT, mais les ontologies classiques présentent des limites dans ce qui est de la gestion de l'incertitude inhérente au domaine médical qui se présume en très grande partie dans l'incertitude des données, ce qui a poussé les chercheurs du domaine à proposer de nouvelles techniques afin de permettre de pallier les lacunes des ontologies classiques.

Dans ce travail, les ontologies floues ont été proposées afin d'atténuer les limitations des ontologies classiques et permettre de gérer l'incertitude. Cependant, le flou n'est pas la seule

forme d'incertitude existante dans le domaine de l'IoMT, c'est pour cela que pour le traitement des connaissances incertaines dans notre travail, l'ontologie floue va être combinée avec l'utilisation d'un réseau bayésien flou qui permettra de prendre en compte les probabilités sur les connaissances floues qui influenceront la décision.

Quant aux raisons qui nous ont poussé à choisir cette approche, elles nous permettront de répondre aux questions posées précédemment :

- L'ontologie floue servira à représenter les connaissances médicales floues liées à la maladie choisie et aux appareils IoMT, les connaissances floues de ces domaines seront structurées et présentées sous formes de concepts floues tout en leur gardant des définitions réelles.
- Exploitation des capacités des RBF (réseaux bayésiens flous) pour représenter des connaissances probabilistes et incomplètes lorsqu'on veut prendre une décision en calculant les probabilités des événements liés avec des dépendances causales et en ayant fixé des observations floues dans certaines connaissances.

Comme le domaine médical est vaste, une modélisation par une ontologie pour toutes les maladies existantes est impossible, ce qui nous conduit à limiter notre cas d'utilisation en choisissant une maladie précise, dans notre travail, le diabète de type 2 a été choisi comme cas d'étude afin de modéliser le système d'aide au diagnostic.

## Organisation du mémoire

Afin d'atteindre le but de notre travail, l'organisation de notre mémoire sera comme suit :

- **Chapitre 1** : Présentation générale du domaine de l'Internet des objets, ainsi que des ontologies classiques.
- **Chapitre 2** : Présentation des ontologies floues et probabilistes, ainsi qu'une étude des travaux existants.
- **Chapitre 3** : Conception du système d'aide au diagnostic médical, ainsi que la construction d'une ontologie floue pour le diabète de type 2 et la construction du réseau bayésien flou.
- **Chapitre 4** : Implémentation et test du système, et démonstration des résultats obtenus par ce dernier.

Nous clôturerons ce mémoire par une conclusion générale, où nous aborderons les points principaux de notre travail, ainsi que des perspectives futures d'amélioration du travail proposé.





# **Représentation des connaissances dans l'internet des objets médicaux (IoMT)**

---

## **1.1 Introduction**

Depuis la fin des années 1980, Internet a évolué de manière spectaculaire. La dernière avancée en date est l'utilisation de ce réseau mondial pour la communication avec des objets ou entre objets, évolution nommée Internet des Objets (IoT pour Internet of Things). L'évolution de l'IoT est rapide : depuis 2014, le nombre d'objets connectés est supérieur au nombre d'humains connectés et il est prévu que 50 milliards d'objets seront connectés d'ici 2025.

Dans ce chapitre, nous allons présenter et définir les notions d'internet des objets et ses différents domaines d'utilisation, ses avantages et inconvénients, ainsi que les ontologies, et leur rôle, nous présenterons aussi les différents composants d'une ontologie et décrirons les différents types d'ontologies, ensuite, nous analyserons quelques travaux existants qui traitent des ontologies classiques dans le domaine de l'IoMT, et pour finir, nous parlerons des connaissances incertaines, leur types et l'incertitude dans le domaine médical.

## **1.2 L'internet des objets (IoT)**

Dans cette section, nous allons faire une brève présentation du domaine de l'internet des objets, nous commencerons par définir ce dernier, ensuite nous donnerons les principaux domaines d'utilisations de l'IoT, ainsi que les technologies nécessaires aux systèmes IoT, et le processus de fonctionnement de ces systèmes. Et pour finir, nous citerons quelques avantages et inconvénients de l'IoT.

## 1.2.1 Définition

L'Internet of Things (IoT), ou Internet des objets, décrit le réseau de terminaux physiques, les « objets », qui intègrent des capteurs, des logiciels et d'autres protocoles et technologies en vue de se connecter à d'autres terminaux et systèmes sur internet et d'échanger des données avec eux, ces terminaux peuvent aller du simple appareil de la vie de tous les jours, aux outils industriels d'une grande complexité [4]. L'IoT est considéré comme l'une des plus importantes technologies du 21<sup>ème</sup> siècle car elle permet de connecter des appareils du quotidien (voitures, appareils électroménagers, thermostats) à internet, et d'échanger des données avec d'autres appareils et d'autres réseaux.

L'Internet des objets n'est plus une fantaisie d'un monde futuriste, mais une réalité de notre vie quotidienne, la preuve étant que nous possédons tous au moins un appareil intelligent, ce qui fait que le rôle que jouera l'IoT dans les décennies à venir sera prédominant et ce dans divers domaines, car l'IoT a révolutionné les rapports entre les êtres humains et la technologie. L'IEEE (Institute of Electrical and Electronics Engineers) définit l'IoT comme un « réseau d'éléments chacun muni de capteurs qui sont connectés à Internet » [5].

## 1.2.2 Domaines d'application de l'IoT

Durant les 2 dernières décennies, l'expansion d'internet à l'échelle planétaire et le développement des technologies de l'information et de la communication ont fait que le nombre d'objets connectés a grandi exponentiellement, ce qui a mené à la diversification des domaines d'applications de l'IoT, des exemples de secteurs où l'IoT a révolutionné ces derniers :

— **Les transports :**

Depuis la création de l'IoT en 1999, le nombre des véhicules intelligents est en croissance, presque tous les véhicules vendus aujourd'hui dans le monde renferment déjà des capteurs et des moyens de communication pour traiter la congestion du trafic, la sécurité, la pollution et le transport efficace des marchandises, etc. L'objectif est qu'une voiture soit capable de communiquer de façon autonome avec d'autres véhicules ou avec une centrale de surveillance pour prévenir les accidents et réduire les coûts d'assurance [6].

— **La santé :**

Un ensemble de capteurs connectés permettent de créer un système de suivi à distance, où chaque patient peut être monitoré sans avoir besoin de se rendre à un centre médical, les données récoltées par ces systèmes serviront à améliorer la qualité des soins et fournir aux centres médicaux des informations en temps réel sur le patient, ce qui leur permettra de prendre des décisions rapides si besoin est. Par conséquent, une base de données massive sera disponible et permettra de demander conseil aux médecins en cas de besoin [6].

— **L'agriculture :**

L'utilisation d'objets connectés devient de plus en plus populaire dans le domaine de

l'agriculture, compte tenu des nombreuses améliorations que peuvent apporter ces derniers en matière de gestion du matériel agricole, prédiction des conditions météorologiques, aide à la maîtrise de l'irrigation, suivi de la croissance des plantes et végétaux, ce qui va profondément affecter et renouveler les techniques agricoles en place [6].

— **Les villes intelligentes :**

L'IoT a permis de nombreuses avancées dans le domaine des habitations intelligentes, et ce grâce à des systèmes qui permettent une automatisation des fonctions de contrôle des bâtiments, par le biais de mécanismes intelligents qui fournissent des données et informations et les interprètent en temps réel, ces dernières s'appliquent aussi dans le contexte de la sécurité, par exemple les systèmes électriques (compteurs. . .), les systèmes de traitement des déchets et des eaux usées, toutes ces données récoltées vont permettre aux différents bâtiments de réagir aux potentiels changements, tout cela combinés à d'autres services tels que les transports permettent de créer des villes intelligentes qui seront le fleuron de l'évolution humaine [7].

### 1.2.3 Fonctionnement de l'IoT

Pour garantir le bon fonctionnement de l'IoT, plusieurs systèmes technologiques sont requis. Dans ce qui suit, nous mettons en lumière certaines de ces technologies qui permettent l'interconnexion d'objets connectés par le biais d'internet [4], ces dernières permettent l'identification, le stockage, traitement et transfert des données dans des environnements physiques et univers virtuels [8], les plus importantes étant :

— **Objets connectés :**

L'IoT repose avant tout sur les objets connectés qui ont la capacité de capter des données et de les envoyer via le réseau Internet ou d'autres technologies, pour que celles-ci soient analysées et visualisées. Les objets connectés sont des dispositifs qui ont la capacité d'interagir avec leur environnement de manière autonome (doivent fonctionner indépendamment), elles possèdent entre autres une mémoire, une bande passante, une durée de vie, ainsi qu'une quantité d'énergie consommée qui vient généralement d'une batterie. Dans l'Internet des Objets, un objet peut être un véhicule, une machine industrielle, un petit appareil, etc [9].

— **RFID (Radio Frequency Identification) :**

Fait référence à l'utilisation des ondes radio pour identifier automatiquement des entités IoT (objets ou des personnes). Cette technologie permet de mémoriser et de récupérer des informations à distance grâce à une étiquette qui émet des ondes radio [10]. Il s'agit d'une méthode utilisée pour transférer les données contenues dans les étiquettes des objets, ou pour identifier ces derniers à distance [4].

— **WSN (Wireless Sensor Network) :**

Décrit un réseau composé d'un ensemble de capteurs qui communiquent à distance,

possédant chacun une capacité de traitement et différents types de mémoires, ainsi qu'une source d'alimentation, ces capteurs sans fils sont nécessaires au bon fonctionnement de l'Internet des objets [11].

— **M2M (Machine To Machine) :**

Réfère à «l'association des technologies de l'information et de la communication avec des objets intelligents dans le but de donner à ces derniers les moyens d'interagir sans intervention humaine avec le système d'information d'une organisation ou d'une entreprise »[12].

— **Schémas d'adressages :**

La capacité d'un système à identifier de manière précise les différents objets est primordiale au bon fonctionnement de l'internet des objets, car cela permet de garder trace et monitorer des millions d'appareils et en prendre le contrôle à distance [13].

— **Stockage et analyse des données :**

Une résultante du développement de ce domaine est la quantité phénoménale de données qui est produite, ce qui fait que le stockage de ces données est devenu un problème critique à mesure que cette quantité de données augmente.

— **La Visualisation :**

C'est un aspect essentiel de l'IoT, car elle représente l'interaction entre l'utilisateur et son environnement, grâce aux récents progrès en matière d'écrans tactiles en tout genres, l'utilisation des téléphones et tablettes est devenue plus intuitive et plus ludique. Créer des visuels attrayants et faciles à appréhender par les utilisateurs est une étape cruciale [13].

## 1.2.4 Le processus de fonctionnement de l'IoT

— **Identification :**

Pour pouvoir faire partie intégrante de l'IoT, les différents composants connectés qui font partie d'un réseau donné doivent être adressables, à l'instar des ordinateurs qui sont adressables et identifiables grâce à leur adresse IP (Internet Protocol), ces dispositifs doivent être accessibles pour garantir le bon fonctionnement du système [14], il existe plusieurs identifiants pour les composants de l'IoT, on peut citer par exemple les identificateurs de ressources uniformes (URI),et les adresses IP qui appartiennent au protocole (IPv6) [15].

— **Détection :**

Afin de pouvoir interagir avec le monde, les systèmes IoT sont équipés de capteurs qui permettent aux différents composants des systèmes de récolter des données, ceci est illustré dans les téléphones portables qui contiennent nombres de capteurs similaires (thermomètres , capteurs d'empreintes digitale, capteurs de position, détecteurs de proximité, podomètres) [16]. Parmi les caractéristiques qui distinguent les capteurs on cite : l'étendue de mesure, la résolution, la rapidité, la précision, le temps de réponse, la fréquence,...etc.

— **Communication :**

Une fois la récolte de données par les capteurs effectuée, il est important de transmettre ces informations à des systèmes capables de les analyser correctement et de s'adapter en fonction des résultats de ces analyses, par exemple si on considère les thermostats intelligents, les données de température ambiante doivent être transmises aux éléments du réseau connecté, ce dernier pourra ensuite prendre une décision sur les réglages du thermostat (température, puissance...etc). Afin de permettre ces échanges, plusieurs technologies de communication ont vu le jour, parmi les plus connues sont le Wi-Fi, Bluetooth, les code QR (Quick Response), ainsi que les étiquettes RFID (Radio Frequency Identification) [16].

— **Informatisation :**

Après les phases de collecte et transmission des données effectuées, il est nécessaire d'informatiser les données pour que le système puisse les traiter, pour ce faire, la plupart des systèmes IoT sont munis de microcontrôleurs, un exemple souvent cité serait les détecteurs de fumée dans les bâtiments et habitations qui peuvent en cas d'incendie automatiquement appeler les pompiers si les habitants ou employés sont absents [16].

— **Services :**

Les services réfèrent aux différentes fonctionnalités que proposent les systèmes d'IoT, et ces derniers dépendent et diffèrent en fonction du système choisi, par exemple un système de maison intelligente pourra modifier l'éclairage d'une chambre, et ceci grâce à des actionneurs [16].

— **Sémantique :**

La sémantique est un concept clé de l'IoT, car elle permet d'avoir un modèle commun de données qui par la suite peut être ré-utilisé et partagé entre des applications et des systèmes. Son utilité réside dans le fait de permettre aux utilisateurs et aux ordinateurs de trier les données de façon plus intuitive [15].

## 1.2.5 Avantages et inconvénients de l'IoT

Comme toute technologie ou avancée scientifique, l'IoT présente des avantages et inconvénients, comme décrit ci-dessous :

**-Avantages :**

- 1- Améliorer l'efficacité de l'entreprise grâce à la prise de décisions plus intelligentes et plus éclairées.
- 2- Permet de gagner du temps, en réduisant les déplacements inutiles.
- 3- Permet d'optimiser les dépenses aussi bien dans les simples ménages que dans les entreprises.
- 4- Améliore la productivité de l'entreprise et l'interaction avec les clients, grâce aux informations sur ces derniers, pour davantage satisfaire les attentes des clients.
- 5- Garantit un accès ubiquitaire à une quantité phénoménale d'informations, pour plus de confort

dans notre vie quotidienne.

**-Inconvénients :**

1- **La sécurité** : L'un des plus gros points faibles de l'IoT est la sécurité, en raison de l'énorme quantité de données manipulée, ce qui peut engendrer des pertes financières considérables et qui peut nuire à la réputation des entreprises.

2- **La confidentialité** : En cas de cyberattaque, la probabilité que les informations concernant l'entreprise (clients, employés..etc) soient modifiées, supprimées ou volées est extrêmement élevée, et ceci concerne davantage les systèmes IoT car toutes les données sont partagées.

3- **Les Défaillances** : Un système IoT qui gère des biens matériels comme une voiture intelligente ou une pompe à insuline, pourrait provoquer des conséquences désastreuses s'il venait à se produire une panne ou une défaillance quelconque [15].

## 1.3 L'internet des objets médicaux (IoMT)

La croissance de la santé numérique et l'utilisation de l'IoT pour créer des services axés sur la santé ont généré une nouvelle branche de l'IoT, à savoir : l'Internet des objets médicaux (IoMT). L'IoMT permet des interconnexions entre des infrastructures de dispositifs médicaux et des applications qui communiquent avec une grande variété de systèmes logiciels de soins de santé. En plus de permettre les interconnexions mentionnées ci-dessus, l'IoMT joue un rôle vital dans l'industrie de la santé pour augmenter la fiabilité, la productivité et la précision des dispositifs médicaux [17].

Comme vu précédemment, l'interopérabilité entre les dispositifs médicaux est indispensable pour réaliser un réseau IoMT, de ce fait, une représentation sémantique des données permettant le partage est nécessaire, l'outil de la modélisation sémantique choisi sera les ontologies. Nous parlerons plus en détail des ontologies dans la section suivante.

## 1.4 Les ontologies

### 1.4.1 Définition

Le mot ontologie est emprunté au mot latin scientifique « ontologia ». Ce dernier est composé du grec onto-, tiré du grec ancien ontos, qui signifie « étant, ce qui est », et de -logia, tiré du grec ancien logos qui signifie « discours, traité ». En philosophie, ce terme désigne une branche fondamentale de la métaphysique qui est selon Aristote « la science de l'être en tant qu'être ». L'informatique a repris ce terme à la philosophie par analogie. La définition informatique de l'ontologie a évolué à travers plusieurs philosophes pour arriver à un système d'information, d'où une ontologie est une spécification formelle et explicite d'une conceptualisation partagée, voici les termes importants dans cette définition à prendre en compte avec leurs significations :

- Le mot **formelle** : se rapporte au fait que l'ontologie devrait être compréhensible par une machine.
- Le mot **explicite** : signifie que le type des concepts utilisés, et les contraintes sur leur utilisation sont explicitement définis.
- Le mot **conceptualisation** : signifie qu'un modèle abstrait des phénomènes est identifié par des concepts appropriés à ces phénomènes.
- Le mot **partagée** : reflète que l'ontologie devrait capturer la connaissance consensuelle admise par les communautés [18].

### 1.4.2 Rôle et intérêt des ontologies

- Les ontologies définissent des vocabulaires structurés, regroupant des concepts utiles d'un domaine et de leurs relations et qui servent à organiser et échanger des informations de façon non ambiguë.
- Fournir une sémantique formelle d'un domaine du monde réel et permettant de lier le contenu exploitable par la machine.
- Les ontologies sont utilisées pour construire des bases de connaissances. Une base de connaissances peut être formée par une ontologie et un ensemble d'instances de ces classes.
- Le partage et la réutilisation d'information ou les ressources d'information peuvent être transmises entre les humains et les agents logiciels [19].
- Interopérabilité de logiciels : L'utilisation d'un modèle comme une spécification d'un logiciel est à la base de l'approche MDA (Model-Driven Architecture), dans cette approche, un modèle est utilisé pour générer le code de l'application. Le lien formel entre le modèle et le code permet alors de faire évoluer ce dernier lorsque la spécification du logiciel évolue.
- Le web : L'ensemble des services qui sont offerts par le web sont parfois assez isolés. Par conséquent, pour arriver au résultat escompté, un utilisateur doit, soit avoir une connaissance approfondie du web, soit passé par une période fastidieuse de navigation sur plusieurs sites web. Cette difficulté vient du fait que les noms utilisés pour décrire un même service sont très différents et dépendent du créateur du service. L'idée du web sémantique [20] est de développer des ontologies de domaine puis d'indexer les services par leur description en termes de ces ontologies, rendant ainsi la recherche automatisable et réalisable par des agents informatiques.
- La communication : Une ontologie facilite la communication en fournissant une spécification explicite d'un domaine.

### 1.4.3 Composants d'une ontologie

#### — Concepts :

Les concepts peuvent être une pensée, un principe, une notion profonde. Ils sont appelés

aussi termes ou classes de l'ontologie, selon Gomez Pérez, ces concepts peuvent être classifiés selon plusieurs dimensions : - Niveau d'abstraction (concret ou abstrait). - Atomicité (élémentaire ou composée). - Niveau de réalité (réel ou fictif) [21].

— **Relations :**

Les relations d'une ontologie désignent les différentes interactions et corrélations entre les concepts de l'ontologie, ces relations englobent les associations suivantes : Sous classe de (spécification ou généralisation), partie de (agrégation ou composition), associé à, instance de, est un ... etc [21].

— **Axiomes :**

Les axiomes sont utilisés pour décrire les assertions de l'ontologie qui seront considérés après comme vrais, cette détermination a pour but de définir les significations des composants d'ontologie, les contraintes sur les valeurs des attributs, et les arguments de relations [21].

— **Fonctions :**

Elles constituent des cas particuliers de relation, dans laquelle un élément de la relation, le n-ième est défini en fonction des n-1 éléments précédents [21].

— **Instances :**

C'est une définition extensionnelle de l'ontologie, par exemple les individus « Kamel » et « Salim » sont des instances du concept « personne » [21].

#### 1.4.4 Les méthodes d'ingénierie ontologique

Il existe plusieurs de méthodes d'ingénierie ontologique qui peuvent être résumées en 3 classes :

**1-Création d'une ontologie depuis le début :** Cela en commençant par un ensemble de connaissances explicites, on suit différentes étapes ordonnées pour concevoir tous les composants d'une ontologie [3].

**2-Fusion avec des ontologies existantes :** La fusion d'ontologies est "la création d'une nouvelle ontologie à partir de deux ou plusieurs ontologies existant avec des parties qui se chevauchent". Cette méthode d'ingénierie par deux grandes phases : La première est l'alignement d'ontologies qui cherche les correspondances entre leurs concepts, la deuxième phase est la fusion d'ontologies se basant sur l'alignement trouvé [22].

**3-Ré-ingénierie d'ontologies existantes :** Cela en commençant par une ontologie existante, et appliquer une méthodologie de construction pour ajouter une ou des extensions à cette dernière [3].

**4-Construction d'un outil d'évaluation d'ontologies :** Cela implique la création de différents outils pour vérifier la validation une ontologie, autrement dit, si une ontologie vérifie quelques critères.



## 1.5 Quelques applications d'ontologies dans l'IoMT

Dans ce qui va suivre, nous nous sommes intéressés à quelques travaux existants dans le domaine de l'internet des objets qui utilisent les ontologie classiques afin de proposer divers systèmes :

### 1.5.1 Travail de : "N.Sharma et al. , 2021"

Le but de cette étude est de proposer une plateforme de monitoring à distance des patients atteints de la covid-19 dans le contexte de l'internet des objets médicaux, les auteurs tentent de proposer une solution facilitée grâce à des capteurs bio-portables pour identifier les patients potentiellement atteint de la covid-19. En outre, les mêmes informations peuvent être utilisées pour avertir les personnes à proximité de faire preuve de prudence et d'adopter des mesures préventives avec le plus grand soin. Le système de surveillance à distance basé sur les ontologies analyse les défis liés à la sécurité et à la protection de la vie privée des patients.

Au cours de la simulation, il a été observé que le modèle proposé a atteint une précision de 96,33%, ce qui démontre l'efficacité de ce dernier. [23]

### 1.5.2 Travail de : "V.Kumar et al. , 2015"

Le but de cette étude est de proposer un système d'information qui propose des soins de santé, qui est également basé sur les ontologies et l'internet des objets. Ce système a pour objectif de tacler les défis en matière de sécurité et de protection de la vie privée des utilisateurs, ces besoins sont analysés par le système d'information. Les services médicaux d'urgence (SMU) sont un type de service d'urgence dédié à fournir des soins médicaux à l'extérieur de l'hôpital, le transport vers les soins définitifs. De plus, un plan d'infrastructure fonctionnelle est fourni pour montrer l'unification entre l'architecture d'application proposée avec l'Internet des objets et la hiérarchie de l'ontologie. [24]

### 1.5.3 Travail de : "F.Alsubaei et al. , 2019"

Le but de cette étude est de proposer un outil de recommandation qui modélise les concepts du domaine de l'internet des objets médicaux et les problèmes de sécurité en plus de recommander des dispositions de sécurité pour les patients. L'outil présenté utilise une ontologie sémantiquement enrichie pour modéliser les composants IoMT, les potentiels problèmes et les mesures de sécurité. L'ontologie développée définit plusieurs règles de sécurité en incluant les différents types de services, de vulnérabilités, de mesures,...etc afin de permettre au système proposé de faire des raisonnements pour bien évaluer plusieurs parties prenantes dans les systèmes IoMT. Les différents tests effectués sur le système (en comparant ses recommandations à des recommandations d'experts) ont démontré un résultat concluant, avec un taux de précision moyen de 94,7%, ce qui prouve l'efficacité de l'outil mis en place. [25]

Voici un tableau pour récapituler différents critères sur les 3 travaux présentés précédemment :

Critères Travaux	Domaine d'application et objectif	Méthode d'ingénierie de l'ontologie	Informations modélisées dans l'ontologie	Gestion de l'incertitude	Évaluation
N.Sharma et al. , 2021	Services de santé, le but de cette étude est de proposer une plateforme de monitoring à distance des patients atteints de la covid-19.	Création d'une nouvelle ontologie	Des comme indicateurs du COVID tel que la température du corps et de la peau, le rythme cardiaque, le taux de respiration,...etc Ces informations sont représentées dans l'ontologie en tant que mesures car elles sont capturées a partir de capteurs.	non	oui
V.Kumar et al. , 2015	Services de santé, le but ce système est dédié à fournir et partager des informations sur plusieurs maladies pour aider les médecins à prendre des décision en cas d'urgence.	Création d'une nouvelle ontologie	L'ontologie contient des informations modélisées représentant des détails sur plusieurs maladies telles que les traitements, les symptômes, et les causes...etc.	non	non
F.Alsubaei et al. , 2019	Sécurité pour les systèmes de santé dans l'IoMT, le but de cette étude est d'évaluer pour trouver des vulnérabilités et pour proposer des mesures de sécurité pour différentes parties prenantes de l'IoMT.	Création d'une nouvelle ontologie	L'ontologie modélise des concepts sur tous les aspects de la sécurité tel que les objectives de la sécurité (ex: confidentialité), les types de services (ex:cloud),dispositifs(ex: Pare-feu), mesures de sécurité(ex :detection, prévention), types de vulnérabilité(ex : Malware),types de parties prenantes(Patient,professionnel médical,administrateur système) ...etc	non	oui

TABLE 1.1 – Quelques utilisations d'ontologies classiques dans l'IoMT

### 1.5.4 Analyse & Discussion

Comme nous pouvons le voir, l'utilisation des ontologies est déjà très présente dans les travaux effectués dans le cadre de l'IoMT.

L'étude des maladies pour améliorer les services de santé font l'objet des travaux de "N.Sharma et al., 2021" et "V.Kumar et al., 2015", le travail de "F.Alsubaei et al., 2019" visait un autre problème de l'IoMT qui est la sécurité, nous remarquons également que les connaissances modélisées dans leurs ontologies développées sont adaptées à leurs besoins.

Cependant, les travaux analysés précédemment sont basés sur des ontologies classiques basées uniquement sur une logique booléenne, et ne reposent en aucun cas sur des connaissances incertaines ce qui peut limiter la contribution des ces travaux.

Dans la section suivante, nous allons introduire les limites des ontologies classiques.

## 1.6 Limites des ontologies classiques

Les langages d'ontologies standards du Web sémantique, tels que OWL et RDF(S), sont basés sur une logique précise et ne peuvent donc pas gérer des connaissances incomplètes ou partielles sur un domaine d'application. Cependant, le raisonnement humain s'appuie fréquemment sur des connaissances ou des données inexacts, incertaines ou imprécises.

Par exemple, en modélisation de domaine, en plus de savoir si « A est une sous-classe de B », on peut aussi savoir et souhaiter exprimer que « A est une petite sous-classe de B » ; ou, dans le cas où A et B ne sont pas logiquement liés, on peut toujours souhaiter exprimer que "A et B se chevauchent largement", ou on peut chercher à connaître le degré de similitude entre 2 concepts même si ces derniers ne sont pas reliés par une relation d'héritage.

De plus, une description (d'une classe ou d'un individu) que l'on souhaite saisir dans une ontologie peut être bruyante et incertaine, ce qui conduit souvent à des conclusions trop généralisées dans un raisonnement basé sur la logique car parfois nous avons besoin de résultats bien précis. L'incertitude devient plus répandue dans la spécification conceptuelle dans plusieurs ontologies utilisées dans plusieurs domaines.

## 1.7 Connaissances incertaines

### 1.7.1 L'incertitude

Plusieurs chercheurs y compris des philosophes ont établi différentes définitions de l'incertitude pour divers domaines, mais de façon générale l'incertitude sera définie comme l'incapacité perçue d'un individu à prédire quelque chose avec précision. Un individu éprouve de l'incertitude parce qu'il se perçoit comme manquant d'informations suffisantes pour prédire avec précision ou parce qu'il se sent incapable de faire la distinction entre les données pertinentes et les données non pertinentes (Gifford, Bobbitt et Slocum, 1979)[26].

### 1.7.2 Types d'incertitude

Il existe plusieurs types d'incertitude selon l'organisme W3C [1] (voir la **figure 1.1**) , que nous présentons brièvement ci-dessous :

— **L'incomplétude :**

On parle de l'incomplétude si les informations sur le monde sont incomplètes, certaines informations sont manquantes [27].

— **L'inconsistance :**

On parle de l'inconsistance si une phrase contient des connaissances et des termes contradictoires [27].

— **L’ambiguïté :**

Les référents des termes dans une phrase sur le monde ne sont pas clairement spécifiés et il ne peut donc pas être déterminé si la phrase est satisfaite [27].

— **L’aléatoire :**

Une classe pour laquelle il n’existe pas une loi statistique régissant si les instances sont satisfaites [27].

— **L’imprécision :**

Il n’y a pas de correspondance précise entre les termes de la phrase et les référents dans le monde [27].

La **figure 1.2** ci-dessous représente une taxonomie pour les types d’incertitudes :

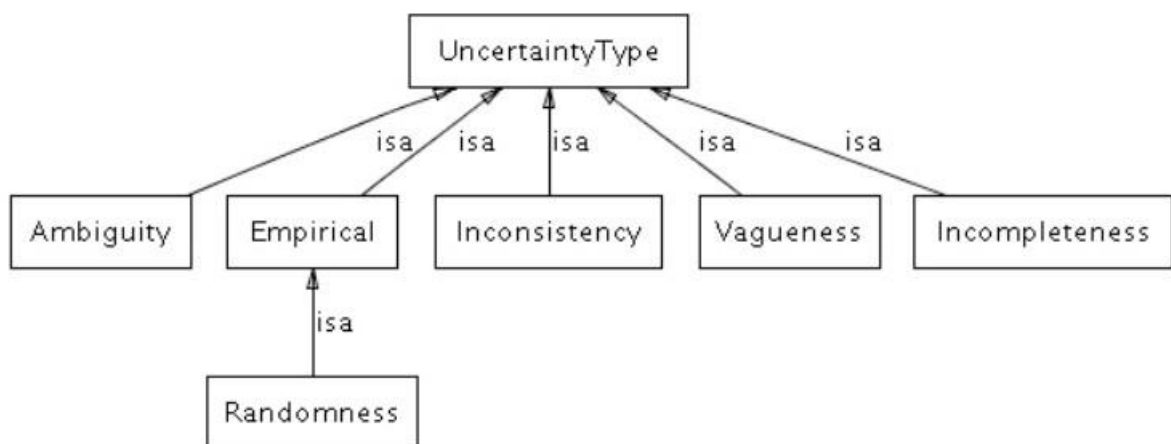


FIGURE 1.1 – Les types d’incertitude [1]

### 1.7.3 L’incertitude en médecine

La pratique de la médecine est rarement simple. L’incertitude peut affecter la prise de décision clinique qui peut avoir des répercussions fondamentales sur la pratique médicale si celle-ci est mauvaise. La prestation de soins optimaux nécessite un équilibre entre l’expertise des cliniciens et les preuves disponibles avec les préférences des patients [28] [29] [30].

En médecine, les sources d’incertitude peuvent provenir de la complexité des informations cliniques (Les données utilisées pour la décision peuvent être contradictoires, ambiguës ou rares.) ; la probabilité de résultats particuliers ; et les caractéristiques individuelles des cliniciens, telles que la tolérance à l’incertitude [31] ou la capacité d’un individu à faire face à la complexité,

au risque et à l'incertitude.[32] Une faible tolérance à l'incertitude a été associée à une utilisation accrue des ressources et à des taux plus élevés d'épuisement professionnel chez les médecins en exercice.

Dans l'enseignement médical, la gestion de l'incertitude et la tolérance à l'incertitude sont souvent enseignées indirectement par l'observation, la modélisation de rôles et des expériences informelles [33] [34]. De nombreux articles ont été consacrés à l'enseignement de la prise de décision et du raisonnement cliniques ainsi qu'à la prise de décision partagée.[30] [35] [36] Cependant, il est important de concevoir un modèle permettant de relier directement les concepts d'incertitude aux ou à un processus décisionnel partagé.

## 1.8 Conclusion

L'internet des objets est devenu un domaine phare dans l'informatique moderne, compte tenu des possibilités diverses qu'offre ce domaine dans différentes disciplines, spécialement dans le domaine médicale, couplé à l'intégration de l'ingénierie ontologique.

Dans ce chapitre, nous avons commencé par définir ce qu'est l'internet des objets. Nous nous sommes intéressés aux divers domaines d'applications de l'internet des objets ainsi que les différents composants qui permettent son fonctionnement, le processus de fonctionnement d'un système IoT ainsi que certains avantages et inconvénients de cette technologie. Ensuite, nous avons défini les ontologies de manière générale, et avons détaillé le rôle de ces dernières et l'intérêt d'utiliser les ontologies dans différents domaines, puis nous avons présenté les différents composants des ontologies ainsi que les différentes méthodes d'ingénierie ontologique, puis, nous avons analysé quelques travaux existants qui traitent des ontologies classiques dans le domaine de l'IoMT, et pour finir, nous avons abordé les connaissances incertaines, en expliquant ce qu'est l'incertitude, en présentant les types d'incertitude, et nous avons parlé de l'incertitude dans le domaine médical.

Dans le chapitre suivant, nous présenterons les formalismes de représentation de connaissances les plus adaptés pour la représentation des connaissances dans l'internet des objets, qui sont les ontologies floues et probabilistes, et analyserons quelques travaux de la littérature.



## *Chapitre 2*

---

# **Approches de modélisation de la connaissance ontologique incertaine dans l'IoMT : Étude comparative**

---

## **2.1 Introduction**

Comme précisé dans le chapitre précédent, les ontologies classiques, qui se basent sur la logique booléenne présentent certaines limites, notamment en ce qui concerne le traitement de l'incertitude et le fait que le comportement binaire propose à la logique booléenne ne s'applique qu'à très peu d'objets ou entités du monde réel, cette problématique est à l'origine de l'apparition de nouvelles formes de logiques capable de traiter cet aspect relatif des objets du monde réel, ce qui nous intéresse dans ce qui suit est la manière de représentation des connaissances floues et probabilistes dans les ontologies.

Dans ce chapitre, nous allons présenter les notions de base sur la logique floue et probabilistes, ainsi que les ontologies floues et probabilistes, et par la même occasion analyser quelques travaux existants afin de mettre en valeur l'apport de ces ontologies dans le domaine de l'internet des objets.

## **2.2 Représentation des connaissances floues dans les ontologies**

La société a radicalement changé au cours des dernières décennies, essentiellement en raison de la possibilité pour pratiquement n'importe qui d'accéder facilement à l'information et à la connaissance. C'est ce qu'on appelle généralement Internet. Cependant, le succès d'Internet provient du fait que les humains ont adopté et utilisé la technologie dans une manière complètement différente que jamais auparavant. Les individus, les groupes et les communautés dépendent

de plus en plus du soutien et de l'aide reçue d'appareils de natures différentes. Le désir et le besoin de nouvelles fonctionnalités ne cesse de croître, en raison à la fois de la curiosité humaine et de la paresse, mais très probablement en raison de l'utilité perçue dans ces avancées. Il en résulte une envie de développer de nouvelles, des méthodes et des technologie révolutionnaires qui peuvent être utilisées par la société.

Pour y parvenir, il est nécessaire d'avoir recours à des solutions de différents domaines de recherche, en particulier les techniques nécessaires pour gérer l'incertitude, l'imprécision et la connaissance tacite, dans ce contexte, les ontologies floues, utilisant la logique floue a connu un essor considérable, ces concepts seront développés plus en détails dans ce qui suit :

### 2.2.1 La logique floue

La théorie des ensembles flous et la logique floue ont originellement été proposées par Zadeh en 1965 [37]. La motivation principale était le fait que les objets du monde réel ont rarement des appartenances clairement définies à des groupes spécifiques. Dans la théorie classique des ensembles, les éléments peuvent exclusivement appartenir à un ensemble ou non, mais avec l'avènement de la théorie floue des ensembles, il devient possible pour ces éléments d'appartenir à des ensembles dans une certaine mesure. La logique floue est une branche de la théorie des ensembles flous étendus, où il représente un système logique capable de traiter l'imprécision et les connaissances incertaines. À tort, la logique floue est parfois utilisée pour désigner toute la théorie des ensembles flous. Il est également important de souligner le fait qu'en mettant en œuvre la logique floue ne signifie pas automatiquement que les résultats finaux seront imprécis ou vagues, on peut produire des réponses aussi précises que l'on le désire [38].

La différence fondamentale entre la logique booléenne et la logique floue est la façon dont le degré d'appartenance à différents groupes est défini et visualisé. La logique booléenne est limitée aux valeurs 0 et 1, ou true ou false, ce qui signifie qu'une instance appartient ou n'appartient pas à un ensemble. Dans le cas de la logique floue, la valeur de l'adhésion d'un élément évolue dans un continuum entre 0 et 1. Cela signifie qu'il peut en même temps appartenir en partie à l'ensemble et en partie ne pas appartenir à ce dernier. Cela donne plus d'options lors de la modélisation de données imprécises, permettant d'indiquer que même si la valeur d'adhésion n'est pas 1, elle a encore quelques similitudes avec l'ensemble.

La **figure 2.1** illustre la différence entre les deux types de logiques :



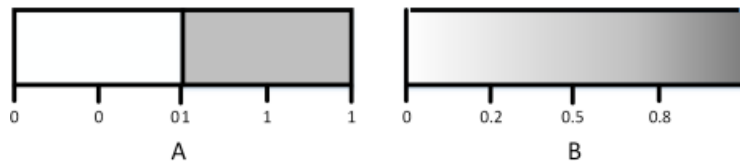


FIGURE 2.1 – (A) logique booléenne (B) logique floue

Cependant, les ontologies classiques ne sont pas appropriées pour représenter des connaissances imprécises et vagues [39] : Pour résoudre ce problème, le concept d’ontologie floue a été introduit, en utilisant les théories sur la logique floue en combinaison avec des ontologies. Avec cette consolidation, il devient possible de modéliser les différents domaines d’études et d’inclure les connaissances tacites et imprécises inhérentes aux domaines en question.

## 2.2.2 Les ontologies floues

Les ontologies floues se basent sur l’intégration de la logique floue à la définition des ontologies classiques pour combler les failles présentes dans les ontologies classiques qui sont principalement l’incertitude et l’imprécision, et ce dans plusieurs domaines, un exemple serait le web sémantique. Il n’existe pas de définition unique de l’ontologie floue : elle varie selon des domaines d’application spécifiques [40]. Bobillo a donc simplifié et généralisé la définition en énonçant qu’une ontologie floue est une ontologie qui utilise la logique floue afin de représenter la connaissance imprécise et faciliter le raisonnement sur celle-ci [40] :

*“An ontology which uses fuzzy logic to provide a natural representation of imprecise and vague knowledge, and eases reasoning over it ”*

Les ontologies floues sont également définies sous forme de quintuple :

$$\mathbf{O.F = (I,C,R,F,A)}$$

où :

**I = ensembles des individus**, dans les ontologies floues, l’appartenance d’un individu ou élément à un composant flou n’est pas sûre, elle est déterminée par une variable probabiliste prenant ses valeurs entre 0 et 1, et déterminée par une fonction d’appartenance [2].

**C = ensembles des concepts flous**, un concept flou est considéré comme un ensemble flou. Les instances de cet ensemble possèdent des degrés d’appartenance qui se situent dans l’intervalle [0,1]. Ces degrés sont calculés grâce aux formules de calcul relatives à la fonction

d'appartenance définie pour le concept [41].

La **figure 2.2** présente deux concepts flous à savoir Employé-Jeune et Employé-Âgé :

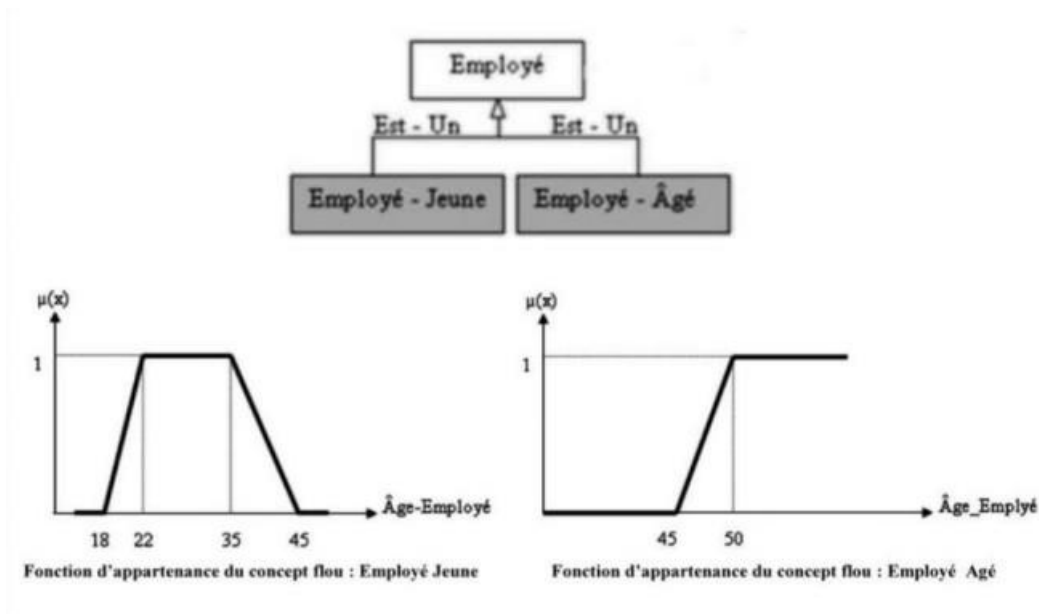


FIGURE 2.2 – Exemple de concepts flous [2]

### R = ensembles des relations

Les relations d'une ontologie désignent les différentes interactions et corrélations entre les concepts de l'ontologie [2].

F = **ensembles des relations floues**, une relation floue peut être considérée comme un ensemble flou. Ses instances sont dotés de degrés d'appartenance avec leurs valeurs dans l'intervalle [0,1]. Ces degrés sont calculés grâce aux formules de calcul de la fonction d'appartenance définie pour la relation [2].

A = **ensembles des axiomes**, en plus de leurs utilisations classiques, les axiomes servent à exprimer les formules de calcul des fonctions d'appartenance des ontologies floues.

Par exemple, les axiomes spécifiques au concept flou « Employé-Jeune » de la figure 2.2 :

- $\text{Emp-Jeune}(\text{Age}) = (\text{Age} - 18) / (22 - 18)$  si Age appartient à [18 ; 22].
- $\text{Emp-Jeune}(\text{Age}) = 1$  si Age appartient à [22 ; 35].
- $\text{Emp-Jeune}(\text{Age}) = (45 - \text{Age}) / (45 - 35)$  si Age appartient à [35 ; 45].
- $\text{Emp-Jeune}(\text{Age}) = 0$  ailleurs. [42]

### 2.2.3 Méthodes de développement des ontologies floues

Jusqu'à présent, les méthodes de développement des ontologies se font encore rares contrairement à celles des ontologies standards. Néanmoins il existe une méthode créée en 2016 par un centre de recherche à l'université de Madrid, elle est nommée "Fuzzy Ontology Development Methodology" (FODM). Cette méthode est construite sur la base de méthodes de développement d'ontologies existantes avec des considérations supplémentaires floues.

Le but du FODM est de guider le développeur de l'ontologie avec des activités qui doivent être réalisées tout au long du processus de développement. Elle vise à présenter la première approche méthodologique pour construire des ontologies floues à partir de zéro, plutôt que de convertir des ontologies existantes en ontologies floues [3].

Le processus de développement de FODM est composé de 11 étapes qui sont structurées comme illustré dans la **figure 2.3** suivante :

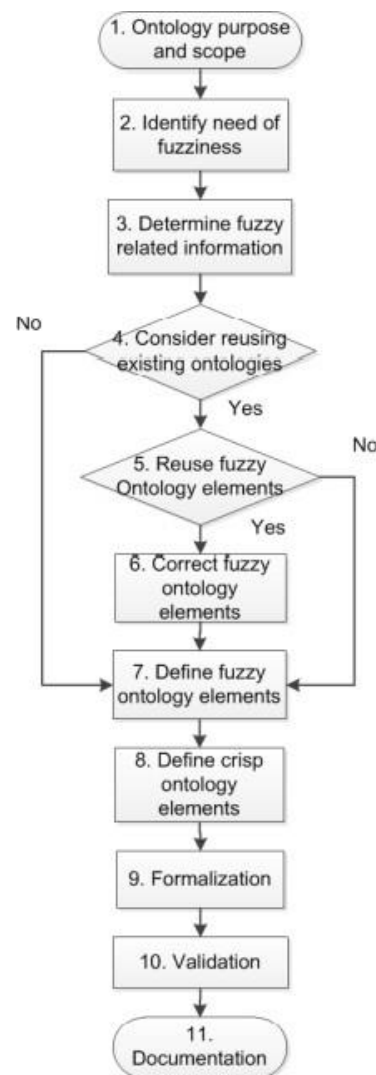


FIGURE 2.3 – Flux de travail du FODM [3]

- **PHASE 1** : Des questions de base doivent être soulevées et traitées explicitement afin de clarifier le but et la portée de l'ontologie.
  - 1) Quel est le domaine ou la portée des informations à modéliser ?
  - 2) L'ontologie est-elle la meilleure technique de modélisation par rapport aux autres solutions ?
  - 3) Quel est le type d'ontologie en fonction de la détermination du domaine ou de la portée?
  - 4) Qui sera impliqué dans le développement de l'ontologie et quels rôles ils vont jouer ?
- **PHASE 2** : La deuxième phase vise à identifier si le flou doit être introduit dans la conception de l'ontologie.
- **PHASE 3** : Suite à l'étape 2, une meilleure compréhension des informations doit être faite dans cette étape, qui vise à établir une distinction entre les informations précises et floues dans le domaine.
- **PHASE 4** : La vérification des ontologies standards existantes pertinentes pour le domaine ou la portée d'intérêt et la détermination de leur réutilisabilité sont les principales tâches définies dans cette phase.
- **PHASE 5** : Similaire à la phase précédente, cette phase vise à vérifier l'existence d'ontologies floues pour réutiliser un ou plusieurs éléments de ces dernières, cependant si aucune ontologie floue n'a été trouvée pour notre domaine, alors cette phase et la phase suivante seront ignorées.
- **PHASE 6** : Si dans la phase précédente, au moins une ontologie floue a été constatée comme utile grâce à ses concepts, on peut remodifier ces concepts pour les adapter à notre ontologie, ceci est le but de cette phase.
- **PHASE 7** : Le but de cette phase est de définir différents éléments d'ontologie floue pour fournir des approximations correctes de la nature des informations vagues et imprécises dans le domaine.
- **PHASE 8** : Cette phase se concentre sur le traitement de certaines connaissances dans le domaine. Outre les informations floues liées, le reste de la base de connaissances du domaine est défini comme des éléments d'ontologie précis en fonction de leurs attributs spécifiques.
- **PHASE 9** : Un certain langage doit être sélectionné pour formaliser l'ontologie conçue dans un format lisible par machine.
- **PHASE 10** : L'ontologie conçue doit passer par une vérification approfondie pour s'assurer qu'elle a représenté le modèle prévu du monde.
- **PHASE 11** : Documentation pour présenter les principes d'ingénierie de l'ontologie conçue, y compris les descriptions des différents éléments de l'ontologie, les détails de conception, la méthode d'utilisation et la maintenance.

## 2.3 Représentation des connaissances probabilistes dans les ontologies

### 2.3.1 La probabilité

L'utilisation du concept de probabilité est indéniable dans la littérature médicale et surtout clinique. Ils sont essentiels à la façon dont les essais cliniques sont compris et interprétés. Ils sont au cœur de la compréhension de comment les risques de santé, les diagnostics, et les pronostics sont communiqués aux patients. Les probabilités sont considérées comme constitutives de l'interprétation rationnelle de tests diagnostiques. Pourtant, les probabilités n'ont jamais eu de définitions spécifiques dans le domaine médical, pour cela il faut se pencher vers les statistiques, les mathématiques et la littérature philosophique pour comprendre ce qu'est exactement la probabilité. Il en existe trois interprétations distinctes de cette dernière : mathématique, subjective et fréquentielle, qui sont présentées ci-dessous : [43]

#### **Interprétation mathématique de la probabilité :**

Dans le début du 20e siècle, le calcul des probabilités a été axiomatisé par Kolmogorov et d'autres, et ce fait la probabilité est devenue une théorie mathématique complète et cohérente. Cette interprétation soutient qu'une probabilité est un entier non négatif qui prend une valeur comprise entre 0 et 1, avec 0 connotant l'impossibilité et 1 la certitude [44].

#### **Interprétation subjective de la probabilité :**

Les probabilités ne sont pas des quantités mesurées d'événements évalués, mais reflètent les croyances subjectives d'une personne sur ce que la probabilité ou la vraisemblance d'un événement peut être [45].

#### **Interprétation fréquentielle de la probabilité :**

L'interprétation fréquentiste de la probabilité admet une variété de constructions différentes. Il existe deux variétés de types de fréquence : fréquence fini et infini. La première est mieux représentée par les problèmes liés à un processus où tout contexte dans lequel il existe une classe bien définie d'événements possibles. La fréquence infinie s'applique aux situations où le jeu de référence est non dénombrable. Dans ce cas, des répétitions infinies ou des essais répétés sont nécessaires pour fixer la valeur de la probabilité. La probabilité est déterminée comme la fréquence limite lorsque la série tend vers l'infini [45].

L'étude des probabilités a connu de nombreux développements au cours des trois derniers siècles en partie grâce à l'étude de l'aspect aléatoire et en partie imprévisible de certains phénomènes, en particulier les jeux de hasard. Ceux-ci ont conduit les mathématiciens à développer des théories qui ont ensuite eu des implications dans des domaines aussi variés que la météorologie,

la finance ou la chimie. L'une de ces théories on y trouve le célèbre théorème de Bayes qui est issu des travaux du révérend Thomas Bayes (1702-1761), présenté à titre posthume en 1763, se basant sur les probabilités conditionnelles. Voici la formule de cette règle : [46]

$$P(A/B)P(B) = P(A \cap B) = P(B/A)P(A)$$

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)}$$

$P(A | B)$  est la probabilité conditionnelle de A sachant B qui signifie (on cherche la probabilité de A quand B est réalisé),  $P(A \cap B)$  est la probabilité de la réalisation des événements A et B en même temps.

Ce théorème sert à déterminer des probabilités fréquentistes d'un échantillon faisant parti d'un ensemble. Encore plus tard, les chercheurs ont voulu exploiter d'avantage ce théorème pour concevoir un modèle de représentation de connaissances pour une finalité d'un outil d'aide à la décision, ce qui amena à l'invention des réseaux bayésiens [47].

## 2.3.2 Les réseaux bayésiens

### 2.3.2.1 Définition et structure :

Un réseau bayésien est un système représentant la connaissance et permettant de calculer des probabilités conditionnelles pour répondre à différentes sortes de problématiques généralement à des fins de prise de décision. Également appelé réseaux de croyances (ou réseaux bayésiens en abrégé), appartiennent à la famille des modèles graphiques probabilistes (GM) [47].

Un réseau bayésien possède une structure simple :

- un graphe acyclique orienté** : C'est un réseau qui ne contient pas de boucles, il contient en effet des nœuds reliés par des arcs dirigés [48].
- Des nœuds** : Les nœuds représentent des variables aléatoires et sont tirés sous forme de cercles étiquetés par les noms de variables [48].
- Des arcs** : Expriment les probabilités conditionnelles entre les variables aléatoires [48].
- Table de probabilités** : Les tables de probabilités sont définies par des statistiques relatives au problème à résoudre (peuvent aussi être déterminées par des experts). Chacune des variables dispose d'une table de probabilités conditionnelles relatives aux variables causales dont elle dépend [49].

La **figure 2.4** montre un exemple d'un réseau bayésien :

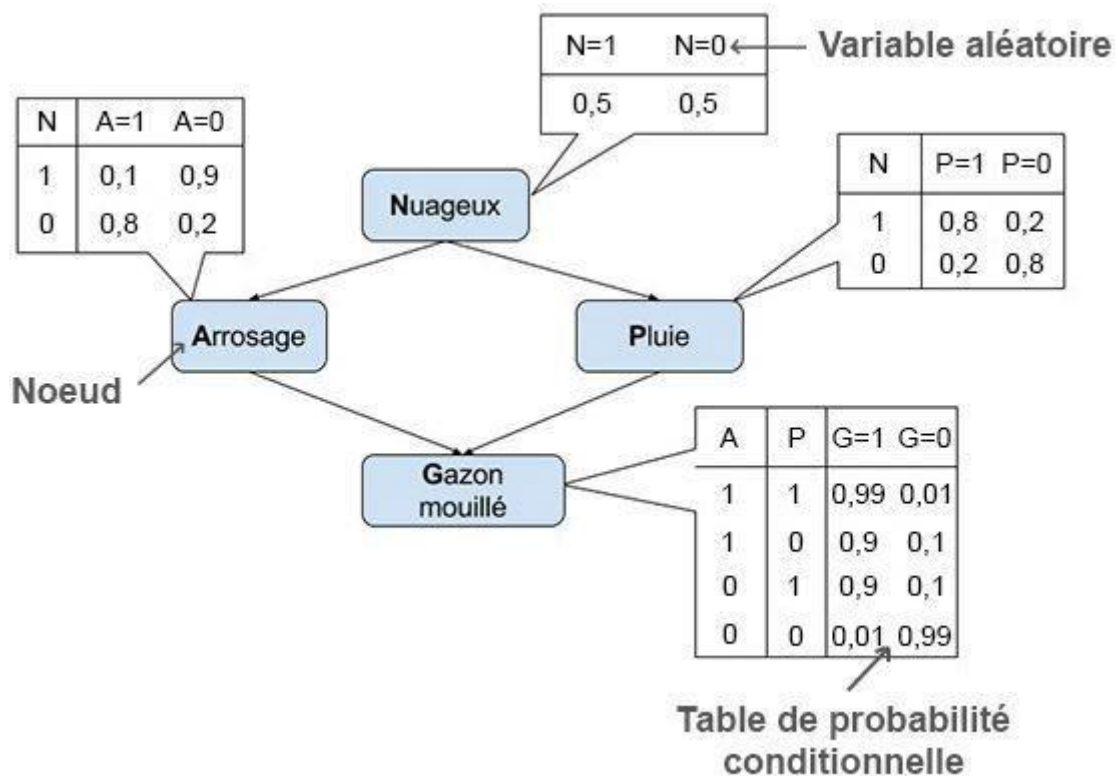


FIGURE 2.4 – Un réseau bayésien simple contenant 4 variables aléatoires

### 2.3.2.2 L'apprentissage :

Comme vu précédemment, un réseau bayésien est à la fois un graphe (description qualitative) et un ensemble de probabilités conditionnelles (description quantitative), l'apprentissage d'un réseau bayésien doit donc répondre aux deux questions suivantes :

- Comment estimer les lois de probabilités conditionnelles ?
- Comment trouver la structure du réseau bayésien ?

Donc l'apprentissage des réseaux bayésiens se compose de deux parties :

- L'apprentissage de la structure**, où le but est de trouver le meilleur graphe représentant la tâche à résoudre.
- L'apprentissage des paramètres**, où nous supposons que la structure du réseau est fixe, et il faudra estimer les probabilités conditionnelles de chaque nœud du réseau [50].

### 2.3.2.3 L'apprentissage de la structure :

L'apprentissage de structure d'un réseau bayésien consiste à identifier les nœuds et les relations possibles entre ces nœuds partir des données d'apprentissage. La recherche de structure des réseaux bayésiens est un problème difficile, principalement à cause du fait que l'espace de recherche est exponentiel en fonction du nombre de variables décrivant le domaine.

Il existe 2 façons pour s'y prendre :

**La construction d'un réseau bayésien en collaboration avec les experts du domaine :** La construction d'un RB en utilisant les connaissances des experts du domaine est une tâche difficile et chronophage. Elle est menée par un ensemble de spécialistes. Un expert en RB dirige le processus de construction, il pose des questions pertinentes, et explique les suppositions encodées dans le modèle au reste du groupe des spécialistes afin de fixer des relations cause à effet entre les noeuds.

**La construction d'un réseau bayésien à partir de données :** La deuxième façon est d'utiliser des algorithmes pour l'apprentissage automatique de la structure comme l'algorithme PC de recherche de causalité[49].

#### **2.3.2.4 L'apprentissage des paramètres :**

L'apprentissage des paramètres d'un réseau bayésien consiste à associer une table de probabilités locales à chaque nœud de la structure du réseau préalablement élaborée ou apprise.

Parmi les méthodes d'apprentissage de paramètres qui sont classées selon la nature des données (données complètes ou incomplètes), on trouve des méthodes d'apprentissage à partir de données complètes (données observées) comme la méthode de maximum de vraisemblance qui se base sur la fréquence d'apparition d'un événement pour calculer sa probabilité, et des méthodes d'apprentissage à partir de données incomplètes (c'est à dire les variables sont complètement manquantes ou ne sont observées que partiellement) comme l'algorithme EM (Esperance-Maximisation) proposé par Dempster et al. 1977[51].

#### **2.3.2.5 L'inférence dans un réseau bayésien :**

L'utilisation essentielle des réseaux bayésiens est donc de calculer des probabilités conditionnelles d'évènements reliés les uns aux autres par des relations de cause à effet. Cette utilisation s'appelle inférence.

Dans un réseau bayésien, l'inférence consiste à y propager une ou plusieurs dépendances causales certaines au sein de ce réseau, pour en déduire comment sont modifiées les probabilités concernant les autres nœuds.

Formellement, si il y a deux (ou plus) sous-ensembles de variables aléatoires nommés par exemple A et B où : A regroupe les variables de question ou cachées qui représentent donc le problème de prédiction ou de diagnostic à résoudre et B qui constitue les variables d'évidence ou les évènements observés, alors l'inférence se réduira au calcul de la probabilité  $P(A/B)$ . Grâce à la loi de Bayes, la factorisation et la marginalisation des probabilités, on peut calculer toutes les probabilités d'intérêt[50]. Pour ce qui est de l'application de l'inférence, il y a deux approches : une inférence exacte, et une inférence approximative. La première approche visant à calculer la distribution de probabilité a posteriori d'un ensemble de variables de requêtes, compte tenu d'un événement observé, c'est-à-dire de certaines affectations de valeurs à des variables de preuve, pour ce type d'approche on y trouve deux types d'inférences : par énumération, par élimination de variable.

Concernant la deuxième approche (approximative), pour certains réseaux bayésiens qui contiennent



des nœuds discrets et des nœuds continus, il n'existe pas d'algorithme d'inférence exacts. Le seul recours pour évaluer des probabilités est alors d'utiliser des algorithmes d'inférence approchés. Une parmi les méthodes proposées est la méthode de propagation cyclique MCMC (pour Markov Chains Monte Carlo) permettant d'échantillonner des variables aléatoires en construisant une chaîne de Markov [52].

### 2.3.2.6 Les réseaux bayésiens flous :

Le réseau bayésien flou est une extension du réseau bayésien standard car il se combine avec la théorie des ensembles flous en gardant une architecture similaire à celle d'un réseau bayésien classique. L'inférence dans les réseaux bayésiens flous est basée sur un raisonnement probabiliste flou. Lorsqu'un nœud est flou, plusieurs états peuvent être observés à la fois avec des degrés d'appartenances différents mais la somme de ces derniers doit être égale à 1. Il existe plusieurs méthodes pour faire une inférence probabiliste floue, nous citons les principales méthodes qui sont : méthode de pondération, méthode de l'évidence virtuelle et la méthode de distribution probabiliste floue.

L'évidence virtuelle d'une variable X d'un réseau bayésien est une observation avec incertitude de la variable.

$$L(X) = (L(X = x_1) : \dots : L(X = x_m)) = (P(\text{obs} | x_1) : \dots : P(\text{obs} | x_m))$$

où les  $L(X = x_i)$  sont des quantités relatives les unes aux autres représentant la probabilité de l'événement observé étant donné que X est dans l'état  $x_i$ .

La méthode de l'évidence virtuelle consiste donc à ajouter pour chaque nœud flou dans le RB, un nœud fils appelé « nœud virtuel » qui sert à stocker les observations incertaines ou floues. En effet, les degrés d'appartenances seront mis dans la CPT (table de probabilités conditionnelles) du nœud virtuel et seront incorporées avec les probabilités conditionnelles comme des poids lors de l'inférence bayésienne en appliquant un algorithme d'apprentissage standard [53].

Nous présentons ci-dessous un exemple de l'application de l'évidence virtuelle :

Dans cet exemple, on modélisera une relation de causalité entre le diabète (D) et la glycémie (GLYC), la glycémie est une condition du diabète, ce qui veut dire que le nœud D dépend du nœud GLYC.

Le graphe RBF qui correspond à cette situation est illustré dans la **figure 2.4** :

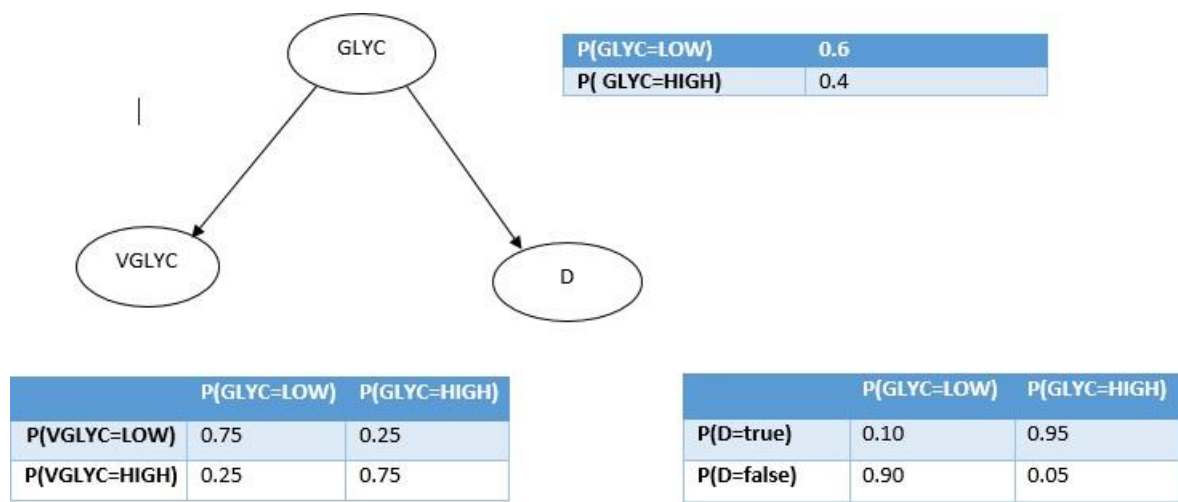


FIGURE 2.5 – Un exemple de l'application de l'évidence virtuelle

La variable GLYC est considérée comme floue, donc ses états vont être fuzzifiés, et auront chacun une fonction d'appartenance. Le tableau suivant illustre cette étape.

Etat	Interval	Type et paramètres de la fonction
LowGlyc	[0.7;1.2]	Trapezoid("low",0.7,0.7,0.9,1.3)
HighGlyc	[1.4;1.8]	Trapezoid("high",0.9,1.3,1.8,1.8)

TABLE 2.1 – Tableau des fonctions d'appartenance de la variable GLYC

On suppose que la valeur de glycémie d'un patient P1 soit mesurée à 1.2 g/l, après le calcul des valeurs d'appartenances à chaque sous-ensemble flou résultent les degrés suivants :

$\mu_{low}(1.2) = 0.25$   
 $\mu_{high}(1.2) = 0.75$

Supposons qu'on veuille calculer la requête suivante :

P(D = true | GLYC) : Cette requête représente la probabilité que le patient P1 soit diabétique sachant que la glycémie est "low" avec un degré de 0.25 et "high" avec un degré de 0.75.

### 2.3.3 Ontologies probabilistes

Jusqu'à présent, il n'y a pas encore de langages pour la représentation d'ontologies probabilistes qui soit standardisée par le W3C. Par contre, certains travaux (p.ex., BayesOWL [54] et PR-OWL) ont proposé des solutions assez pertinentes ; le travail le plus documenté, qui a déjà un noyau de projet largement utilisé, est PR-OWL qui est déjà à sa version 2.0 compatible avec OWL [55].

P. Costa, un des premiers instigateurs du projet PR-OWL, définit une ontologie probabiliste comme étant une représentation formelle et explicite de connaissances concernant un domaine

d'une application associée à l'incertitude, ce type d'ontologies comporte :

- a. Types d'entités existant dans le domaine.
- b. Propriétés de ces entités.
- c. Relations entre entités.
- d. Les processus et les événements qui se produisent avec ces entités.
- e. Les régularités statistiques qui caractérisent le domaine.
- f. Une définition pour des connaissances incomplètes, peu fiables, ambiguës et peu concluantes.
- g. L'incertitude concernant toutes les formes de connaissances précitées [55].

La plupart de ces propriétés sont présentes dans une ontologie classique comme nous pouvons le remarquer. Mais les ontologies probabilistes vont au-delà de la simple annotation des ontologies avec des probabilités pour fournir un moyen d'exprimer les caractéristiques subtiles requises pour exprimer une théorie bayésienne de premier ordre.

Les langages d'ontologies probabilistes se font encore rares, cependant on peut citer PR-OWL qui a été développé comme une extension permettant aux ontologies OWL de représenter des modèles probabilistes bayésiens complexes d'une manière suffisamment flexible pour être utilisée par divers outils probabilistes bayésiens (par exemple, PRM, BN, etc.). Plus précisément, PR-OWL est une ontologie supérieure pour les systèmes probabilistes, il est suffisamment expressif pour représenter même les modèles probabilistes les plus complexes [56].

Un autre moyen existant pour les ontologies probabilistes on trouve aussi Bayes-OWL et PR-OWL.

## 2.4 Analyse des travaux existants et discussion

Dans ce qui va suivre, nous allons analyser quelques articles de la littérature qui traitent sur le domaine de l'internet des objets couplés avec les ontologies pour développer différents systèmes qui servent plusieurs causes, comme vu précédemment, les ontologies classiques présentent certaines limites, notamment dans ce qui est de traiter l'incertitude, l'objectif de ce travail est de mettre en exergue l'apport des ontologies floues et probabilistes dans le domaine de l'internet des objets et des avantages de ces dernières dans la représentation des différentes connaissances issues des données récoltées par les différents composants des systèmes de l'internet des objets.

### 2.4.1 Travail de : "N.Selvan et al. , 2018"

Dans cet étude, l'Internet des objets (IoT) est exploité dans un contexte medical (IomT) pour créer des systèmes de surveillance pour l'état des patients chroniques tout en minimisant les efforts manuels déployés pour les recommandations et soins. Cependant, de nombreux facteurs incertains sont impliqués chez ces patients, et les modèles existants sont incapables de les gérer efficacement conduisant souvent à de mauvais résultats parfois intolérables. Les travaux existants

rencontrent d'autres problèmes importants comme la croissance rapide du nombre de dossiers des patients, ce qui rend leur gestion et leur contrôle par les systèmes de soins et de santé difficile. Aussi, nous citons que les dossiers médicaux recueillis à partir de différents appareils sont dans différents formats ce qui conduit à une hétérogénéité de données qui empêche l'utilisation de ces dernières pour des analyses ultérieures.

Le but de ce travail est de proposer un système de recommandation basé sur une ontologie floue en utilisant la logique floue de type 2 pour recommander des aliments et des médicaments à un patient chronique (diabétique). La logique floue permettra de résoudre le problème des connaissances incertaines chez les patients, pour les dossiers des patients, les clouds seront utilisés pour le stockage des données. Pour homogénéiser les dossiers, l'extraction des facteurs de risques pour les patients chroniques est réalisée via des capteurs portables et les dossiers médicaux électroniques des patients basés sur l'IoT sont liés à des données ouvertes liées (LOD qui sont des BD avec des licences libres ouvertes) pour créer une base de connaissances. [57]

#### **2.4.2 Travail de : "R. Gorrepati et al. , 2021"**

Dans cette étude, nous utilisons l'ontologie pour décrire la représentation sémantique des objets médicaux et leurs données dans la sauvegarde U-IoMT dans cette analyse. Le système de santé vise à examiner les ordonnances des patients et les dossiers de gestion de la chaîne d'approvisionnement en médicaments. Les services de soins de santé d'urgence sont connectés aux appareils portables des patients sur smartphone à des fins de surveillance afin de réduire les cas d'urgence et de conserver les dossiers électroniques des patients. L'application de santé mobile vise à maintenir l'état de santé d'un patient où qu'il se trouve. L'architecture des réseaux de capteurs sémantiques (SSN) utilise la communication peer-to-peer pour réaliser l'interopérabilité sémantique, qui est décrite comme des plates-formes IoMT interopérables. Un cadre de SSN et la couche de sensibilisation au contexte ont également été créés pour la visualisation de la surveillance de la santé à distance des patients, la gestion des médicaments, l'analyse du moment du patient et le suivi des patients ont été surveillés en utilisant divers appareils, la santé en ligne a été utilisée pour démontrer les maladies chroniques [58].

#### **2.4.3 Travail de : "M.Lesani et al. , 2019"**

Le but de cette étude est de concevoir une ontologie floue pour pour les réseaux de capteurs sémantiques (SSN) pour les maisons intelligentes dans le contexte de l'internet des objets en deux phases. Dans la première phase, à l'aide de l'ontologie WordNet, l'emplacement et le type d'un objet sont identifiés à l'aide d'une interface graphique. Cet objet et ses synonymes sont ajoutés à la liste des objets connus définis. Ensuite, la relation de l'objet avec d'autres groupes est évaluée sur la base d'une mesure de similitude en plus d'utiliser l'ontologie floue. Dans la deuxième phase, les capteurs avec des informations erronées sont identifiés en trouvant une

relation entre certains facteurs spécifiques. À cette fin, la température, l'humidité et la lumière sont prises en compte et le système d'inférence neuronal (ANFIS) est incorporé. [59]

#### **2.4.4 Travail de : "Y.Benazzouz et al. , 2014"**

Le but de cette étude est de proposer une ontologie qui sera utilisée via une architecture de service web pour diagnostiquer et tester des réseaux à capteurs sans fil (WSN) à grande échelle afin qu'ils soient plus tolérants aux pannes. Pour y parvenir, on utilisera une ontologie floue de représentation des capteurs proposée déjà existante et l'améliorant et en y intégrant les domaines de fautes sur les WSN c'est à dire dans le contexte de l'IoT réseaux, en plus de pouvoir supporter l'hétérogénéité des appareils et les réseaux et les types d'erreurs. Cette ontologie finalisée sera intégrée dans une architecture qui récoltera des connaissances nécessaires pour évaluer un large nombre de systèmes distribués et non-familiers. Donc comme ça, on n'aura pas besoin de prendre en compte l'infrastructure interne du réseau particulier. [60]

#### **2.4.5 Travail de : "G.Lee et al. , 2016"**

le but de cette étude est de développer une plateforme qui coopère avec des applications et services pour évaluer le niveau de confiance entre des entités dans le contexte du S-IoT(social internet of things), en utilisant un modèle de confiance et en intégrant des propriétés pour les objets comme la réputation (qui sera tirée des opinions d'autres entités), la recommandation (qui représente l'avis d'une entité de vouloir faire confiance ou pas à une autre entité), et la connaissance (les données partagées publiquement sur le réseau concernant un objet), ce système est modélisé et conçu en utilisant les ontologies floues. [61]

#### **2.4.6 Travail de : "S.Chabridon et al. , 2017"**

le but de cette étude est de proposer une extension aux travaux existants dans le cadre de l'identification de situations qui appartient au domaine de l'IoT, en y intégrant le méta-modèle QoCIM (Quality Of Context Information Model) , et en utilisant le framework muSIC dédié à l'identification de situations, afin de modéliser la qualité d'un contexte ainsi que la qualité d'une situation en même temps, cette recherche est basée sur l'utilisation d'ontologies floues qui permettent d'effectuer des raisonnements complexes et manipuler des données sous forme numérique ou sous forme de texte.

#### **2.4.7 Travail de : "L.Varghese et al. , 2021"**

Le but de cet étude est de proposer une plateforme interopérable grâce à l'internet des objets dans une maison intelligente afin de contrôler les appareils électroménagers de n'importe où en offrant une opérabilité à distance et réduisant ainsi l'intervention humaine,les données collectées sont fournies à des programmes tiers si nécessaire, pour que les tiers puissent créer

des applications et des services sur mesure qui répondent aux besoins du consommateur. Une ontologie floue est utilisée pour la vérification et l'analyse en temps réel des données collectées à partir des capteurs intégrés dans le système pour la prise correcte de décision par le biais de l'intelligence artificielle, cela permettra au travail effectué de gérer les ressources comme l'électricité, l'eau, la nourriture etc de façon économique et également prévoir les incidents potentiels et les éviter, tels que les incendies, le vol, etc. Le système a montré une précision d'environ 96%, ce qui est très bon. [62]

#### **2.4.8 Travail de : "S.Titi et al. , 2020"**

Le but de cette étude est de concevoir un système de surveillance et d'évaluation des patients diabétiques en utilisant le concept de l'IoT pour permettre d'avoir de meilleures prises de décisions sur les soins de santé. La majorité des systèmes de surveillance de la santé basés sur l'IoT adoptent des ontologies pour représenter et interagir avec les données collectées. Les ontologies classiques ne peuvent pas traiter de manière appropriée les connaissances imprécises et ambiguës. Les ontologies floues peuvent résoudre efficacement les problèmes de connaissances incertaines. Ce système se compose de trois couches : la couche de données, le réseau, et la gestion. La couche de données fait référence aux données spécifiques au patient (son profil, mode de vie, régime alimentaire, exercices physiques, traitement) collectées par différents capteurs médicaux. Ensuite ces données sont transférées au serveur en utilisant les technologies de l'internet des objets qui constituent la couche réseau. En dernier lieu, la couche gestion est dédiée aux services et à la manipulation des données qui seront considérées comme des entrées pour pouvoir les interpréter et aider l'utilisateur dans sa prise de décision. Ce système a été évalué et a montré des résultats plus précis qu'un système utilisant une ontologie classique. [63]

#### **2.4.9 Travail de : "G.Lunardi et al. , 2018"**

Le but de cette étude est de concevoir un système d'assistance à l'autonomie à domicile, basé sur la technologie AAL ( Active and Assisted Living) qui est un domaine de l'IoT, ce système est une extension du modèle HBMS (Human Behavior Monitoring and Support), qui fournit un modèle conceptuel représentant le comportement d'un utilisateur dans un environnement, grâce à des ontologies probabilistes qui représentent des données collectées de sources IoT différentes et qui permettent de faire des raisonnements sur des situations incertaines, comme décider laquelle des prochaines actions possibles est la mieux adaptée à l'utilisateur dans un contexte donné, La performance de cette approche a été évaluée à l'aide de données réelles collectées à partir d'un prototype de maison intelligente équipé de capteurs installables et d'appareils IoT. Les expériences ont fourni des résultats prometteurs avec un taux de précision dépassant les 70% et ce avec une petite quantité de données. [64]

## 2.4.10 Tableau comparatif

Afin d'avoir une vue globale sur les travaux réalisés précédemment, quelques critères ont été sélectionnés pour comparer ces derniers dont :

- Le type d'ontologies utilisé exprimant les types de connaissances.
- Le domaine d'application et l'objectif de l'étude.
- La méthode d'ingénierie qui constitue la base de construction de notre ontologie.
- Évaluation : Pour finir sur les résultats démontrés d'un travail si celui-ci a été évalué.

Cette comparaison est illustrée dans le **tableau 2.2** ci-dessus :

Critères Travaux	Types d'ontologies	Domaine d'application et objectif	Méthode d'ingénierie de l'ontologie	Évaluation
N.Selvan et al. , 2018	Floue	Services de santé, ce système propose aux patients atteints de diabète chronique des aliments et médicaments.	Création d'une nouvelle ontologie	oui
R. Gorrepati et al. , 2021	Floue	Services de santé, l'objectif du travail est de proposer une plateforme pour, l'analyse des données des patients et leurs suivis même à distance, et la gestion des traitements.	Création d'une nouvelle ontologie	oui
M.Lesani et al. , 2019	Floue	Maisons intelligentes, le but du système est de proposer une ontologie pour les réseaux de capteurs sémantiques dans les maisons intelligentes	Création d'une nouvelle ontologie	oui
Y.Benazzouz et al. , 2014	Floue	Test et diagnostics de réseaux IoT, le but du système proposé est de pouvoir faire des diagnostics et tests dans les réseaux à capteurs sans-fil.	Création d'une nouvelle ontologie	non
G.Lee et al. , 2016	Floue	S-IoT (trust service platform), la plateforme proposée offre une évaluation du niveau de confiance entre deux entités dans le contexte de l'internet des objets sociaux.	Création d'une nouvelle ontologie	non
S.Chabridon et al. , 2017	Floue	L'identification de situations (IoT), ce travail propose une extension aux travaux existants dans le contexte de L'identification de situations dans l'internet des objets.	Création d'une nouvelle ontologie	non
L.Varghese et al. , 2021	Floue	Maisons intelligentes, la plateforme proposée permet aux propriétaires d'un logement de le monitorer de n'importe où dans le monde.	Création d'une nouvelle ontologie	non
S.Titi et al. , 2020	Floue	Services de santé, le but du système proposé est de permettre le monitoring des patients atteints du diabète.	Création d'une nouvelle ontologie	oui
G.Lunardi et al. , 2018	probabiliste	vie active et assistée (AAL), le but est de proposer une extension aux travaux existants visant à aider les personnes âgées souffrants de troubles de la mémoire dans leur vie quotidienne.	Utilisation d'une ontologie existante en lui ajoutant des extensions.	oui

TABLE 2.2 – Tableau comparatif

## 2.4.11 Discussion

À la lumière de l'analyse effectuée sur les divers travaux existants de la littérature énoncés dans la figures 2.1 , nous pouvons remarquer que :

- Les ontologies floues sont les plus souvent utilisées comme c'est le cas des systèmes "N.Selvan et al., 2018", "R. Gorrepati et al. , 2021", "M.Lesani et al. , 2019", "Y.Benazzouz et al. , 2014", "G.Lee et al. , 2016", "S.Chabridon et al. , 2017", "L.Varghese et al. , 2021", "S.Titi et al. , 2020", et pour les ontologies probabilistes, elles ne sont utilisées que par "G.Lunardi et al. , 2018".
- Les systèmes ci-dessus traitent de plusieurs domaines différents tels que les maisons intelligentes, l'identification de situation dans le contexte de l'IoT, les tests et diagnostics des réseaux IoT...etc, on remarque cependant que seulement trois systèmes ("N.Selvan et al. , 2018", "S.Titi et al. , 2020", et celui de "R. Gorrepati et al. , 2021") traitent du domaine médical.
- Quant à la méthode d'ingénierie de l'ontologie, tous les systèmes ont opté pour la création de leur propre ontologie sans utiliser d'ontologie déjà existante, sauf pour "G.Lunardi et al., 2018" qui ont exploité une ontologie existante en y ajoutant des extensions.
- La plupart des systèmes ont été évalués et ont démontré de bons résultats, sauf pour les systèmes ("Y.Benazzouz et al. , 2014", "G.Lee et al. , 2016", "S.Chabridon et al. , 2017", "L.Varghese et al. , 2021").

Comme nous l'avons indiqué, plusieurs travaux dans le cadre de l'IoMT ont été développés comme les travaux de "N. Senthil Selvan et al., 2018" et "Sondes Titi et al. , 2020". Ces systèmes sont similaires dans certaines caractéristiques. Malgré cela, leur étude nous a permis d'identifier un ensemble de limites :

- Traitement d'un seul type d'incertitude seulement pour l'IoMT qui est le flou.
- Absence de la notion de probabilité dans les travaux de l'IoMT.

## 2.5 Conclusion

Dans ce chapitre nous avons commencé par présenter une forme d'incertitude qui est la logique floue qui sert à résoudre le problème d'incertitude concernant l'appartenance d'une instance à un concept particulier, et puis son intégration dans les ontologies pour en faire des ontologies floues. Ces dernières ont une structure un peu différente des ontologies normales quant aux nouveaux concepts flous apportant une solution à l'imprécision.

Ensuite, nous avons introduit et vu les probabilités qui est une autre forme d'incertitude aussi importante que la précédente en rappelant aussi le célèbre théorème de Bayes, ensuite nous avons vu les réseaux bayésiens : en quoi ils consistent, leur structure, l'inférence dans ces derniers qui constitue la base du calcul des probabilités, et aussi nous avons brièvement parlé



de l'apprentissage pour les réseaux bayésiens. Comme pour la logique floue, l'utilisation des probabilités à été combinée aux ontologies ce qui donne des ontologies probabilistes, nous avons également présenté deux outils utilisés pour les ontologies probabilistes qui sont PR-OWL et Bayes-OWL.

Pour conclure ce chapitre, nous avons mentionné quelques applications des ontologies floues et probabilistes dans différents domaines. Dans le chapitre suivant, nous allons présenter la conception de notre système de modélisation en choisissant l'un des types d'ontologies afin de réaliser notre modélisation de connaissances dans le contexte de l'IoMT.



## Chapitre 3

---

# Conception du système

---

### 3.1 Introduction

Après avoir vu les principaux domaines de notre système en commençant par l'internet des objets, les ontologies floues, et les réseaux bayésiens, maintenant il est temps de passer à la conception de notre système, c'est à dire décider de comment employer ces derniers pour parvenir à notre but.

Nous allons présenter dans ce chapitre la motivation du travail, ainsi que notre solution proposée, la conception du système présentée avec un schéma global montrent les différentes phases de conception.

Tout au long de ce chapitre, nous allons présenter notre système, et détailler la solution proposée pour ce système de diagnostic médical dans le contexte de l'IoMT, qui vise essentiellement à tirer profit des avantages offerts par l'internet des objets, les ontologies floues et des réseaux bayésiens flous cités ci-dessus pour faire des inférences sur les patients. Nous commencerons par la présentation des données utilisées ainsi que les grandes étapes de conception comprenant le développement de l'ontologie floue, la construction du réseau bayésien flou (RBF), et enfin, l'algorithme appliqué pour le diagnostic de l'état du patient en se basant sur le RBF et l'ontologie floue.

### 3.2 Motivation derrière le travail

En Algérie, le diabète reste une réalité préoccupante, étant la deuxième maladie chronique après l'hypertension. Le nombre de diabétiques en Algérie est passé d'un million de personnes en 1993, à plus de 2 500 000 en 2007, soit 10% de la population en 2010. À l'échelle mondiale, le nombre de diabétiques devrait augmenter de 130% au cours des 25 prochaines années[65].

Le diabète de type 2 (diabète non insulino-dépendant) est la forme la plus répandue des diabètes en représentant plus de 90% des diabétiques[66], il s'agit d'une forme multifactorielle. Par conséquent, le choix du diabète type 2 comme porté de domaine d'étude peut être justifié.

Et entraîne des complications microvasculaires et macrovasculaires qui engendrent par la suite une anomalie psychologique et physique chez les personnes atteintes.

Malgré l'augmentation des connaissances sur les facteurs de risque du diabète de type 2 et les preuves de la réussite des programmes de prévention, l'incidence et la prévalence de la maladie continuent d'augmenter à l'échelle mondiale. La détection précoce grâce à des programmes de dépistage et la disponibilité de thérapies sûres et efficaces réduisent la morbidité et la mortalité en prévenant ou en retardant les complications.

Comme vu précédemment dans **la section 2.4**, la majorité des systèmes d'aide au diagnostic médical en IoMT ne sont conçus que pour faire face qu'à un seul type d'incertitude (le flou). Cependant, nous pouvons développer une meilleure approche pour prendre en compte différentes incertitudes qui seront le flou et l'aléatoire (probabilités). Or une ontologie ne peut être floue et probabiliste en même temps, donc le choix le plus judicieux est d'utiliser une ontologie floue en la combinant à un réseau bayésien pour remplacer l'extension probabiliste de l'ontologie, à noter que le RB (réseau bayésien) doit être flou pour être adapté aux connaissances de l'ontologie.

### 3.3 Étapes de construction du système

Notre objectif consiste à développer un système d'aide au diagnostic médical dans le contexte de l'IoMT pour faire le diagnostic sur la maladie du diabète de type 2 fondé sur un modèle d'ontologie floue et un moteur d'inférence basé principalement sur l'inférence fournie par le réseau bayésien flou. Les différentes parties composant notre système sont les suivantes :

**-Préparation des données :** Dans cette phase on doit trouver des données pertinentes qui peuvent être utilisées en tant que données de patients, le dataset et les données qu'il contient sont décrits dans la **section 3.4**.

**-Construction de l'ontologie floue :** Servira comme base de connaissances pour les données du patient relatives au diabète, des informations médicales de l'IoMT impliquées. Dans cette section, la conception et la création de cette ontologie floue seront détaillées.

**-Construction du réseau bayésien flou :** Contendra le moteur d'inférence qui permettra de diagnostiquer un patient selon ses informations stockées dans l'ontologie floue.

**-Couche application :** Les informations du patient seront capturées grâce à des capteurs connectés, et seront transférées à l'application qui affichera le résultat de l'inférence sur son interface.

Le détail de chaque module sera représenté dans les sections suivantes, un schéma récapitulant ces différents points est illustré dans la **figure 3.1** :

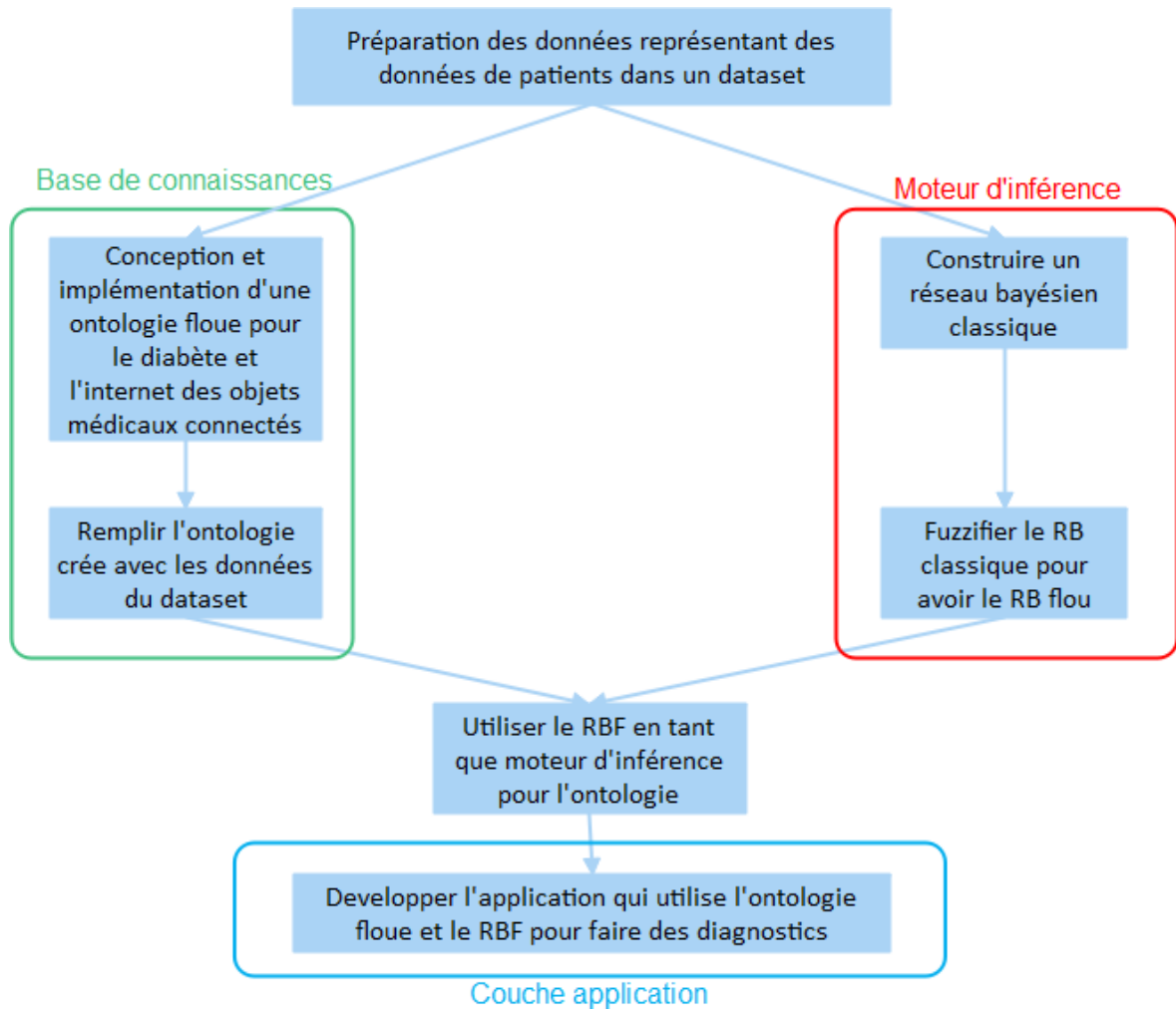


FIGURE 3.1 – Processus de développement du système

### 3.4 Préparation des données

Nous représentons ci-dessous le dataset utilisé, issu de "L'Institut national du diabète et des maladies digestives et rénales".

La figure 3.2 représente un fragment du dataset utilisé :

	patient_number	age	gender	cholesterol	glucose	systolic_bp	diastolic_bp	height	weight	bmi	waist	hip	diabetes
0	1	19	female	193	77	118	70	154.94	53.55	22.306548	32	38	No diabetes
1	2	19	female	146	79	108	58	152.40	60.75	26.156302	33	40	No diabetes
2	3	20	female	217	75	110	72	170.18	84.15	29.056084	40	45	No diabetes
3	4	20	female	226	97	122	64	162.56	51.30	19.412881	31	39	No diabetes
4	5	20	female	164	91	122	86	177.80	63.45	20.070959	32	39	No diabetes

FIGURE 3.2 – Dataset utilisé

Ce dataset contient 12 colonnes, dont la 12ème colonne est celle du diagnostic, les autres colonnes représentent les facteurs de risques qui seront utilisés dans notre étude :

- Le cholestérol** : Représente le taux de cholestérol total sérique dans le sang, exprimée en cg/l (centigramme/litre) comme unité de mesure utilisée.
- Glucose** : Le taux du sucre dans le sang, exprimée en cg/l.
- Âge** : L'âge du patient en années.
- Genre** : Si le patient est un homme ('male') ou une femme ('female').
- La taille** : La taille d'une personne en centimètre.
- Le poids** : Le poids d'une personne en kilogramme.
- L'indice de masse corporelle** : L'IMC permet de déterminer si l'on est situation de maigre, de surpoids ou d'obésité grâce à la taille et au poids.
- Tension artérielle systolique** : Est la pression du sang lorsque le cœur se contracte, en mmHg (millimètre de mercure).
- Tension artérielle diastolique** : Est la pression du sang dans les artères quand le cœur se relâche, en mmHg.
- le tour de taille** : en cm.
- le tour de hanche** : en cm.

La **figure 3.3** affiche les informations du dataset après une analyse :

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 390 entries, 0 to 389
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   patient_number        390 non-null    int64
1   cholesterol            390 non-null    int64
2   glucose                390 non-null    int64
3   age                   390 non-null    int64
4   gender                 390 non-null    object
5   height                390 non-null    int64
6   weight                390 non-null    int64
7   bmi                   390 non-null    float64
8   systolic_bp           390 non-null    int64
9   diastolic_bp          390 non-null    int64
10  waist                 390 non-null    int64
11  hip                   390 non-null    int64
12  diabetes               390 non-null    object
dtypes: float64(1), int64(10), object(2)
memory usage: 39.7+ KB
```

FIGURE 3.3 – Propriétés du dataset

Comme nous le pouvons le voir, ce dataset ne contient pas de valeurs nulles. Il faut rappeler que les informations contenues dans ce dataset doivent être vérifiées par un expert (un médecin)

afin qu'on puisse l'utiliser dans notre conception.

## 3.5 Construction d'une ontologie floue pour le diabète de type 2 et l'emploi de l'internet des objets médicaux

La méthodologie FODM a été proposée dans un cadre général pour le développement des ontologies floues. Dans notre étude, nous allons suivre cette méthodologie afin de construire une ontologie floue du diabète de type 2 et de l'IoMT, formalisée grâce au plug-in Fuzzy-OWL 2.

### 3.5.1 Phase 1 : Objectif et portée de l'ontologie

Dans le but de décrire le domaine cible, la motivation est de modéliser le diagnostic du diabète de type 2 avec l'aide de l'internet des objets médicaux IoMT. Une liste de questions est esquissée comme suit et les réponses à celles-ci peuvent être utiles pour déterminer l'objectif et la portée de l'ontologie de manière précise.

**Q1** : Quel type d'information doit être modélisé? Compte tenu de notre domaine d'étude, les informations modélisées seront : les facteurs risques du diabète de type 2, ses symptômes, et aussi des composantes et des caractéristiques de l'IoMT.

**Q2** : L'ontologie est-elle la meilleure technique de modélisation par rapport aux autres solutions comme la modélisation clé-valeur, UML, la modélisation graphique ou la modélisation multidisciplinaire ? Oui, pour différentes raisons :

- Premièrement, elle permet de caractériser des informations afin de relier des connaissances et de les inférer pour bien représenter le domaine très large de l'IoMT.
- Deuxièmement, elle est un modèle réutilisable qui permet le partage des connaissances afin de constituer une base commune à l'ensemble du domaine pour l'échange et l'interprétation des informations.
- Troisièmement, grâce à son langage formel, la communication entre le médecin et le patient est possible, de plus son extensibilité permet d'élargir la portée globale des connaissances modélisées.

**Q3** : Quel sera le type de l'ontologie envisagée? Le type de l'ontologie est une « ontologie du domaine », elle est utilisée pour représenter un domaine sous forme de base de connaissances. Elle présente les concepts-clés, des attributs et les instances relatifs au domaine.

**Q4** : Qui seront impliqués dans le développement de l'ontologie et quels rôles ils vont jouer? Les développeurs (nous) pourraient être les principaux participants, ainsi que les médecins qui auront participé à la contribution des connaissances nécessaires pour la modélisation.

**Q5** : Comment assurer une collaboration étroite entre les différents participants afin de

garantir un développement réussi de l'ontologie ? Nous les développeurs devons être à l'écoute des avis des médecins pour assurer une modélisation juste des connaissances dans l'ontologie et y appliquer une validation pour éliminer d'éventuelles contradictions et vérifier la satisfiabilité de certains critères de validation avec les autres participants.

### 3.5.2 Phase 2 : Identifier les besoins flous

Selon l'objectif de notre ontologie, le diagnostic médical du diabète de type 2, et les informations de modélisation prévues seront les facteurs risques de la maladie qui seront analysés grâce à des objets connectés médicaux pour plusieurs facteurs risques, tandis que les symptômes seront analysés par le médecin.

Cependant il est obligatoire de prendre en compte l'implication de l'incertitude et surtout le flou donc il est à noter de concevoir des concepts flous sur chaque concept précis pour prendre en compte tous les cas de degrés d'appartenance d'une instance (patient) à ce concept par exemple pour l'âge, on s'y prendra ainsi en créant les concepts flous suivants : "jeune", "vieux" ainsi que les attributs flous spécifiques à ces derniers, ce qui fait que chaque patient aura des degrés d'appartenance à chacun de ces concepts. En se basant sur cette approche, le flou est nécessaire pour gérer l'imprécision inhérente à la classification du diabète de type 2 chez les patients.

De ce fait, on conclut à partir de cette phase qu'une ontologie floue est bien meilleure qu'une ontologie classique pour modéliser le domaine du diabète de type 2.

### 3.5.3 Phase 3 : Déterminer les informations floues et précises

Dans cette étape, il est convenable d'établir une distinction claire entre les informations floues et les informations précises. Après avoir recueilli les connaissances et les suggestions de médecins, nous répertorions les connaissances comme suit :

**Connaissances floues :** Des facteurs risques causant le diabète seront considérées comme obligatoires à y appliquer la modélisation floue conseillée par le médecin, ces facteurs sont :

- L'âge
- La taille
- Le poids
- L'indice de masse corporelle (BMI)
- Tour de taille
- Tour de hanche

D'autres facteurs à évaluer seront pris en compte et considérés comme des mesures dans le cadre de l'IoMT car seront capturés par des objets médicaux connectés qui seront :



- Le taux de cholestérol total
- La glycémie
- La tension artérielle systolique
- La tension artérielle diastolique

Nous avons précédemment évoqué tous ces connaissances dans la **section 3.4**.

**Connaissances précises :** Pour commencer on aura les symptômes du diabète de type 2 qui seront précisés, car ceux-ci se caractérisent par leur présence ou pas, il y aura aussi les types d'acteurs impliqués dans le diagnostic qui seront le docteur et le patient évidemment. À propos du cadre IoMT, premièrement les types d'appareils feront partie de l'ontologie floue où on citera les dispositifs de détection, d'actionnement, et d'identification. Des caractéristiques d'objets médicaux jugés importantes pour le diagnostic seront modélisées, ainsi que d'autres paramètres de l'IoMT comme l'historique, le temps, la localisation, . . . etc.

### **3.5.4 Phase 4 : Envisager la réutilisation des ontologies existantes**

À propos des ontologies déjà existantes, une ontologie classique nous aura aidé pour tirer des connaissances dans le domaine d'IoMT comme les types d'appareils, les caractéristiques, et les paramètres contextuels. Voici le lien URI de cette ontologie :

<http://www.semanticweb.org/hp/ontologies/2018/4/Health-Iot>

### **3.5.5 Phase 5 et 6 : Réutiliser les éléments d'ontologies floues et les corriger**

Étant donné qu'aucun élément d'ontologie floue existant n'est considéré comme pouvant être réutilisé dans l'ontologie IoMT Diabète, ces deux phases vont être ignorées.

### **3.5.6 Phase 7 : Définir les éléments flous de l'ontologie**

Pour les éléments flous, on aura besoin des axiomes pour établir les liaisons entre les concepts flous et les attributs flous.

Les concepts flous ainsi que leurs attributs, les fonctions d'appartenance des données floues et les axiomes sont représentés dans le **tableau 3.1** :

Concept flou	Attribut du concept	Sous concepts flous	Attributs des sous concepts	Fonction d'appartenance	Axiomes
<b>Sous classes de « Measurement »</b>					
Cholesterol	Has_Cholesterol	Low_Cholesterol	Low_Cholesterol_rate	Leftshoulder(78,443,180.36,25 1.11)	Low_Cholesterol $\exists$ Cholesterol $\cap \exists$ Has_Cholesterol . Low_Cholesterol_rate
		High_Cholesterol	High_Cholesterol_rate	rightshoulder(78,443,180.36,2 51.11)	High_Cholesterol $\exists$ Cholesterol $\cap \exists$ Has_Cholesterol . High_Cholesterol_rate
Glucose	Has_Glucose	Low_Glucose	Low_Glucose_rate	Leftshoulder(48,385,89.97,242 .93)	Low_Glucose $\exists$ Glucose $\cap \exists$ Has_Glucose . Low_Glucose_rate
		High_Glucose	High_Glucose_rate	Rightshoulder(48,385,89.97,2 42.93)	High_Glucose $\exists$ Glucose $\cap \exists$ Has_Glucose . High_Glucose_rate
systolic_bp	Has_SBP_rate	Low_systolic_bp	Low_Sbp_rate	Leftshoulder(90,250,125.14,16 4.15)	Low_systolic_bp $\exists$ systolic_bp $\cap \exists$ Has_SBP_rate . Low_Sbp_rate
		High_systolic_bp	High_Sbp_rate	Rightshoulder(90,250,125.14, 164.15)	High_systolic_bp $\exists$ systolic_bp $\cap \exists$ Has_SBP_rate . High_Sbp_rate
diastolic_bp	Has_DBP_rate	Low_diastolic_bp	Low_dBP_rate	Leftshoulder(48,124,73.23,94. 89)	Low_diastolic_bp $\exists$ diastolic_bp $\cap \exists$ Has_DBP_rate . Low_dBP_rate
		High_diastolic_bp	High_dBP_rate	Rightshoulder(48,124,73.23,9 4.89)	High_diastolic_bp $\exists$ diastolic_bp $\cap \exists$ Has_DBP_rate . High_dBP_rate
<b>Sous classes de « Risk_Factor »</b>					
Height	Has_height	Low_Height	Has_low_height	Leftshoulder(122.08,203.4, 159.74,176.02)	Low_Height $\exists$ Height $\cap \exists$ Has_height . Has_low_height
		High_Height	Has_high_height	Rightshoulder(122.08,203.4,1 59.74,176.02)	High_Height $\exists$ Height $\cap \exists$ Has_height . Has_high_height
Weight	Has_weight	Low_Weight	Has_low_weight	Leftshoulder(40, 150, 68.86, 98.48)	Low_Weight $\exists$ Weight $\cap$ Has_weight . Has_low_weight
		High_weight	Has_high_weight	Rightshoulder(40, 150, 68.86, 98.48)	High_weight $\exists$ Weight $\cap$ Has_weight . Has_high_weight
Waist	Has_waist	Thin_Waist	Has_thin_waist	Leftshoulder(26,56,33.74,43.5 5)	Thin_Waist $\exists$ Waist $\cap \exists$ Has_waist . Has_thin_waist
		Large_Waist	Has_large_waist	Rightshoulder(26,56,33.74,43. 55)	Large_Waist $\exists$ Waist $\cap \exists$ Has_waist . Has_large_waist
Hip	Has_hip	Thin_Hip	Has_thin_hip	Leftshoulder(30,64,39.52,48.9 8)	Thin_Hip $\exists$ Hip $\cap \exists$ Has_hip . Has_thin_hip
		Large_Hip	Has_large_hip	Rightshoulder(30,64,39.52,48. 98)	Large_Hip $\exists$ Hip $\cap \exists$ Has_hip . Has_large_hip
Age	Has_Age	AgeYoung	has_youngAge	Leftshoulder(19,92,34,63) (Valeurs arrondies)	AgeYoung $\exists$ Age $\cap \exists$ Has_Age . has_youngAge
		AgeOld	has_oldAge	Rightshoulder(19,92,34,63) (Valeurs arrondies)	AgeOld $\exists$ Age $\cap \exists$ Has_Age . has_oldAge
BMI	Has_BMI	BMIslim	Has_slimBMI	Leftshoulder(15.2,55.8,24.74,3 5.94)	BMIslim $\exists$ BMI $\cap \exists$ Has_BMI . Has_slimBMI
		BMIobese	Has_obeseBMI	Rightshoulder(15.2,55.8,24.74 ,35.94)	BMIobese $\exists$ BMI $\cap \exists$ Has_BMI . Has_obeseBMI

TABLE 3.1 – Concepts, attributs flous ainsi que leurs fonctions d'appartenance

### — 3.5.6.1 DATIL

Pour la définition des types de données flous et leurs fonctions d'appartenance respectives, nous avons utilisé DATIL (DATatypes with Imprecision Learner).

DATIL est un logiciel qui apprend automatiquement les types de données flous pour les ontologies floues à partir de différents types d'entrées (3 formats d'entrées supportés : ".owl", ".fld", ".csv"). DATIL implémente plusieurs algorithmes de clustering non supervisés : k-means, fuzzy c-means et mean-shift. L'outil est accessible au public à partir du site : <http://webdiis.unizar.es/ihv->

dis/Datil.html.

Pour chaque propriété de données dans une ontologie avec une plage numérique (ou avec des assertions impliquant des valeurs numériques), DATIL collecte un tableau de nombres réels correspondant aux valeurs de la propriété pour différents individus.

Un algorithme de clustering fournit un ensemble de centroïdes à partir de ces tableaux de valeurs. Ces centroïdes sont utilisés comme paramètres pour construire des fonctions d'appartenance floues partitionnant le domaine [67].

Pour notre cas d'utilisation, l'entrée utilisée est le dataset, et l'algorithme de clustering choisi est « Fuzzy C-means ». Cet algorithme fonctionne en attribuant un degré d'appartenance à chaque point de données correspondant à chaque centre de cluster sur la base de la distance entre le centre du cluster et le point de données. Plus les données sont proches du centre du cluster, plus leur appartenance au centre du cluster particulier est grande.

De toute évidence, la somme des différents degrés liés aux centroïdes attribués à chaque point de données doit être égale à 1.

Après avoir exécuté l'algorithme sur le Dataset, DATIL nous fournit les résultats sous forme d'un fichier OWL, les paramètres des fonctions d'appartenance sont décrites en tant qu'axiomes. La **figure 3.4** montre les fonctions d'appartenance issues de l'application de DATIL sur la colonne "Âge" :

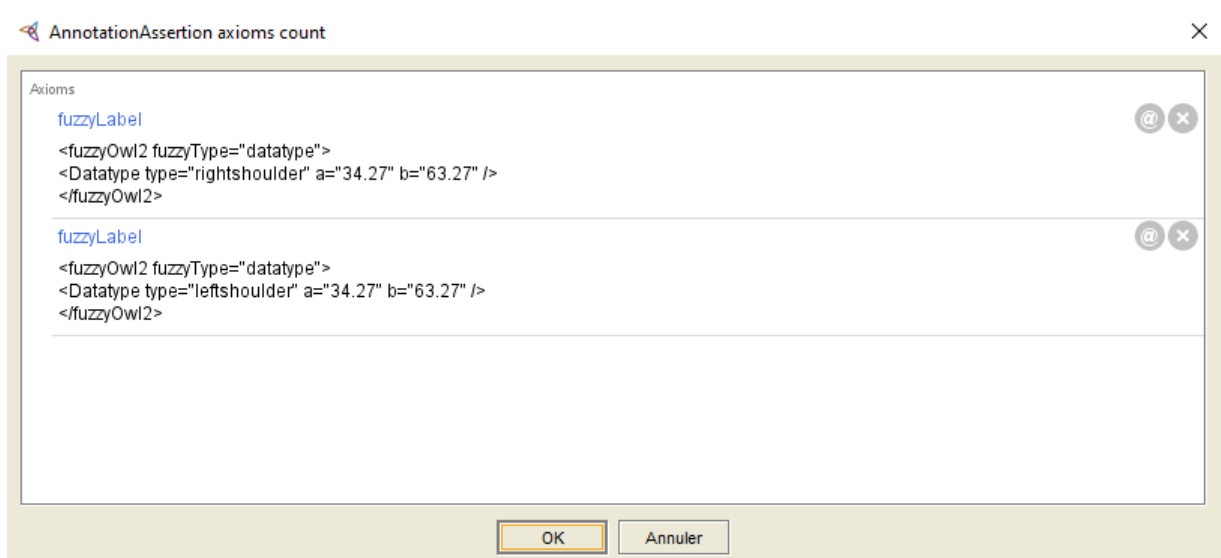


FIGURE 3.4 – Fonctions d'appartenance de : "AgeYoung" et "AgeOld"

### 3.5.7 Phase 8 : Définir des éléments précis de l'ontologie

Nous présentons dans les **tableaux 3.2 & 3.3** les concepts ainsi que les attributs précis :

Concept	Description	Attributs	Type des attributs
Medical_Object	Un objet connecté médical est utilisé et surveillé une mesure chez le patient.	Id_object Name_object	Int String
lot_Context	Un ensemble des faits ou des circonstances qui entourent une utilisation d'un objet médical.	-	-
Location	L'emplacement d'un patient/appareil est une information importante à prendre en compte.	-	-
geo-point	Position par point-GPS (norme WSG84) défini par les 2 coordonnées: latitude, longitude.	Latitude longitude	Float Float
location-address	Position adresse	adress	String
Historic	Pour tracer les diagnostics effectués d'un patient.	-	-
Network_Property	Pour décrire quelques propriétés réseaux.	-	-
latency	Durée de transmission des données d'une entité IoT à une autre.	ping	int
Network-capability	Capacité du réseau où le nombre maximal de nœuds possibles dans le réseau.	netcapability	int
Time		-	-
Time:duration	Durée de travail d'un appareil.	start_time end_time	Xsd:dateTimeStamp Xsd:dateTimeStamp
Time:instant	Contient la date et l'heure d'une action par exemple.	instant	Xsd:dateTimeStamp (de la forme "YYYY-MM-DDThh:mm:ss")
IoT_devices	Différentes technologies des objets connectés, ces technologies sont indispensables pour le fonctionnement de ces derniers.	-	-
Tagging_devices	Les technologies permettant aux objets de s'identifier, cela peut être des étiquettes RFID, des codes-barres,...etc.	-	-
Actuator	une technologie complémentaire aux capteurs, un actionneur convertit l'énergie électrique en mouvement ou énergie mécanique pour appliquer une action.	-	-
Sensor	dispositif utilisé pour détecter un événement ou une grandeur physique, tels que luminosité, température, humidité du sol, pression, etc.	-	-
Symptoms	Les effets apparents de la maladie d'étude (diabète type 2, ces symptômes sont représentés en ces sous concepts.	-	-
Blurred_Vision	vision floue	has_blurred_vision	Boolean
Dry_Skin	peau sèche	has_dry_skin	Boolean
Frequent_Infections	infections fréquentes	has_frequent_infections	Boolean
Frequent_Urination	besoin fréquent d'uriner	has_frequent_urination	Boolean
Increased_Hunger	augmentation de la faim	has_increased_hunger	Boolean
Increased_Thirst	augmentation de la soif	has_increased_thirst	Boolean
Neurological_Disorders	insensibilité ou fourmillement des mains et/ou des pieds	has_neurological_disorders	Boolean
Asthenia:Tiredness	fatigue	has_tiredness	Boolean
Slow_Wound_Healing	cicatrisation lente des blessures	has_slow_wound_healing	Boolean
Person	tous les acteurs participant dans le diagnostic.	-	-
Doctor	Un docteur est nécessaire pour tirer quelques informations d'un patient, aussi des symptômes que seul lui peut reconnaître.	family_name first_name	String String
Patient	Est la personne qui sera diagnostiquée bien sûr.	family_name first_name	String String

TABLE 3.2 – Concepts et attributs précis - partie 1

Concept	Description	Attributs	Type des attributs
IoT_characteristics	Des critères spécifiques aux entités IoT et aux capteurs, les concepts ci-dessous de ce tableau héritent tous de cette classe.	-	-
Communication-protocol	les protocoles qui peuvent être utilisés pour que les objets communiquent entre eux, cela peut par exemple un Bluetooth, Zigbee,...etc.	id_protocol Type_protocol	int string
Coverage	Représente « l'espace physique » couvert par une entité IoT (par exemple un capteur d'humidité à l'intérieur d'une pièce a une couverture cette pièce)	coverage	string
Energy_source	Source d'énergie du capteur qui est généralement une batterie intégrée.	energy_source	string
Frequency	La fréquence de collecte des données.	frequency	float
Lifetime	Durée de vie de l'appareil exprimée en mois.	lifetime	int
Memory_capability	Capacité de la mémoire de l'objet medical connecté exprimé en (X unit).	memory_capability	float
Unit_of_measure	Il faut une unité de mesure pour représenter la donnée collectée à partir du capteur	unit_of_measure	string
Device_state	On/off	device_state	{"on"} or {"off"}
Sensing_Measurement_Range	L'étendue de mesure est la différence entre la limite supérieure et la limite inférieure de la grandeur mesurable par un capteur.	sensing_measurement_range	float
Sensing_Measurement_Time	Le temps de réaction du capteur, il est liée à la bande-passante.	sensing_measurement_time	float
Resolution	Le plus petit changement dans une mesure que le capteur est capable de détecter.	-	-
Measurement	Représente la donnée qui est capté par un objet connecté, la mesure est caractérisée par une valeur elle peut être: la glycémie, les 2 tensions artérielles, le cholestérol, et le cholestérolHDL qui ont été présentés comme concepts flous précédemment.	-	-
Risk_Factors	Les facteurs risques qui sont à la cause de l'atteinte d'une personne du diabète type 2.	-	-
gender		gender	{"male"} or {"female"}

TABLE 3.3 – Concepts et attributs précis - partie 2

Les **tableaux 3.4 & 3.5** montrent les relations établies entre différents concepts de l'ontologie :

Concept Domaine	La relation	Concept Range
Doctor	<i>Observes</i>	Glucose
Doctor	<i>Observes</i>	Cholesterol
Doctor	<i>Observes</i>	Diastolic_bp
Doctor	<i>Observes</i>	Systolic_bp
Doctor	<i>Examines</i>	Patient
Doctor	<i>Identifies</i>	Symptoms
Doctor	<i>Identifies</i>	Asthenia:Tiredness
Doctor	<i>Identifies</i>	Blurred_Vision
Doctor	<i>Identifies</i>	Frequent_Infections
Doctor	<i>Identifies</i>	Frequent_Urination
Doctor	<i>Identifies</i>	Increased_Hunger
Doctor	<i>Identifies</i>	Increased_Thirst
Doctor	<i>Identifies</i>	Dry_Skin
Doctor	<i>Identifies</i>	Neurological_Disorders
Doctor	<i>Identifies</i>	Slow_Wound_Healing
Patient	<i>Has_historic</i>	Historic
Patient	<i>Has_object</i>	Medical_Object
Patient	<i>Has_blurred_vision</i>	Blurred_Vision
Patient	<i>Has_dry_skin</i>	Dry_Skin
Patient	<i>Has_frequent_infections</i>	Frequent_Infections
Patient	<i>Has_frequent_urination</i>	Frequent_Urination
Patient	<i>Has_increased_hunger</i>	Increased_Hunger
Patient	<i>Has_increased_thirst</i>	Increased_Thirst
Patient	<i>Has_neurological_disorders</i>	Neurological_Disorders
Patient	<i>Has_tiredness</i>	Asthenia:Tiredness
Patient	<i>Has_slow_wound_healing</i>	Slow_Wound_Healing
Patient	<i>Has_SBP</i>	Systolic_bp
Patient	<i>Has_DBP</i>	Diastolic_bp
Patient	<i>Has_Glucose</i>	Glucose
Patient	<i>Has_age</i>	Age
Patient	<i>Has_bmi</i>	BMI
Patient	<i>Has_cholesterol</i>	Cholesterol
Patient	<i>Has_gender</i>	Gender
Patient	<i>Has_height</i>	Height
Patient	<i>Has_weight</i>	Weight
Patient	<i>Has_waist</i>	Waist
Patient	<i>Has_hip</i>	Hip

TABLE 3.4 – Relations - partie 1

<b>Concept Domaine</b>	<b>La relation</b>	<b>Concept Range</b>
Medical_Object	<i>sendsData</i>	Doctor
Medical_Object	<i>Has_sensor</i>	Sensor
Medical_Object	<i>Has_actuator</i>	Actuator
Medical_Object	<i>Has_tagging_devices</i>	Tagging_devices
Medical_Object	<i>Has_battery_level</i>	Battery_level
Medical_Object	<i>Communicate_via</i>	Communication-protocol
Medical_Object	<i>Has_coverage</i>	Coverage
Medical_Object	<i>Has_device_state</i>	Device_state
Medical_Object	<i>Has_energy_source</i>	Energy_source
Medical_Object	<i>Has_lifetime</i>	Lifetime
Medical_Object	<i>Has_memory_capability</i>	Memory_capability
Medical_Object	<i>Has_location</i>	Location
Medical_Object	<i>Has_geo-point</i>	Geo-point
Medical_Object	<i>Has_location-address</i>	Location-address
Medical_Object	<i>Has_networkproperty</i>	NetworkProperty
Medical_Object	<i>Links_data</i>	Measurement
Sensor	<i>Has_frequency</i>	Frequency
Sensor	<i>Has_accuracy</i>	Accuracy
Sensor	<i>Has_resolution</i>	Resolution
Sensor	<i>Has_sensing_measurement_range</i>	Sensing_Measurement_Range
Sensor	<i>Has_sensing_response_time</i>	Sensing_Response_Time
Sensor	<i>Has_unit_of_measure</i>	Unit_of_measure
Measurement	<i>Monitored_during</i>	Time:duration
Diastolic_bp	<i>Monitored_during</i>	Time:duration
Systolic_bp	<i>Monitored_during</i>	Time:duration
Glucose	<i>Monitored_during</i>	Time:duration
Cholesterol	<i>Monitored_during</i>	Time:duration
Measurement	<i>Monitored_At</i>	Time:instant
Diastolic_bp	<i>Monitored_At</i>	Time:instant
Systolic_bp	<i>Monitored_At</i>	Time:instant
Glucose	<i>Monitored_At</i>	Time:instant
Cholesterol	<i>Monitored_At</i>	Time:instant

TABLE 3.5 – Relations - partie 2

### 3.5.8 Phase 9 : Formalisation

Au niveau de cette phase, un langage doit être sélectionné pour formaliser l'ontologie conçue dans un format lisible par machine. Pour notre ontologie, OWL 2 est sélectionné pour représenter le modèle d'ontologie conçu. Pour effectuer facilement la transformation du modèle conceptuel en expressions au format OWL2, l'éditeur d'ontologie Protégé et son extension FuzzyOWL sont utilisés dans cette étape. Protégé permet des implémentations visualisées et faciles de l'ontologie de représentation du diagnostic du diabète de type 2, tandis que FuzzyOWL est responsable de la représentation des éléments flous de l'ontologie, l'implémentation de notre ontologie est illustrée dans la **figure 3.5** ci-dessous :

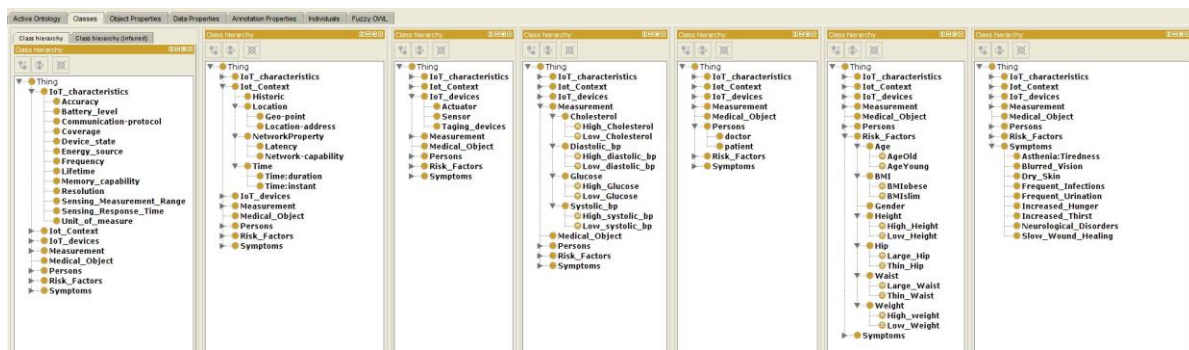


FIGURE 3.5 – Représentation taxonomique de l'ontologie dans Protégé

### 3.5.9 Phase 10 : Validation

Les résultats de la validation servent de preuve de l'utilité de l'ontologie développée. Dans cette phase, la validation de notre ontologie est effectuée en vérifiant des caractéristiques qui nous citons ci-dessus :

**-Cohérence & Consistance** : La cohérence et la consistance d'une ontologie sont deux caractéristiques qui prouvent l'absence de contradiction entre les concepts d'ontologies. Une ontologie est cohérente si tous ses concepts sont satisfiables, l'incohérence signifie l'existence un concept qui n'a pas d'interprétation et qui ne peut accepter aucune instance. Ainsi, la cohérence d'une ontologie est déterminée par l'absence de contradictions logiques entre les représentations qu'elle contient et elle est liée à la TBox. En revanche, la consistance de l'ontologie est liée à la ABox et consiste à ce que tous les individus de l'ontologie soient valides.

Pour prouver ces deux caractéristiques dans notre ontologie, nous avons utilisé Hermit 1.3.8 qui est un raisonneur qui détecte automatiquement la présence des inconsistances et incohérences, en cas de présence d'un de ces types d'erreurs, une fenêtre est affiché pour expliquer l'erreur en question. Aucune erreur n'a été trouvée au niveau de toute l'ontologie.

La **figure 3.6** montre le lancement du raisonneur :



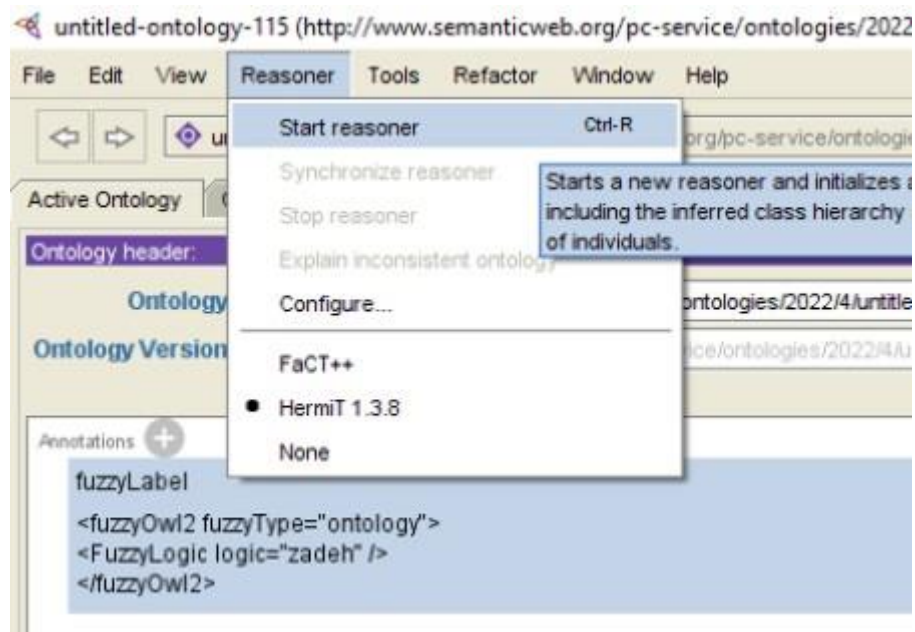


FIGURE 3.6 – Activation du reasonner

La figures 3.7 montre le reasonner en cours ne signalant aucune erreur :

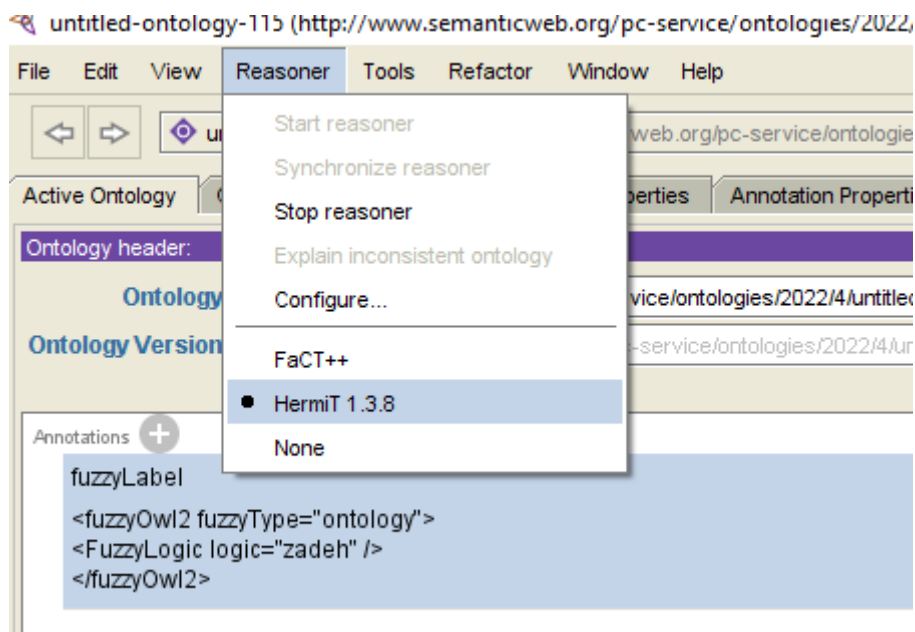


FIGURE 3.7 – L'ontologie en mode reasonner activé

-**Complétude** : La fonctionnalité d'exhaustivité garantit que l'ontologie conçue a pu couvrir tous les aspects de l'information appartenant au domaine cible.

-Selon le but de notre système qui est de diagnostiquer un patient sur le diabète de type 2, les facteurs de risques jugés les plus importants selon les médecins ayant participé à l'identification des besoins ont été modélisés, ces facteurs sont jugés importants même selon "l'Institut national du diabète" propriétaire du dataset utilisé dans notre étude.

**-Compréhensibilité** : Les éléments de l'ontologie doivent être facilement compréhensibles pour toutes les parties prenantes, y compris les médecins, les ingénieurs en ontologie,...etc.

-De ce qui est de l'ontologie, nous avons évité les termes compliqués, pour la terminologie concernant l'IoMT, les termes sont déjà faciles à comprendre, et pour les facteurs de risques du diabète et les symptômes, nous avons essayé d'employer le minimum possible de termes médicaux, les quelques termes médicaux employés sont généralement connus, et tous les termes sont clairs excepté quelques-uns qui ont été expliqués.

### 3.5.10 Phase 11 : Documentation

Le but de cette étape est de fournir des documentations sur tous les éléments de cette ontologie qui sera publié publiquement, cela peut être déjà réalisé en publiant notre mémoire.

### 3.5.11 Instanciation de l'ontologie floue

Après l'étape de conception et d'implémentation de l'ontologie, rappelons que nous devons remplir l'ontologie avec des instances venant du dataset, c'est l'instanciation.

---

**Algorithm** InstanciationOntologie

---

```
1 INPUT Ontology O + Dataset D
2 OUTPUT Ontology O
3 BEGIN
4 D.load()
5 for each linetext; from D do
6 Data[] ← linetext.split(",")
7 //Creation des noms pour les instances
8 patienti = "patient" + i;
9 agei ← "Age_patient" + i;
10 sexei ← "Sexe_patient" + i;
11 choli ← "Cholesterol_patient" + i;
12 glui ← "Glucose_patient" + i;
13 tasi ← "Tension_artrielle_systolique_patient" + i;
14 tadi ← "Tension_artrielle_diastolique_patient" + i;
15 taillei ← "taille_patient" + i;
16 poidsi ← "Poids_patient" + i;
17 imci ← "IMC_patient" + i;
18 tti ← "Tour_de_taille_patient" + i;
19 thi ← "Tour_de_hanches_patient" + i;
20 //Creation et ajout des instances l'ontologie
21 O.addIndividual(patient1, class : "patient");
22 O.addIndividual(agei, class : "Age");
```

```

23 O.addIndividual(sexei, class : "Gender");
24 O.addIndividual(choli, class : "Cholesterol");
25 O.addIndividual(glui, class : "Glucose");
26 O.addIndividual(tasi, class : "Systolic_bp");
27 O.addIndividual(tadi, class : "Diastolic_bp");
28 O.addIndividual(taillei, class : "Height");
29 O.addIndividual(poidsi, class : "Weight");
30 O.addIndividual(imci, class : "BMI");
31 O.addIndividual(tti, class : "Waist");
32 O.addIndividual(thi, class : "Hip");
33 //ajout des DataProperty aux instances de classes
34 O.addDataProperty(agei, "Has_Age", Data[1]);
35 O.addDataProperty(sexei, "gender", Data[2]);
36 O.addDataProperty(choli, "Has_Cholesterol", Data[3]);
37 O.addDataProperty(glui, "Has_Glucose", Data[4]);
38 O.addDataProperty(tasi, "Has_SBP_rate", Data[5]);
39 O.addDataProperty(tadi, "Has_DBP_rate", Data[6]);
40 O.addDataProperty(taillei, "Has_height", Data[7]);
41 O.addDataProperty(poidsi, "Has_weight", Data[8]);
42 O.addDataProperty(imci, "Has_BMI", Data[9]);
43 O.addDataProperty(tti, "Has_waist", Data[10]);
44 O.addDataProperty(thi, "Has_hip", Data[11]);
45 O.addDataProperty(patienti, "Diagnostic_Final", Data[12]);
46 //creer les relations entre instances
47 O.linkIndividuals(patienti, age1, "Has_age");
48 O.linkIndividuals(patienti, sexe1, "Has_gender");
49 O.linkIndividuals(patienti, chol1, "Has_cholesterol");
50 O.linkIndividuals(patienti, glu1, "Has_glucose");
51 O.linkIndividuals(patienti, tas1, "Has_SBP ");
52 O.linkIndividuals(patienti, tad1, "Has_DBP ");
53 O.linkIndividuals(patienti, taille1, "Has_Height");
54 O.linkIndividuals(patienti, poids1, "Has_Weight");
55 O.linkIndividuals(patienti, imc1, "Has_bmi");
56 O.linkIndividuals(patienti, tt1, "Has_Waist");
57 O.linkIndividuals(patienti, th1, "Has_Hip");
58 end for
59 END

```

Dans ce qui suit, nous allons expliquer ce que fait chaque bloc de l'algorithme précédent.

- ligne 1-2 : spécifier les entrées (ontologie floue vide et dataset) et sorties (ontologie floue remplie) du code.
- ligne 4 : chargement du dataset dans le programme.
- ligne 5-58 : Dans la ligne 5 la création d'une boucle qui se ferme à la ligne 58, la boucle sert à lire chaque ligne du dataset pour utiliser les données de cette dernière pour créer des instances, dataproperty, et relations d'instances liées à un patient.
- ligne 6 : divise chaque ligne du dataset en plusieurs valeurs qui seront stockées dans le vecteur Data [ ], la division de la ligne se base sur les virgules.
- ligne 8-19 : création de noms pour les instances des classes créés ultérieurement.
- ligne 21-32 : création et ajout des instances dans l'ontologie grâce à la méthode "addIndividual".
- ligne 34-45 : ajout des data property aux instances des classes grâce à la méthode "addDataProperty".
- ligne 47-57 : création des relations entre les instances grâce à la méthode "linkIndividuals".
- ligne 59 : Fin de l'algorithme.

## 3.6 Construction du réseau Bayésien Flou

Dans cette section nous représentons les étapes de construction de notre réseau bayésien en commençant par l'application de l'apprentissage de la structure et des paramètres de notre réseau bayésien et en finissant par la fuzzification.

### 3.6.1 Catégorisation du dataset

Afin de pouvoir appliquer plus tard l'apprentissage sur notre réseau bayésien, le dataset ne doit contenir que des valeurs discrètes, c'est là que la catégorisation entre en jeu, nous allons catégoriser les colonnes contenant des valeurs continues, donc toutes les colonnes seront catégorisées excepté le genre et la colonne du diagnostic final, le **tableau 3.6** représente la table des catégorisations des colonnes du dataset :

Colonnes	Normalisation
Cholesterol	Low_Cholesterol < 215 High_Cholesterol => 215
Glucose	Low_Glucose < 115 High_Glucose => 115
Systolic_bp	Low_systolic_bp < 160 High_systolic_bp => 160
Diastolic_bp	Low_Diastolic_bp < 85 High_Diastolic_bp => 85
Age	AgeYoung < 50 AgeOld => 50
Height	Low_height < 170 High_height => 170
Weight	Low_weight < 85 High_weight => 85
Bmi	Bmslim < 30 Bmiobese => 30
Waist	Thin_Waist < 41 Large_Waist => 41
Hip	Thin_Hip < 47 Large_Hip => 47

TABLE 3.6 – Table des catégorisations de chaque colonne du dataset

Maintenant nous avons notre dataset catégorisé (voir la **figure 3.8**) et prêt à être utilisé pour l'apprentissage des paramètres.

	cholesterol	glucose	age	gender	height	weight	bmi	systolic_bp	diastolic_bp	waist	hip	diabetes
0	Low_Cholesterol	Low_Glucose	AgeYoung	female	Low_Height	Low_Weight	BMslim	Low_systolic_bp	Low_diastolic_bp	Thin_Waist	Thin_Hip	No
1	Low_Cholesterol	Low_Glucose	AgeYoung	female	Low_Height	Low_Weight	BMslim	Low_systolic_bp	Low_diastolic_bp	Thin_Waist	Thin_Hip	No
2	High_Cholesterol	Low_Glucose	AgeYoung	female	High_Height	High_weight	BMIobese	Low_systolic_bp	Low_diastolic_bp	Large_Waist	Large_Hip	No
3	High_Cholesterol	Low_Glucose	AgeYoung	female	Low_Height	Low_Weight	BMslim	Low_systolic_bp	Low_diastolic_bp	Thin_Waist	Thin_Hip	No
4	Low_Cholesterol	Low_Glucose	AgeYoung	female	High_Height	Low_Weight	BMslim	Low_systolic_bp	Low_diastolic_bp	Thin_Waist	Thin_Hip	No

FIGURE 3.8 – Dataset avec des valeurs discrètes

### 3.6.2 Apprentissage de la structure

L'algorithme d'apprentissage dans un réseau bayésien sert à trouver une structure idéale au problème posé. À l'aide d'un médecin considéré comme expert du domaine, nous avons pu construire le réseau bayésien de notre système.

La **figure 3.9** représente la structure du réseau Bayésien.

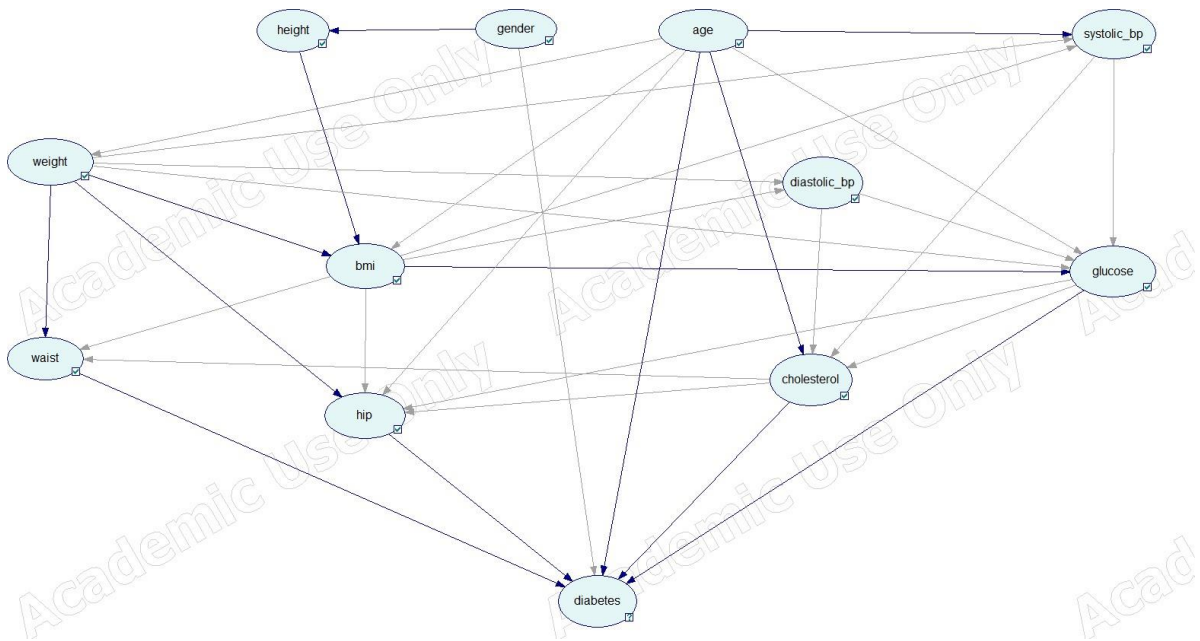


FIGURE 3.9 – Structure de réseau bayésien

### 3.6.3 Apprentissage des paramètres

Maintenant que la structure du réseau a été réalisée, il faudra estimer les probabilités conditionnelles de chaque nœud du réseau. Nous avons catégorisé notre base d'apprentissage dans la première étape, qui ne contient que des données discrètes. Il faut maintenant estimer la probabilité d'un événement en se basant sur la fréquence d'apparition de l'événement dans la base d'apprentissage, donc on utilisera l'approche maximum de vraisemblance (MV).

La **figure 3.10** représente les résultats obtenus des probabilités pour chaque nœud.

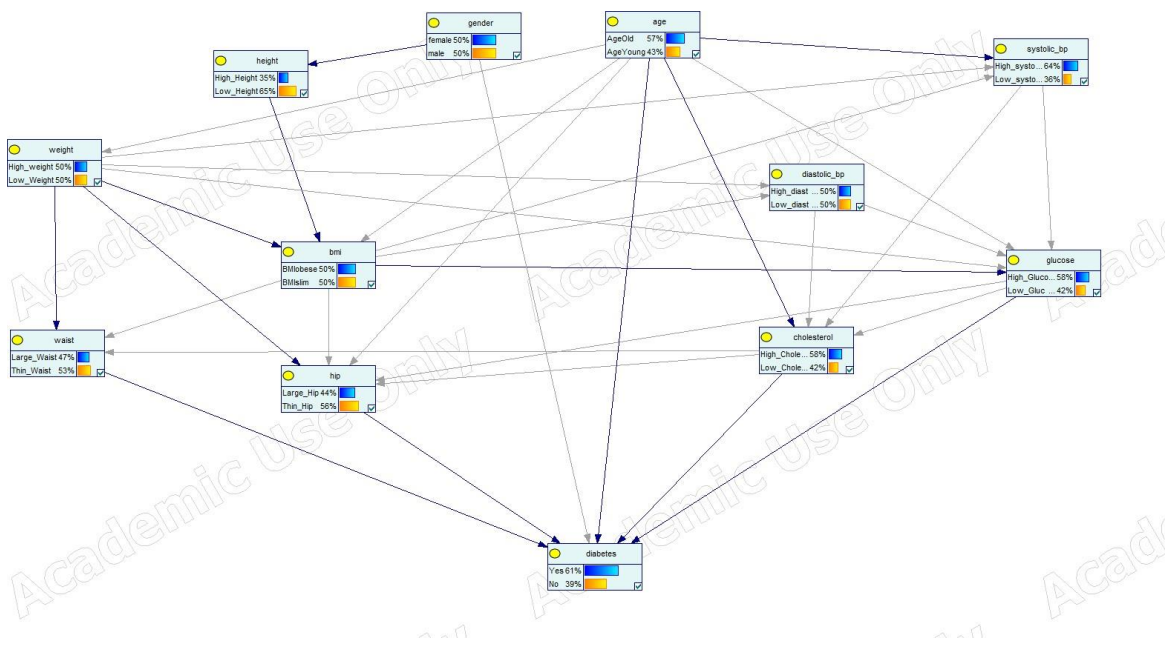


FIGURE 3.10 – Le réseau bayésien après l'apprentissage des paramètres

### 3.6.4 Fuzzification du RB

Et pour finir, pour transformer ce réseau bayésien en un RB flou, la fuzzification consiste à transformer chaque noeud du réseau bayésien en un noeud flou dont chaque état est défini avec une fonction d'appartenance. Les fonctions d'appartenance utilisées pour les noeuds sont les mêmes décrites dans la **section 3.5.6**, celles qui ont été trouvées avec DATIL.

## 3.7 Processus de diagnostic de l'état du patient

Dans cette section nous décrivons le diagnostic de l'état du patient en passant par l'ontologie floue et le RBF avec un diagramme représenté dans la **figure 3.17**. Ce processus se déroule comme suit :

- 1- Capturer les données du patient grâce aux capteurs connectés.
- 2- Stocker les données capturées dans l'ontologie.
- 3- Récupérer les données du patient à diagnostiquer à partir de l'ontologie.
- 4- Ensuite, ces données seront fuzzifiées selon les fonctions d'appartenance décrites précédemment.
- 5- Après la fuzzification, ces données sont prêtes à être envoyées au RBF pour qu'on puisse exécuter l'inférence avec ce dernier, pour donner les résultats des probabilités des deux états du noeud diagnostique qui sont YES (positif au diabète) et NO (négatif au diabète).
- 6- Après le processus choisira automatiquement la plus grande probabilité des deux pour être le résultat final du diagnostic.

7- Afficher les résultats de l'inférence.

La **figure 3.11** représente le processus complet de diagnostic d'un patient.

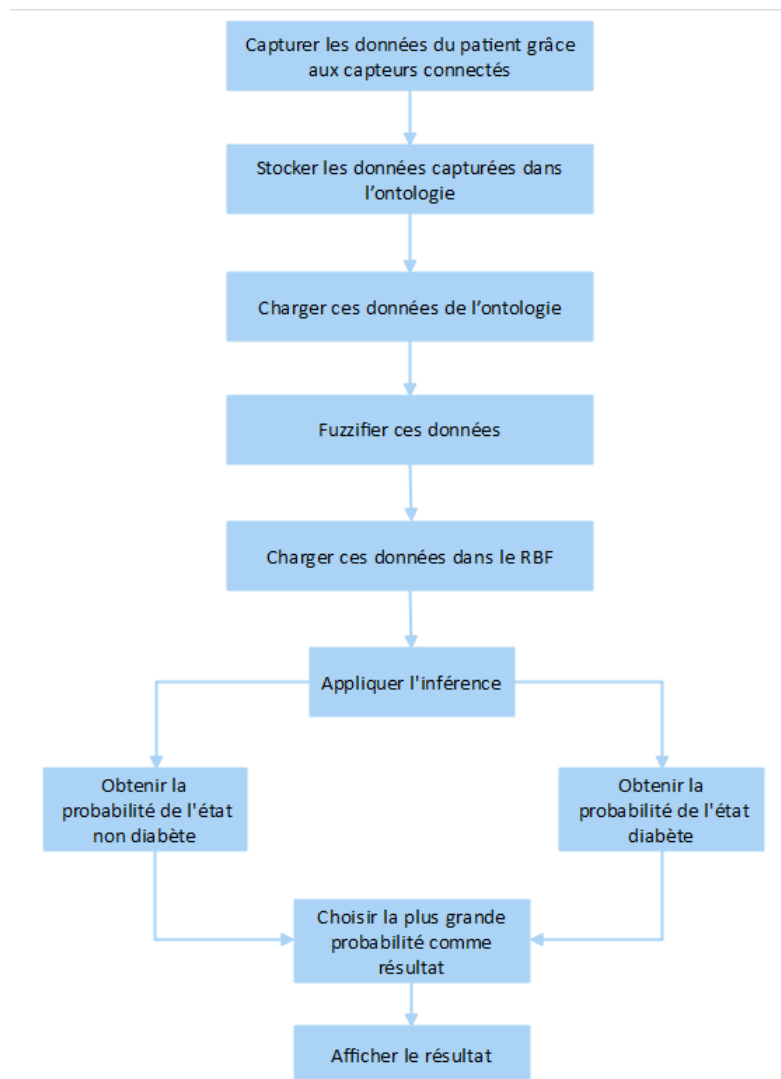


FIGURE 3.11 – Le processus complet de diagnostic incluant l'inférence du RBF et l'ontologie

### 3.8 Conclusion

Nous avons présenté dans ce chapitre les parties composant notre système avec le schéma global, à commencer par le dataset, ensuite le processus de création de l'ontologie floue sur le diabète de type 2 et l'IoMT en utilisant la méthodologie FODM.

Nous avons procédé par la suite à la construction de réseau Bayésien flou, l'explication du déroulement du processus global du diagnostic de l'état du patient en se basant sur le rôle de chacun de l'ontologie floue et du modèle RBF.

Dans le chapitre suivant, nous allons passer à l'implémentation de l'ontologie floue modélisée dans ce chapitre, pour finaliser notre système, le tester et valider les résultats trouvés.



## Chapitre 4

---

# Implémentation et test du système

---

## 4.1 Introduction

Après avoir montré et expliqué la conception du système, nous allons passer à l'étape d'implémentation de notre ontologie floue ainsi que de notre système dans son ensemble.

Dans ce qui va suivre, nous allons détailler l'implémentation de l'ontologie et du système, nous montrerons aussi les différents outils utilisés dans la réalisation de l'application, et pour finir, les tests effectués sur le système et la validation de nos résultats.

## 4.2 Implémentation du système

Le processus d'implémentation d'une ontologie floue est le même que celui d'une ontologie classique, la seule différence étant qu'il faut prendre en compte et ajouter la partie floue.

Afin de réaliser l'implémentation, nous avons utilisé l'éditeur Protégé version 4.3, et le plug-in **Fuzzy OWL 2** pour modéliser la partie floue.

### 4.2.1 Outils de programmation

Dans ce qui suit, nous allons présenter les divers outils utilisés lors du processus de programmation du système.

#### 4.2.1.1 Protégé

Protégé est un éditeur d'ontologies open-source qui supporte le langage OWL (Web Ontology Language). Il est disponible en téléchargement gratuitement sur le site officiel de protégé à savoir (stanford.edu). L'installation de protégé 4.3 est extrêmement simple. Il faut double-cliquer sur le programme d'installation téléchargé à partir du site. Ensuite, choisir l'emplacement d'installation sur l'ordinateur et quelle JVM (Java Virtual machine) va être utilisée, puis il ne reste qu'à suivre le reste des étapes.

La figure 4.1 représente la fenêtre principale de Protégé 4.3.

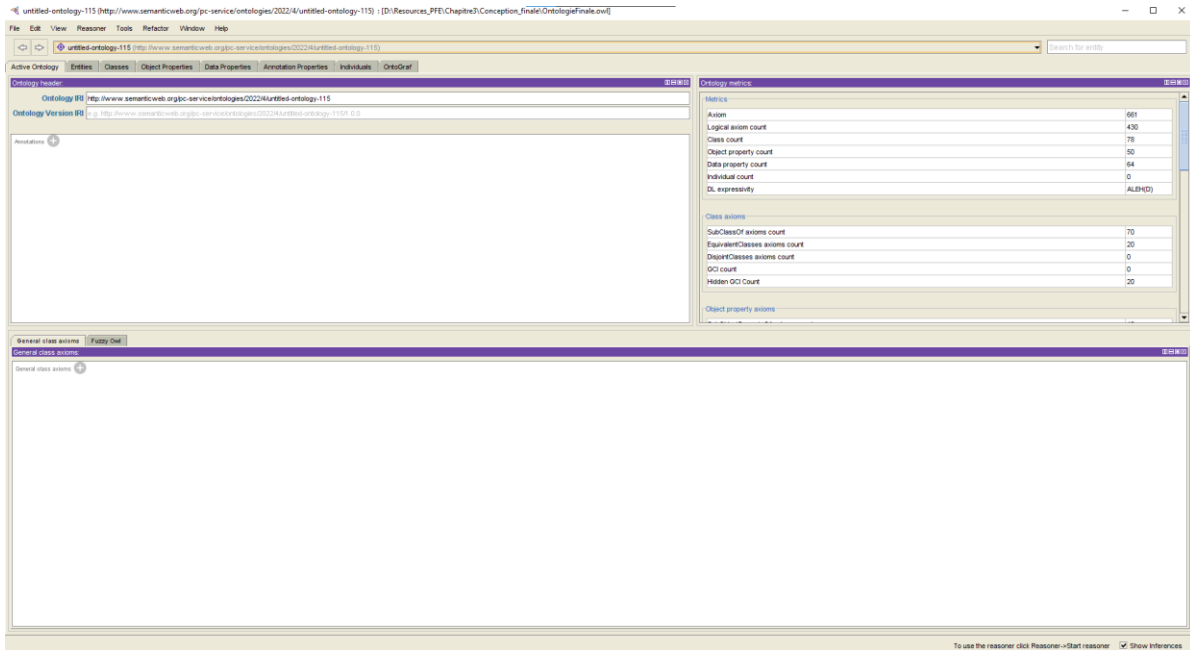


FIGURE 4.1 – Interface graphique de Protégé

## Définition des concepts

La figure 4.2 représente la hiérarchie des concepts de notre ontologie.

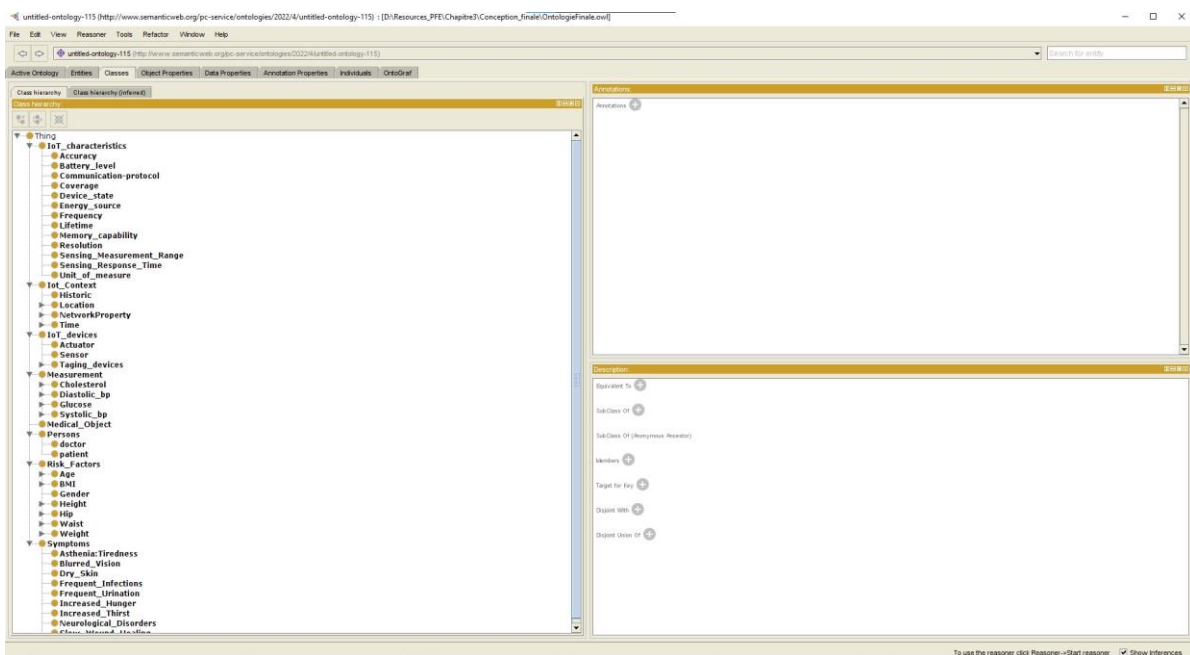


FIGURE 4.2 – Hiérarchie des concepts

## \_\_Définition des relations entre les concepts

La figure 4.3 représente les relations (propriétés de l'objet) entre les concepts.

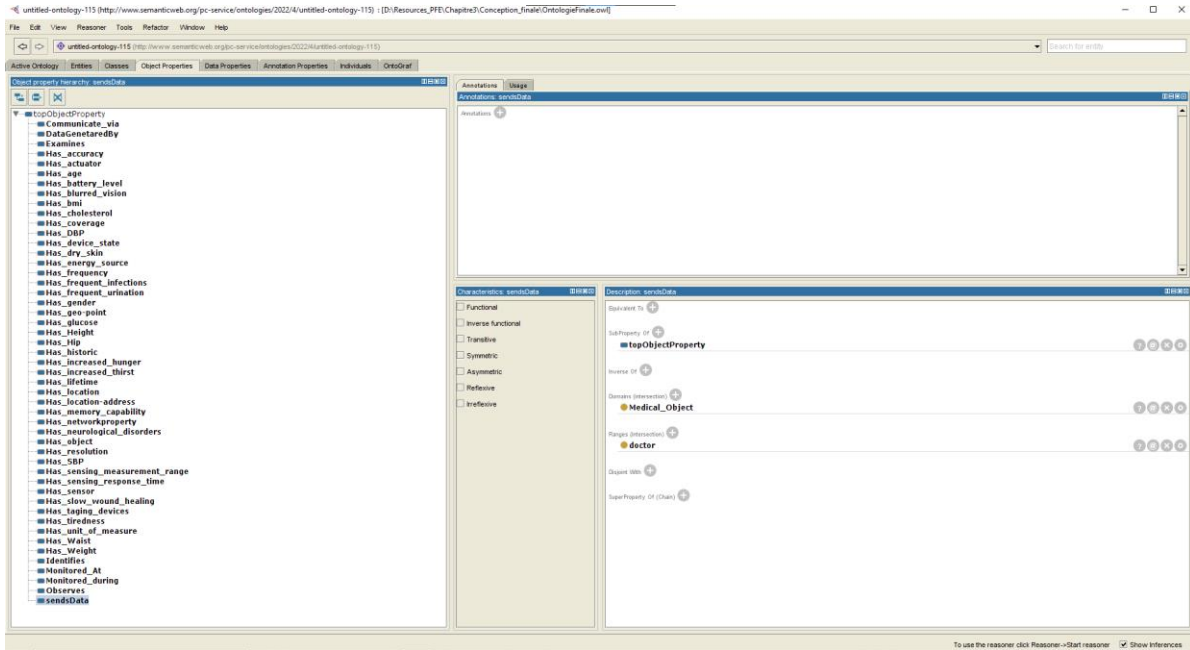


FIGURE 4.3 – Relations entre les concepts

## \_\_Définition des attributs des concepts

La figure 4.4 représente les attributs des concepts.

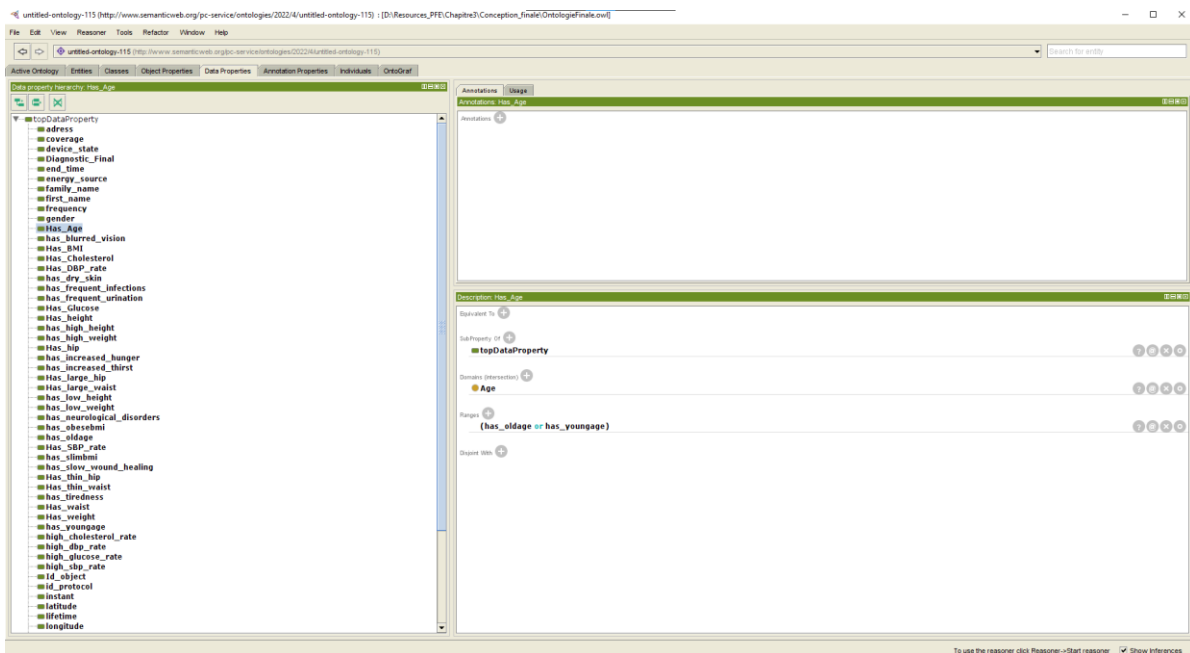


FIGURE 4.4 – Attributs des concepts

#### **4.2.1.2 Plug-in Fuzzy OWL 2**

Fuzzy OWL 2 est un plug-in conçu pour l'éditeur Protégé, qui permet d'effectuer des opérations sur des ontologies Fuzzy OWL 2 et d'écrire des requêtes grâce au moteur d'inférence **Fuzzy DL**. Il a été développé par Fernando Bobillo et Umberto Straccia en 2011 [26].

Pour installer Fuzzy OWL 2, il suffit de copier le fichier "FuzzyOWL.jar" dans le dossier "plugins" de Protégé et copier le fichier "dlib" dans le dossier principal de Protégé. Après avoir ouvert Protégé, un nouvel onglet nommé « fuzzy OWL » se rajoute en plus des onglets déjà existants, c'est un moyen de vérifier que le plugin Fuzzy OWL 2 est bel et bien installé.

#### **4.2.1.3 SMILE**

SMILE est un moteur de raisonnement et d'apprentissage et de découverte causale pour les modèles graphiques, on compte parmi eux les réseaux bayésiens. Il s'agit d'une bibliothèque de classes C++ qui peut être intégrée dans un logiciel utilisateur existant via son API dédiée, améliorer les produits utilisateur avec des capacités de modélisation décisionnelles. SMILE est entièrement portable et disponible pour la plupart des plates-formes, du centre de données à l'embarqué. Des wrappers pour SMILE sont disponibles pour permettre de l'utiliser à partir de Java, Python, R, .NET et d'autres environnements de développement.

#### **4.2.1.4 OWLAPI**

"OWLAPI" ou "org.semanticweb.owlapi" est une API (Application Programming Interface) pour Java possédant des méthodes pour créer, et manipuler des ontologies OWL, cette API est disponible depuis 2003, et a été développée à l'université de Manchester[68].

#### **4.2.1.5 GeNIe**

GeNIe Modeler est une interface utilisateur graphique (GUI) pour SMILE Engine qui permet la construction et l'apprentissage interactifs de modèles. Il est écrit pour l'environnement Windows mais peut également être utilisé sur macOS et Linux sous "Wine". Il a été minutieusement testé sur le terrain depuis 1998, et a reçu une large acceptation dans les universités et l'industrie de manière générale, et compte des milliers d'utilisateurs dans le monde entier.

#### **4.2.1.6 JAVA**

Java est un langage de programmation orienté-objet et une plate-forme de calcul lancé par Sun Microsystems en 1995. Une grande partie du monde numérique dépend de Java : de nombreux services et applications reposent sur cette plate-forme fiable. De même, de nouveaux produits et services numériques innovants et tournés vers l'avenir dépendent aussi de Java.

#### **4.2.1.7 IntelliJ IDEA**

IntelliJ IDEA est un environnement de développement intégré (IDE) écrit en Java pour le développement de logiciels informatiques écrits en Java, Kotlin, Groovy et d'autres langages

"jar". Il est développé par JetBrains (anciennement connu sous le nom d'IntelliJ), et est disponible en tant qu'édition communautaire sous licence Apache 2, et en édition commerciale propriétaire. Les deux peuvent être utilisés pour le développement commercial. L'IDE fournit certaines fonctionnalités comme la complétion de code en analysant le contexte, la navigation de code qui permet de passer directement à une classe ou à une déclaration dans le code, la refactorisation de code, le débogage de code, le linting (analyse du code source pour trouver des erreurs de programmation) et les options pour corriger les incohérences via des suggestions.

## 4.3 Interfaces utilisateur (UI)

Dans ce qui suit, nous dévoilerons les différentes interfaces graphiques qui composent le système.

### 4.3.1 Interface l'accueil

L'interface principale de notre système, ludique et simple à appréhender, elle met en évidence les différentes fonctionnalités du système. Les boutons affichés permettent de naviguer vers les différentes pages et fonctionnalités (voir la **figure 4.5**).

**Bouton "Chercher"** : permet de rechercher un patient en introduisant son nom dans le champ de texte à gauche.

**Bouton "Afficher"** : permet d'afficher la liste des patients sauvegardés.

**Bouton "Ajouter"** : permet d'ouvrir une interface d'ajout pour permettre à l'utilisateur d'introduire un nouveau patient avec tous ses informations nécessaires et de le stocker dans l'ontologie floue.

**Bouton "Modifier"** : permet d'ouvrir une interface de modification pour permettre à l'utilisateur de modifier les informations d'un patient précis.

**Bouton "Supprimer"** : permet d'afficher un message de confirmation si l'utilisateur veut supprimer le patient, après la confirmation de la suppression, nous allons supprimer le patient et ses informations de l'ontologie floue.

**Bouton "Inférence"** : permet d'ouvrir une interface d'inférence pour permettre à l'utilisateur de prédire et afficher les probabilités d'appartenance d'un patient aux deux cas positif et négatif au diabète.

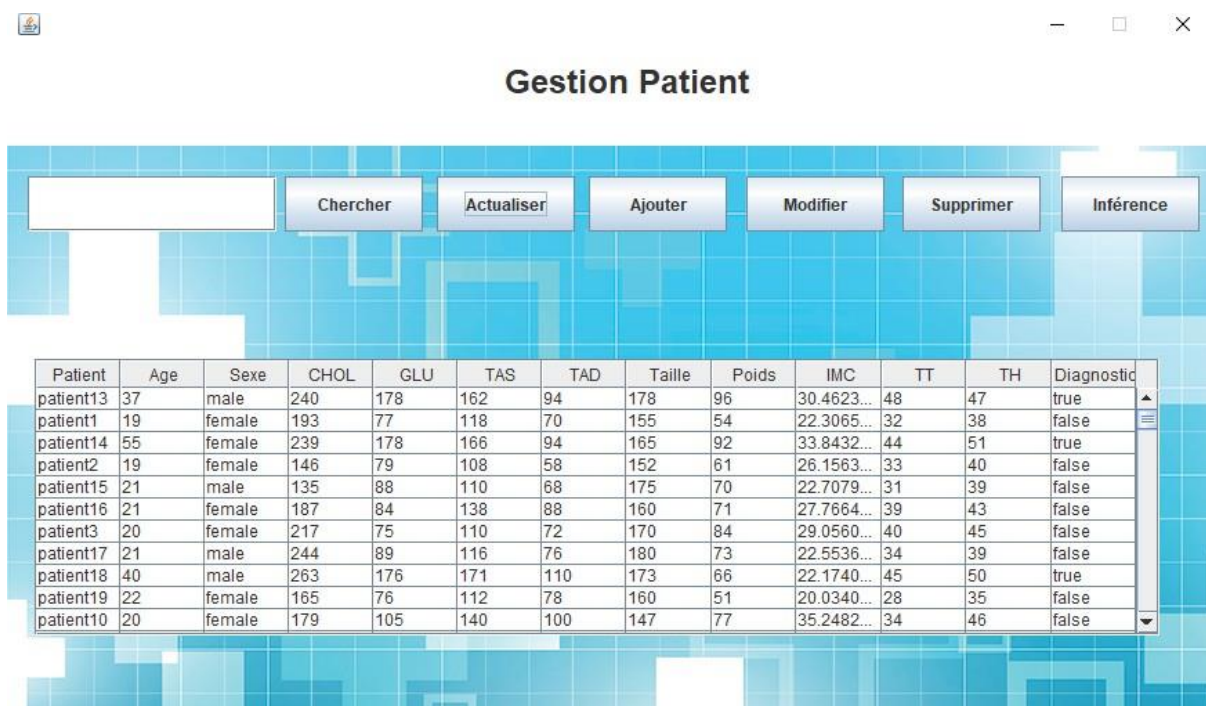


FIGURE 4.5 – Interface d'accueil

### 4.3.2 Interface d'ajout de patient

Cette interface est affichée après avoir cliqué sur le bouton "ajouter" de l'interface d'accueil, elle permet d'ajouter un nouveau patient en fournissant des informations sur ce dernier tels que le nom et les différentes données nécessaires à l'inférence. Ces données seront automatiquement chargées dans l'ontologie floue en cliquant sur le bouton "Ajouter" de l'interface, elle contient aussi un bouton "Accueil" pour revenir à l'interface principale (voir la **figure 4.6**).

**Ajouter un patient**

Nom :  TAS :  TT :

Age :  TAD :  TH :

Sexe : Homme  Taille :

CHOL :  Poids :

GLU :  IMC :

FIGURE 4.6 – Interface d’ajout de patient

### 4.3.3 Interface de modification de patient

Cette interface est affichée après avoir cliqué sur le bouton "modifier" de l’interface d’accueil (voir la **figure 4.5**), elle permet de modifier les informations d’un patient déjà présent dans l’ontologie floue, et ce en remplissant les nouvelles valeurs dans les champs réservés à ces dernières, et en cliquant sur le bouton "Modifier" (voir la **figure 4.7**).

**Modifier un patient**

Nom :  TAS :  TT :

Age :  TAD :  TH :

Sexe : Homme  Taille :

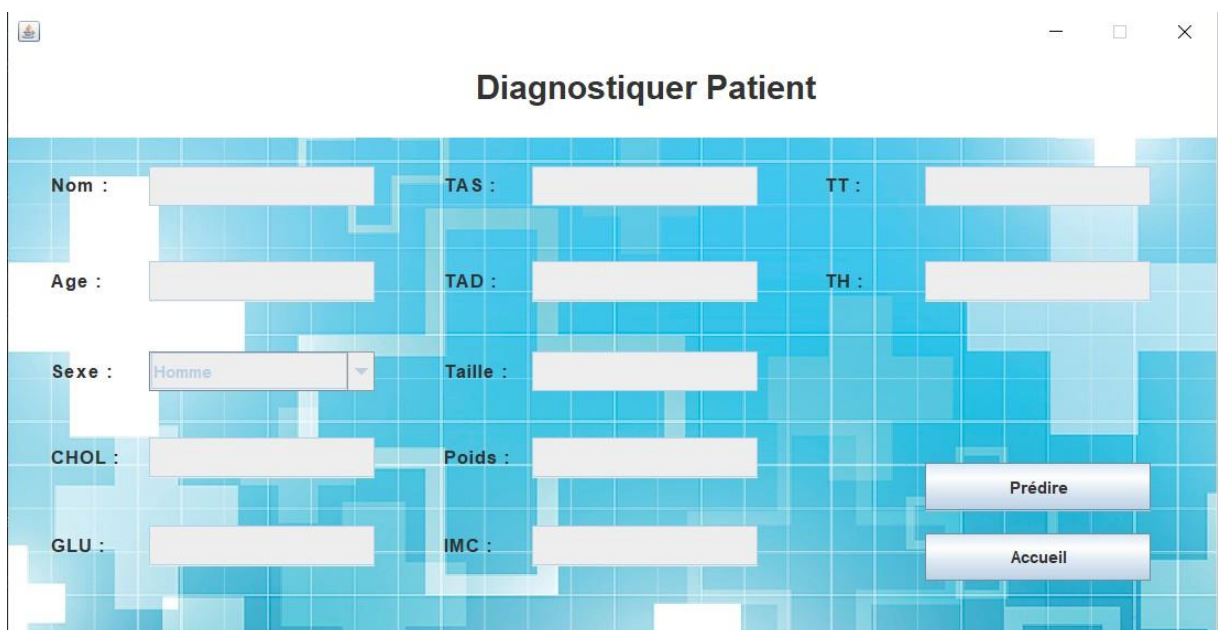
CHOL :  Poids :

GLU :  IMC :

FIGURE 4.7 – Interface de modification de patient

### 4.3.4 Interface d'inférence patient

Si un utilisateur souhaite lancer une prédiction, il doit se rendre sur l'interface d'accueil, puis doit sélectionner un patient sur lequel il souhaite faire un diagnostic à partir du tableau, ensuite il lui suffit de cliquer sur le bouton "inférence" (voir la **figure 4.8**), l'interface d'inférence lui sera affichée et il ne lui reste qu'à appuyer sur le bouton "Prédire" pour confirmer l'action (voir la **figure 4.8**), cette opération va afficher l'interface des résultats (voir la **figure 4.8**).



The screenshot shows a window titled "Diagnostiquer Patient". The interface contains several input fields for patient information: "Nom", "Age", "Sexe" (with a dropdown menu showing "Homme"), "CHOL", "GLU", "TAS", "TAD", "Taille", "Poids", and "IMC". There are also three empty input fields labeled "TT", "TH", and "IMC". On the right side, there are two buttons: "Prédire" and "Accueil". The background of the form area has a light blue grid pattern.

FIGURE 4.8 – Interface d'inférence de patient

### 4.3.5 Interface de résultats

Cette interface permet d'afficher le résultat de l'inférence précédemment évoquée (voir la **figure 4.9**)





FIGURE 4.9 – Interface de résultat

## 4.4 Test et Validation

Afin de tester et d'évaluer les performances de notre système, différentes mesures doivent être définies, nous commencerons cette sous-section par présenter ces mesures et quelques notions essentielles pour mener l'évaluation, ensuite nous montrerons les résultats obtenus.

### 4.4.1 Mesures d'évaluation

L'outil utilisé pour mesurer les performances de notre système qui est avant tout un système de prédiction sera la table de confusion aussi appelée table de contingence, celle-ci résume les résultats de prédiction en vérifiant notamment à quelle fréquence ses prédictions sont exactes par rapport à la réalité. Les éléments de ce tableau sont obtenus en comparant le résultat de prédiction avec la classification réelle sur des données choisies, qui sont différentes des données utilisées pour l'apprentissage du modèle qui est le RBF dans ce cas.

Le **tableau 4.1** présente la matrice de confusion connue du machine learning :

		Classe réelle	
		- (Non diabétique)	+ (Diabétique)
Classe prédite	-	TN (Vrais négatifs)	FN (Faux négatifs)
	+	FP (Faux positifs)	TP (Vrais positifs)

TABLE 4.1 – Matrice de confusion

Ce tableau se compose de 4 éléments :

-**TP (Vrais positifs)** : les cas où la prédiction et la valeur réelle sont toutes les deux positives.

Exemple : le patient est jugé diabétique, et c'est bel et bien la vérité.

-**TN (Vrais négatifs)** : les cas où la prédiction et la valeur réelle sont toutes les deux négatives.

Exemple : le patient est jugé non diabétique, et c'est bel et bien la vérité.

-**FP (Faux positifs)** : les cas où la prédiction est positive, mais où la valeur réelle est négative.

Exemple : le patient est jugé diabétique, alors qu'il n'est pas diabétique en réalité.

-**FN (Faux négatifs)** : les cas où la prédiction est négative, mais où la valeur réelle est positive.

Exemple : le patient est jugé non diabétique, alors qu'il est diabétique en réalité.

#### 4.4.2 Résultats obtenus

Afin de mener à bien les tests de notre système, la méthode choisie pour notre système d'inférence est la validation croisée (K-folds), qui permet d'utiliser l'ensemble des données du dataset sur les patients à notre disposition pour l'entraînement et le test en même temps.

Le principe de la méthode est le suivant : d'abord, nous allons diviser le dataset en k parties, aussi appelées **fold**s, à chaque itération de cette méthode, nous allons choisir une des parties obtenues et la désigner comme partie de test, et le reste (c'est à dire l'union des autres parties k-1) sera utilisé pour entraîner le modèle.

Nous avons mené un ensemble de tests pour plusieurs valeurs de K, et pour chaque K choisi, nous avons appliqué une validation croisée k-fold pour notre RBF et pour un RB classique ayant

la même structure (au niveau des arcs), des tests ont été faits avec l'inférence d'un algorithme d'apprentissage supervisé du machine learning appelé "Gaussian Naive Bayes" exécuté aussi avec les mêmes valeurs de K. Après cela, nous avons calculé les critères de performances pour comparer les résultats des tests sur notre système, avec ceux du RB classique et de l'algorithme "Gaussian Naive Bayes".

Les critères de performances mises en oeuvre sont :

**-L'accuracy** : Représente le pourcentage des bonnes prédictions, c'est un indicateur simple et l'un des plus utilisées pour l'évaluation des performances. L'accuracy doit être élevée pour prouver que le système fait de bonnes prédictions avec un taux d'erreurs minimale.

Formellement, la formule en est ainsi :

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

**-Le rappel(recall)** : Est le ratio des instances positives jugées correctement par rapport à toutes les instances positives. Cet indicateur est très important car la baisse de celui ci signifie que le nombre des faux négatifs est élevé, dans notre cas d'étude qui est une maladie on doit donc minimiser ces faux négatifs où le nombre de personnes diabétiques qui ne sont pas jugées ainsi. La formule du rappel est :

$$Recall = \frac{TP}{TP + FN}$$

**-La précision** : La précision est le ratio des instances positives jugées correctement par rapport à toutes les instances jugées positives. Une précision basse signifie qu'il y a beaucoup de personnes non diabétiques qui sont jugées comme diabétiques, on ne peut non plus tolérer ce type d'erreurs, donc ce critère est très important lui aussi.

La formule de la précision est :

$$Precision = \frac{TP}{TP + FP}$$

#### 4.4.2.1 Présentation des résultats :

Maintenant que le choix de ces critères est justifié, nous présentons ci-dessous les résultats obtenus après les tests pour ces derniers.

Le **tableau 4.2** représente la comparaison des accuracy par les classifieurs pour les différentes valeurs de K :

K	Notre système	k fold avec un RB classique	Gaussian Naive Bayes
2	<b>77,40107727</b>	76,47995758	71,3761
3	75,80645243	<b>79,74910736</b>	78,003
4	<b>86,16718292</b>	84,21066093	82,6349
5	<b>89,9642868</b>	86,18181763	84,601
6	<b>89,78494517</b>	86,55913798	85,3294
7	<b>90,50824193</b>	86,52014596	86,0375
8	<b>90,33088017</b>	86,34979153	86,3975
9	<b>90,68100315</b>	88,35125393	86,397
10	<b>90,68783188</b>	87,83068695	87,1104
moyenne	<b>86,81465575</b>	84,69250665	83,09853333

TABLE 4.2 – Tableau de comparaison des 'accuracy' (%) obtenues par chaque classifieur

La **figure 4.10** illustre un histogramme pour les valeurs d'accuracy en fonction des K pour les 3 classifieurs :

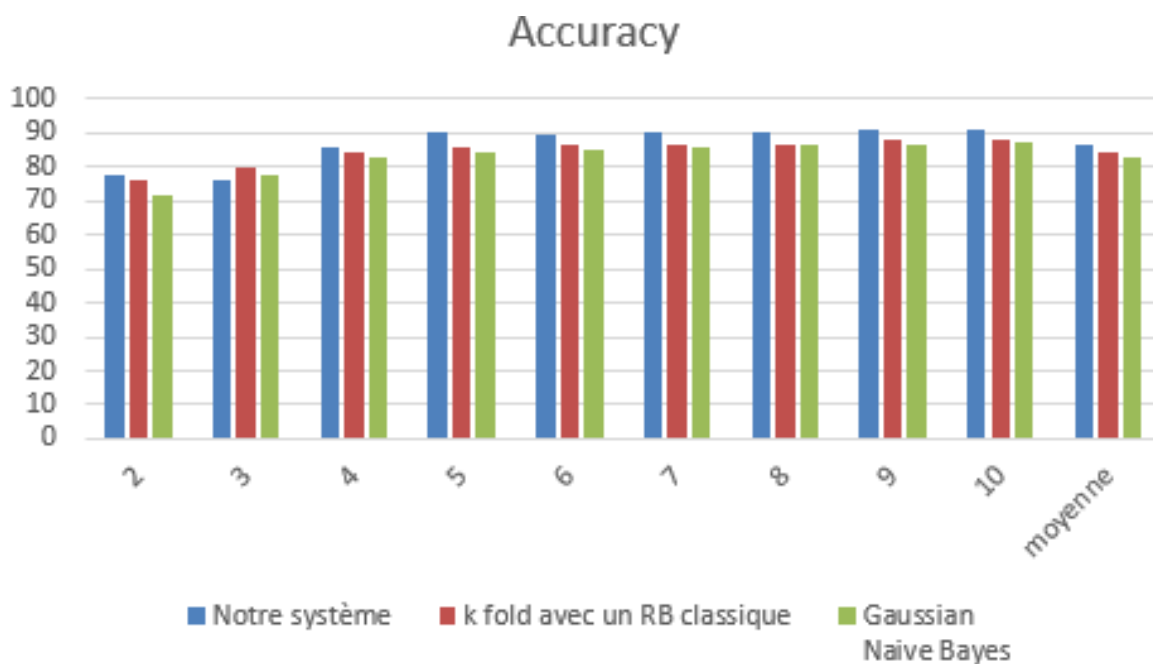


FIGURE 4.10 – Accuracy obtenus par chaque classifieur sous forme de graphe

Le **tableau 4.3** représente la comparaison des rappels par les classifieurs pour les différentes valeurs de K :

K	Notre système	k fold avec un RB classique	Gaussian Naive Bayes
2	<b>81,28774261</b>	66,59873199	73,52
3	77,41354243	<b>83,89560699</b>	78,33
4	<b>84,29096985</b>	81,30155945	83,8
5	<b>88,2421051</b>	82,55374603	83,7
6	<b>87,33502579</b>	83,90348689	84,54
7	<b>87,90729305</b>	82,9046075	86,01
8	<b>88,23221874</b>	81,02326107	86
9	<b>87,32527415</b>	84,24256643	86,71
10	<b>86,42102432</b>	83,05897942	83,39
moyenne	<b>85,38391067</b>	81,0536162	82,88888889

TABLE 4.3 – Tableau de comparaison des 'rappels' (%) obtenues par chaque classifieur

La **figure 4.11** illustre un histogramme pour les valeurs de rappel en fonction des K pour les 3 classifieurs :

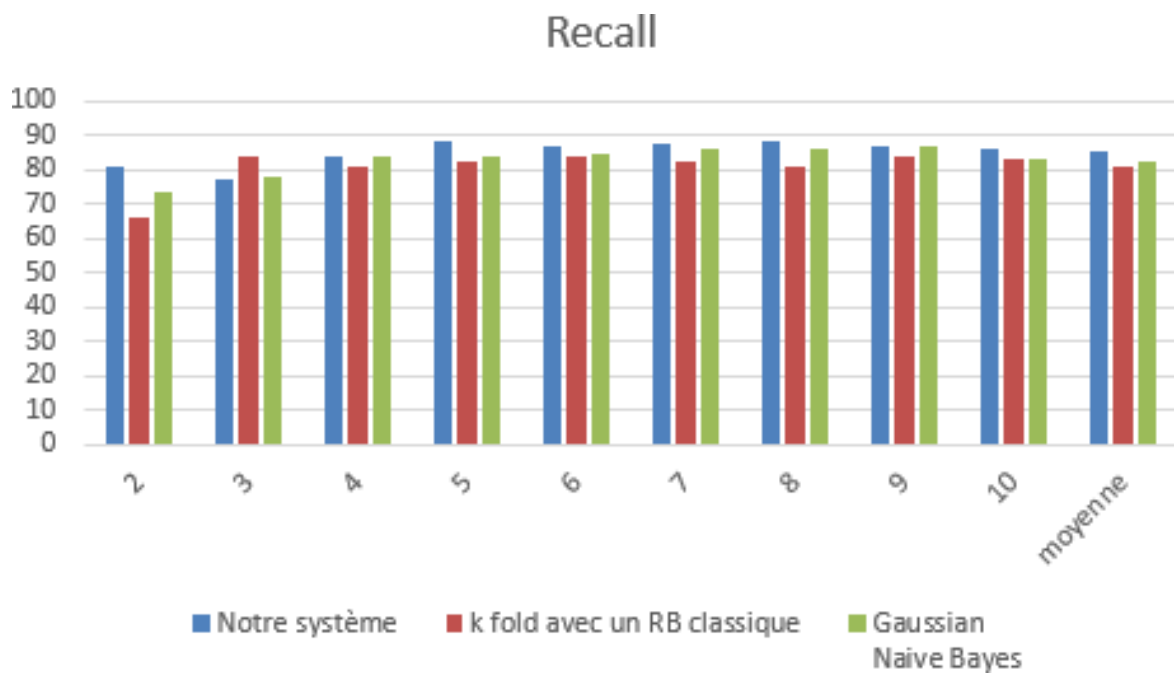


FIGURE 4.11 – Rappels obtenus par chaque classifieur sous forme de graphe

Le **tableau 4.4** représente la comparaison des précisions par les classifieurs pour les différentes valeurs de K :

K	Notre système	k fold avec un RB classique	Gaussian Naive Bayes
2	76,32769394	<b>84,26470566</b>	76,82
3	79,24687831	76,9003245	<b>82,75</b>
4	<b>86,18225098</b>	81,42091751	84,31
5	<b>89,0799408</b>	84,85438995	86,64
6	<b>87,56818263</b>	83,32218043	86,91
7	<b>89,8550742</b>	82,68179921	86,44
8	<b>88,59578037</b>	80,02795792	86,66
9	<b>88,04012426</b>	82,75110711	86,24
10	<b>89,2713501</b>	79,66460571	86,66
moyenne	<b>86,01858617</b>	81,765332	84,82555556

TABLE 4.4 – Tableau de comparaison des 'précisions' (%) obtenues par chaque classifieur

La **figure 4.12** illustre un histogramme pour les valeurs de précision en fonction des K pour les 3 classifieurs :

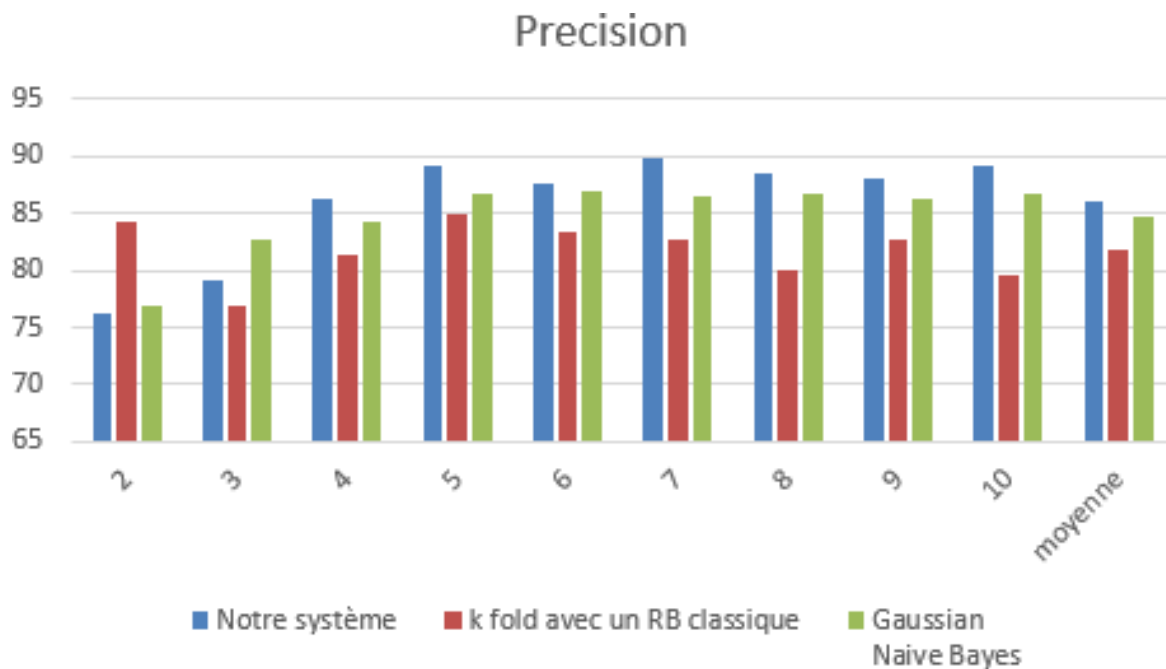


FIGURE 4.12 – Précisions obtenues par chaque classifieur sous forme de graphe

#### 4.4.2.2 Analyse et confirmation des résultats

Comme nous pouvons le voir à propos des résultats obtenus, notre système a présenté de meilleurs résultats que les 2 autres classifieurs au niveau des trois critères évalués, les moyennes calculées prouvent cette affirmation :

-Selon la **figure 4.11** : notre système possède une accuracy moyenne de 86,81% qui est supérieure aux deux autres moyennes du RB classique et de "Gaussian Naive Bayes" qui sont respectivement de 84,69% et 83,09%.

-Selon la **figure 4.13** : notre système possède un rappel avec une moyenne de 85,38% qui est supérieure aux deux autres moyennes du RB classique et de "Gaussian Naive Bayes" qui sont respectivement de 81,05% et 82,88%.

-Selon la **figure 4.15** : notre système possède une precision moyenne de 86,01% qui est supérieure aux deux autres moyennes du RB classique et de "Gaussian Naive Bayes" qui sont respectivement de 81,76% et 84,82%.

De ce qui est du choix de la valeur du K, la valeur de K ne doit être ni trop basse ni trop haute, en pratique on choisit généralement une valeur comprise entre 5 et 10 en fonction de l'envergure du dataset, car le choix d'une valeur trop basse ou trop haute peut mener à un sous-apprentissage (underfitting) ou sur-apprentissage (overfitting) respectivement.

En effet, c'est dans la plage de valeurs de K entre 5 et 10 que les performances sont les plus optimales (qui restent aussi supérieurs aux autres classifieurs). Pour le K égal à 5, l'accuracy de système notre est de 89,96% face à 86,18% d'accuracy pour le RB classique et 84,60%



d'accuracy pour le "Gaussian Naive Bayes", le rappel est de 88,24% face à 82,55% de rappel pour le RB classique et 83,7% de rappel pour le "Gaussian Naive Bayes", et la précision et de 89,07% face à 84,85% de précision pour le RB classique et 86,64% de précision pour le "Gaussian Naive Bayes".

Pour le K égal à 10, l'accuracy de système notre est de 90,68% face à 87,83% d'accuracy pour le RB classique et 87,11% d'accuracy pour le "Gaussian Naive Bayes", le rappel est de 86,42% face à 83,05% de rappel pour le RB classique et 83,39% de rappel pour le "Gaussian Naive Bayes", et la précision et de 89,27% face à 79,66% de précision pour le RB classique et 86,66% de précision pour le "Gaussian Naive Bayes".

Pour conclure, un bon modèle doit avoir des critères de performance proches de 100 %, l'accuracy, le rappel, et la précision de notre système sont proches de la valeur de perfection, ce qui signifie que les prédictions basées sur notre système sont proches de la réalité c'est à dire des prédictions d'un médecin.

## **4.5 Conclusion**

Dans ce chapitre, nous avons énuméré tous les outils utilisés dans ce travail incluant l'environnement de développement, le langage de programmation et les bibliothèques utilisées ainsi que les outils de création de l'ontologie floue et du réseau bayésien flou. Ensuite, nous avons présenté les interfaces de notre application et leurs rôles. Et pour clôturer le travail, nous avons évalué notre système afin de le valider.



---

## Conclusion générale

---

L'objectif de notre travail était la proposition et l'implémentation d'un système reposant sur un nouveau modèle sémantique qui permet de représenter les connaissances médicales incertaines dans le domaine de l'IoMT.

Le nouveau modèle sémantique proposé a été conçu à l'aide d'une ontologie floue, nous avons suivi une méthode récente de développement d'ontologie complète, dans la mesure où en partant des données brutes représentées comme une source de connaissances, il permet d'arriver à une ontologie opérationnelle. Pour ce faire, la méthodologie « FODM » a été employée afin d'explicitier et de guider grâce à ses différentes phases la construction de l'ontologie floue depuis la collection des termes du domaine jusqu'à la validation de l'ontologie. À l'aide d'un l'expert du domaine, nous avons modélisé et construit la structure du RBF en énumérant les facteurs de risques ainsi que les liens de causalité entre eux. L'apprentissage des paramètres a été fait à partir d'une base d'apprentissage (dataset utilisé) des données complètes en utilisant l'approche maximum de vraisemblance (MV). Ensuite pour arriver à un réseau bayésien flou, chacun des noeuds à été transformé en un noeud flou pour ensuite l'utiliser pour l'inférence.

Une fois le système implémenté, nous avons mis en place des tests permettant la vérification des performances en diagnostiquant par inférence plusieurs patients atteints ou pas du diabète de type 2, tout en ayant leurs informations biologiques et leurs classifications réelles pour la comparaison des résultats.

Afin d'enrichir ce projet et dans le souci d'élargir et d'améliorer notre travail, nous proposons l'amélioration de ce modèle sémantique au contexte du diagnostic médical par le biais de l'IoMT par :

- Expérimenter ce système dans des situations réelles pour ajuster les paramètres de l'incertitude.
- Expérimentation du système dans le cadre d'un suivi à domicile pour les patients atteints du diabète de type 2.
- Enrichissement de l'ontologie par d'autres maladies en ajoutant leurs symptômes, et leurs facteurs d'incertitude appropriés.
- Ajouter une extension à la couche application pour permettre de capter des données de patients grâce à des objets médicaux connectés.



---

# Bibliographie

---

- [1] Kenneth J Laskey and Kathryn B Laskey. Uncertainty reasoning for the world wide web : Report on the urw3-xg incubator group. In *URSW*. Citeseer, 2008.
- [2] Hanène Ghorbel, Afef Bahri, and Rafik Bouaziz. Fuzzy protégé for fuzzy ontology models. *Age*, 12(18) :30, 2009.
- [3] Xin Li, José-Fernán Martínez, and Gregorio Rubio. A new fuzzy ontology development methodology (fodm) proposal. *IEEE Access*, 4 :7111–7124, 2016.
- [4] M Han and H Zhang. Business intelligence architecture based on internet of things. *Journal of Theoretical & Applied Information Technology*, 50(1) :90–95, 2013.
- [5] Amy Nordrum et al. Popular internet of things forecast of 50 billion devices by 2020 is outdated. *IEEE spectrum*, 18(3), 2016.
- [6] MIMI Amani. Suivi médicale à distance dans l’internet des objets.
- [7] Ferhane Souhayla. *l’internet des objets révolutionne notre vie quotidienne : application pour une maison intelligent*. PhD thesis, Universite laarbi tebessi tebessa, 2021.
- [8] Pierre-Jean Benghozi, Sylvain Bureau, and Françoise Massit-Folea. L’internet des objets. quels enjeux pour les européens ? 2008.
- [9] SYNOXgroup. 4 choses à savoir sur l’internet des objets disponible sur : [https ://www.synox.io/4-choses-a-savoir-sur-linternet-des-objets/](https://www.synox.io/4-choses-a-savoir-sur-linternet-des-objets/).
- [10] DAHMANE Abir BOUZIDI Meriem ELABED Sid et al. *Implémentation d’n Réseau Prototype pour l’Internet des Objets*. PhD thesis, Faculté des Sciences et Technologies, 2021.
- [11] John A Stankovic. Wireless sensor networks. *computer*, 41(10) :92–95, 2008.
- [12] N Daniel, R Marcel, and K Daniel. Livre blanc machine to machine enjeux et perspectives : Orange business services. *Syntec informatique, Fing*, 2006.

- [13] Nawaf Alharbe, Anthony S Atkins, and Akbar Sheikh Akbari. Application of zigbee and rfid technologies in healthcare in conjunction with the internet of things. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, pages 191–195, 2013.
- [14] Ricardo Neisse, Maarten Wegdam, and Marten van Sinderen. Trust management support for context-aware service platforms. In *User-Centric Networking*, pages 75–106. Springer, 2014.
- [15] Benjamin Fabian. Secure name services for the internet of things. 2008.
- [16] Shoumen Palit Austin Datta. L’internet des objets : la troisième révolution industrielle. *Logistique & Management*, 23(3) :29–33, 2015.
- [17] Gulraiz J Joyia, Rao M Liaqat, Aftab Farooq, and Saad Rehman. Internet of medical things (iomt) : Applications, benefits and future challenges in healthcare domain. *J. Commun.*, 12(4) :240–247, 2017.
- [18] Abdeltif Elbyed. *ROMIE, une approche d’alignement d’ontologies à base d’instances*. PhD thesis, Institut National des Télécommunications, 2009.
- [19] Imane GHERZOULI. *Système d’aide au diagnostic médical à base d’ontologie*. PhD thesis, FACULTE DES MATHEMATIQUES ET D’INFORMATIQUE-UNIVERSITE MOHAMED BOUDIAF-M’SILA, 2017.
- [20] Tim Berners-Lee. J, handler, o lassila. the semantic web. *Scientific American*, 284(5) :35–43, 2001.
- [21] Jean Charlet, Bruno Bachimont, and Raphaël Troncy. Ontologies pour le web sémantique. *Revue I3, numéro Hors Série «Web sémantique*, pages 43–63, 2004.
- [22] Michel Klein. Combining and relating ontologies : an analysis of problems and solutions. In *OIS@ IJCAI*, 2001.
- [23] Nonita Sharma, Monika Mangla, Sachi Nandan Mohanty, Deepak Gupta, Prayag Tiwari, Mohammad Shorfuzzaman, and Majdi Rawashdeh. A smart ontology-based iot framework for remote patient monitoring. *Biomedical Signal Processing and Control*, 68 :102717, 2021.
- [24] Vinoth Kumar et al. Ontology based public healthcare system in internet of things (iot). *Procedia Computer Science*, 50 :99–102, 2015.
- [25] Faisal Alsubaei, Abdullah Abuhussein, and Sajjan Shiva. Ontology-based security recommendation for the internet of medical things. *IEEE Access*, 7 :48948–48960, 2019.

- [26] Frances J Milliken. Three types of perceived uncertainty about the environment : State, effect, and response uncertainty. *Academy of Management review*, 12(1) :133–143, 1987.
- [27] ISHAK RIALI. *Modélisation de la connaissance incertaine dans le web sémantique*. PhD thesis, Université Blida1-Saad Dahlab.
- [28] Pat Croskerry. The theory and practice of clinical decision-making. *Canadian Journal of Anesthesia*, 52(1) :R1–R8, 2005.
- [29] Suzette Cooke and Jean-Francois Lemay. Transforming medical assessment : integrating uncertainty into the evaluation of clinical reasoning in medical education. *Academic Medicine*, 92(6) :746–751, 2017.
- [30] Pat Croskerry. A model for clinical decision-making in medicine. *Medical science educator*, 27(1) :9–13, 2017.
- [31] Marij A Hillen, Caitlin M Gutheil, Tania D Strout, Ellen MA Smets, and Paul KJ Han. Tolerance of uncertainty : Conceptual analysis, integrative model, and implications for healthcare. *Social Science & Medicine*, 180 :62–75, 2017.
- [32] Paul KJ Han, Daniel Schupack, Susannah Daggett, Christina T Holt, and Tania D Strout. Temporal changes in tolerance of uncertainty among medical students : insights from an exploratory study. *Medical education online*, 20(1) :28285, 2015.
- [33] Vera P Luther and Sonia J Crandall. Commentary : ambiguity and uncertainty : neglected elements of medical education curricula ? *Academic Medicine*, 86(7) :799–800, 2011.
- [34] CJ Ledford, Dean A Seehusen, Alexander W Chessman, and Navkiran K Shokar. How we teach us medical students to negotiate uncertainty in clinical care. *Family medicine*, 47(1) :31–36, 2015.
- [35] Erica S Spatz, Harlan M Krumholz, and Benjamin W Moulton. Prime time for shared decision making. *Jama*, 317(13) :1309–1310, 2017.
- [36] Geoffrey Norman. Research in clinical reasoning : past history and current trends. *Medical education*, 39(4) :418–427, 2005.
- [37] Lotfi A Zadeh. Fuzzy sets. In *Fuzzy sets, fuzzy logic, and fuzzy systems : selected papers by Lotfi A Zadeh*, pages 394–432. World Scientific, 1996.
- [38] Arthur Kordon. *Applying computational intelligence : how to create value*. Springer Science & Business Media, 2009.
- [39] Elie Sanchez. *Fuzzy logic and the semantic web*. Elsevier, 2006.

- [40] Fernando Bobillo, Miguel Delgado, and Juan Gómez-Romero. Representation of context-dependent relevance relations with fuzzy ontologies. In *Proceedings of the 1st Workshop on Advancing Reasoning on the Web : Scalability and Commonsense (ARea2008)*, 2008.
- [41] Umberto Straccia. Fuzzy alc with fuzzy concrete domains. In *Proceedings of the International Workshop on Description Logics (DL-05)*, volume 147, pages 96–103. Citeseer, 2005.
- [42] Chaour Moundir and Kermani Mohamed Hachem. Développement d’une ontologie floue et classification d’instances : Application au domaine des analyses médicales, projet de fin d’études pour l’obtention du diplôme de master recherche en informatique, université constantine 2. 2014.
- [43] Ross EG Upshur. A short note on probability in clinical medicine. *Journal of Evaluation in Clinical Practice*, 19(3) :463–466, 2013.
- [44] Glenn Shafer and Vladimir Vovk. The sources of kolmogorov’s grundbegriffe. *Statistical Science*, 21(1) :70–98, 2006.
- [45] Alan Hájek. Interpretations of probability. 2002.
- [46] Valentin Roussel. Introduction à l’inférence bayésienne-cours 2 : le théorème de bayes. 2019.
- [47] AlainMILLE. Réseaux bayésiens, recherche connaissance et raisonnement. 2006.
- [48] Dan Geiger and David Heckerman. Knowledge representation and inference in similarity networks and bayesian multinets. *Artificial Intelligence*, 82(1-2) :45–74, 1996.
- [49] Olivier François and Philippe Leray. Etude comparative d’algorithmes d’apprentissage de structure dans les réseaux bayésiens. *Rencontres des Jeunes Chercheurs en IA*, 2003.
- [50] Tayeb Kenaza. *Modèles graphiques probabilistes pour la corrélation d’alertes en détection d’intrusions*. PhD thesis, Artois, 2011.
- [51] Ph Leray. *Réseaux bayésiens : Apprentissage et diagnostic de systemes complexes*. PhD thesis, Université de Rouen, 2006.
- [52] Morgan Chopin and Pierre-Henri Wuillemin. Optimisation de l’inférence dans les réseaux bayésiens dynamiques. *LIP6, université de paris, France. ROADEF*, 2010.
- [53] Ali Ben Mrad, Véronique Delcroix, Sylvain Piechowiak, Philip Leicester, and Mohamed Abid. An explication of uncertain evidence in bayesian networks : likelihood evidence and probabilistic evidence. *Applied Intelligence*, 43(4) :802–824, 2015.



- [54] Zhongli Ding, Yun Peng, and Rong Pan. A bayesian approach to uncertainty modeling in owl ontology. Technical report, MARYLAND UNIV BALTIMORE DEPT OF COMPUTER SCIENCE AND ELECTRICAL ENGINEERING, 2006.
- [55] P.C.G.costa. Bayesian semantics for the semantic web. 2005.
- [56] Rommel Novaes Carvalho. *Probabilistic ontology : representation and modeling methodology*. George Mason University, 2011.
- [57] N Senthil Selvan, Subramaniaswamy Vairavasundaram, and Logesh Ravi. Fuzzy ontology-based personalized recommendation for internet of medical things with linked open data. *Journal of Intelligent & Fuzzy Systems*, 36(5) :4065–4075, 2019.
- [58] Rajani Reddy Gorrepati, Do-Hyeun Kim, and Sitaramanjaneya Reddy Guntur. Ubiquitous internet of medical things on interoperable ssn ontology platforms for e-health monitoring of connected objects. 2021.
- [59] Milad Lesani, Marjan Naderan, and Seyed Enayatallah Alavi. Fuzzy ontology with anfis neural network for semantic sensor networks in smart homes based on internet of things. *International Journal of Web Research*, 2(1) :26–38, 2019.
- [60] Yazid Benazzouz, Ioannis Parissis, et al. A fault fuzzy-ontology for large scale fault-tolerant wireless sensor networks. *Procedia Computer Science*, 35 :203–212, 2014.
- [61] Nguyen B Truong, Tai-Won Um, and Gyu Myoung Lee. A reputation and knowledge based trust service platform for trustworthy social internet of things. *Innovations in clouds, internet and networks (ICIN), Paris, France*, pages 104–111, 2016.
- [62] Levin Varghese, Gerard Deepak, and A Santhanavijayan. A fuzzy ontology driven integrated iot approach for home automation. In *International Conference on Digital Technologies and Applications*, pages 271–277. Springer, 2021.
- [63] Sondes Titi, Hadda Ben Elhadj, and Lamia Chaari Fourati. A fuzzy-ontology based diabetes monitoring system using internet of things. In *International Conference on Smart Homes and Health Telematics*, pages 287–295. Springer, 2020.
- [64] Gabriel Machado Lunardi, Fadi Al Machot, Vladimir A Shekhovtsov, Vinícius Maran, Guilherme Medeiros Machado, Alencar Machado, Heinrich C Mayr, and José Palazzo M de Oliveira. Iot-based human action prediction and support. *Internet of Things*, 3 :52–68, 2018.
- [65] M Dali-Sahi, D Benmansour, A Aouar, and N Karam. Type 2 dans des populations endogames de l’ouest algérien. *Leban Sci J*, 13(2) :17, 2012.

- [66] Mickael Canouil. *Développement et application de méthodologies statistiques pour études multi-omiques dans le diabète de type 2 : au-delà de l'ère des études d'association pangénomiques*. PhD thesis, Université du Droit et de la Santé-Lille II, 2017.
- [67] Ignacio Huitzil, Umberto Straccia, Natalia Díaz-Rodríguez, and Fernando Bobillo. Datil : learning fuzzy ontology datatypes. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 100–112. Springer, 2018.
- [68] Matthew Horridge and Sean Bechhofer. The owl api : A java api for owl ontologies. *Semantic web*, 2(1) :11–21, 2011.