

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique Université A. Mira de
Béjaïa
Faculté des Sciences Exactes
Département d'Informatique



Mémoire de fin d'étude

En vue de l'obtention du diplôme de Master Académique en Systèmes
d'information avancés

Thème

Un système de recommandation pour les villes intelligentes

Réalisé par

BOUREDJIOUA Fayçal

TAGZIRT Dyhia

Devant le jury composé de

Professeur AMROUN Kamal : **Président**

Dr CHIBANI Samia : **Examinatrice**

Dr EL BOUHISSI Houda Epse BRAHAMI : **Encadrante**

Mr ZIANE Amine (Doctorant) : **Invité**

Promotion 2022/2023

Remerciements

Tout d'abord, nous remercions Allah le tout puissant de nous avoir donné le courage et la patience nécessaires à mener ce travail à son terme.

Nous tenons à remercier tout particulièrement notre encadrante Mme. **Houda El-Bouhissi** pour l'aide compétente qu'elle nous a apportée, pour sa patience et son encouragement. Son œil critique nous a été très précieux pour structurer le travail et pour améliorer la qualité des différentes sections, la qualité de son suivie ainsi que pour tous les conseils et les informations qu'elle nous a prodigués avec un degré de patience et de professionnalisme sans égal.

Nous tenons aussi à remercier les membres du jury qui trouvent, ici, l'expression de nos sincères remerciements pour l'honneur qu'ils nous font en prenant le temps de lire et d'évaluer ce travail.

Un très grand remerciement pour tous nos professeurs de l'université d'Abd Rahmane Mira pour leur formation de qualité depuis la première année.

Nous souhaitons aussi remercier l'équipe pédagogique et administrative du département informatique pour leurs efforts dans le but de nous offrir une excellente formation.

Nous tenons à remercier Mme. Kahina la formatrice et la directrice de l'école "Amis de JAVA" pour ses formations de qualité, sa disponibilité et ses orientations.

Nous tenons aussi à remercier notre cher ami KOLLI Hamid pour son grand aide et ses contributions qui ont été précieuses et nous lui sommes reconnaissants pour son soutien tout au long du processus.

Pour finir, nous souhaitons remercier toutes les personnes ayant contribué de près ou de loin à la réalisation de ce travail.

Dédicaces

Avec tous mes sentiments de tendresse, d'amour et de respect, je dédie ce travail à :

Ma source de tendresse, de douceur et de bienveillance, ma chère mère Farida, et mon père Mustapha, source de sagesse, de conseils et de fierté. Leurs prières ont sans aucun doute contribué à ma réussite. C'est à eux que je dis « MERCI » et je vous aime. Que Dieu les garde pour moi.

À ma sœur Rima, à mon frère Saïd et à toute la famille Tagzirt.

À mon binôme et ami Fayçal, avec lequel j'ai pris beaucoup de plaisir à travailler. Nous avons formé une belle équipe.

À ma deuxième famille : Lyna, Kahina, Melissa, Kenza, Alicia, Chakib, Rabah, Yanis, Lyes (Eto), MizMiz et tous mes camarades de la section... et en particulier Hamid et sa famille.

Sans oublier tous les professeurs, que ce soit du primaire, du collège, du lycée ou de l'enseignement supérieur.

À tous ceux qui me sont chers, à vous tous, merci.

-Dyhia

Dédicaces

Avec tous mes sentiments de tendresse, d'amour et de respect, je dédie ce travail à :

Ma source de tendresse, de douceur et de bienveillance, ma chère mère Nadia, et mon père Hakim, source de sagesse, de conseils et de fierté. Leurs prières ont sans aucun doute contribué à ma réussite. C'est à eux que je dis « MERCI » et je vous aime. Que Dieu les garde pour moi.

À mon frère Sofiane et ma sœur Rym, à et à toute la famille Bouredjioua et Belaid.

À ma binôme et amie Dyhia, avec laquelle j'ai pris beaucoup de plaisir à travailler. Nous avons formé une belle équipe.

À ma deuxième famille : Billal, Ayman, Riad, Aimad, Walid, Karim, Koukous, Ridha, Zakary, Wassim, Bizak, Kader, Rahim, Karim, Mohamed, Samir, et tous mes camarades de la section... et en particulier Alicia et sa famille.

Sans oublier tous les professeurs, que ce soit du primaire, du collège, du lycée ou de l'enseignement supérieur.

À tous ceux qui me sont chers, à vous tous, merci.

-Fayçal

Table des matières

Remerciements	I
Dédicaces	II
Dédicaces	III
Liste des figures	VIII
Liste des tableaux	X
Liste des abréviation	XII
Chapitre 1 : Introduction générale	1
1.1 Introduction	1
1.2 Motivation	1
1.3 Contributions	3
1.4 Méthodologie de travail	3
1.5 Organisation du mémoire	4
Chapitre 2 : Notions fondamentales	5
2.1 Introduction	5
2.2 Systèmes de recommandation	5
2.3 Types de systèmes de recommandation	6
2.3.1 Filtrage basé sur le contenu	7
2.3.2 Filtrage collaboratif	8

2.3.3	Approche hybride	9
2.3.4	Autres types de recommandations	10
2.4	Défis des systèmes de recommandations	12
2.5	Villes intelligentes	13
2.5.1	Technologies de l'information et de la communication (TIC) et les technologies liées aux villes intelligentes . . .	14
2.5.2	Caractéristiques	14
2.5.3	Dimensions	15
2.5.4	Systèmes de recommandation et villes intelligentes . . .	16
2.5.5	Conclusion	18
Chapitre 3 : L'état de l'art		19
3.1	Introduction	19
3.2	Travaux connexes	20
3.2.1	Filtrage collaboratif	20
3.2.2	Filtrage hybride	22
3.2.3	Filtrage de contenu	25
3.3	Tableaux comparatifs	26
3.4	Discussion et comparaison	29
3.4.1	Mesure de la similarité	30
3.4.2	Matégorisation des utilisateurs ou des éléments	30
3.4.3	Techniques de pondération des termes (TF-IDF)	30
3.4.4	Représentations de réseaux neuronaux	30
3.5	Conclusion	35

Chapitre 4 : Contributions	36
4.1 Introduction	36
4.2 Approche proposée	37
4.2.1 Création du profil utilisateur	39
4.2.2 Recherche de profils similaires	40
4.2.3 Clustering	45
4.2.4 Recommandation	48
4.3 Diagramme d'activité	52
4.4 Conclusion	53
Chapitre 5 : Expérimentation	55
5.1 Introduction	55
5.2 Description du Dataset	55
5.3 Environnement de développement	58
5.3.1 Environnement matériel	58
5.3.2 Environnement logiciel	58
5.4 implémentation	63
5.4.1 Page d'accueil	64
5.4.2 Trouvez un médecin	65
5.4.3 Choisissez un médecin	66
5.4.4 Notez votre médecin	67
5.4.5 A propos	68
5.5 Evaluations	69
5.5.1 Evaluation des algorithmes (K-means, PSO)	69

5.5.2	Évaluation des performances du modele	72
5.6	Conclusion	75
Chapitre 6 : Conclusion générale et perspectives		76
6.1	Introduction	76
6.2	Perspectives et travaux futurs	78
References		80
Résumé		85
Abstract		86

Liste des figures

1.1	Ville intelligente	2
2.1	Aperçu d'un système de recommandation	7
2.2	Filtrage basé sur le contenu	8
2.3	Filtrage collaboratif	9
2.4	Filtrage hybride	10
2.5	Caractéristiques d'une ville intelligente	15
2.6	Dimensions des villes intelligentes	15
4.1	Approche proposée	38
4.2	Calcul de prédiction	42
4.3	Similitude de jaccard pour chaque médecin	45
4.4	Initialisation optimale des K-means à l'aide de Particle Swarm Optimization	46
4.5	Système de recommandation à base de connaissance	49
4.6	Graphe distance	50
4.7	Diagramme d'activité	52
5.1	Dataset medecins	57
5.2	Dataset specialite	58

5.3	React	59
5.4	Axios	59
5.5	CSS	60
5.6	TinyDB	60
5.7	VS code	60
5.8	Flux de contrôle Django	61
5.9	Avantages de Django	62
5.10	Schéma de navigateur	63
5.11	Diagramme de sequence	64
5.12	Accueil	65
5.13	Trouver un médecin	66
5.14	Liste des médecins similaire	67
5.15	Noter un médecin	68
5.16	A propos	68
5.17	Regroupement des médecins selon leur spécialités	69
5.18	Regroupement des médecins selon leur spécialités	70
5.19	Similarité de Jaccard pour chaque médecin en utilisant l'algorithme K-means	70
5.20	Similarité de Jaccard pour chaque médecin en utilisant l'algorithme K-means et PSO	71
5.21	Résultat avec k-means	74
5.22	Résultat avec k-means et PSO	74

Liste des tableaux

2.1	Avantages et inconvénients des systèmes de recommandation . . .	11
2.2	Domaines clés des villes intelligentes, descriptions et exemples d'applications	17
3.1	État des lieux de certaines contributions relatives aux systèmes d'orientation des soins de santé (filtrage collaboratif).	27
3.2	État des lieux de certaines contributions relatives aux systèmes d'orientation des soins de santé (filtrage hybride).	28
3.3	État des lieux de certaines contributions relatives aux systèmes d'orientation des soins de santé (filtrage de contenu).	29

List of Algorithms

1	Conception du SR	39
2	Comparaison des profils	43
3	Recommandation	48
4	A*MD	51

Liste des abréviations

Abréviation	Description
BDD	Base de données
CSS	Cascading Style Sheets
CSV	Comma Separated Values
FBC	Filtrage basé sur le contenu
FC	Filtrage collaboratif
HTML	Hyper Text Markup Language
IDO	Internet des objets
KNN	K-Nearest Neighbors
ML	Machine learning
MVC	Model-View-Controller
MVT	Model-View-Template
NUMPY	NUMerical PYthon
PSO	Particle Swarm Optimization
SR	Système de recommandation
TF-IDF	Term Frequency-Inverse Document Frequency
TIC	Technologies de l'information et de la Communication
TSV	Tab Separated Values

Chapitre 1

Introduction générale

1.1 Introduction

La santé est un domaine essentiel qui influe grandement sur la qualité de vie de la population. Cependant, l'augmentation du nombre de personnes atteintes de maladies chroniques et de troubles de santé rend de plus en plus difficile l'accès rapide et efficace aux soins médicaux nécessaires. Dans ce contexte, les systèmes de recommandation dans le domaine de la santé gagnent en importance. Les villes intelligentes, où les technologies de l'information et de la communication sont largement déployées pour améliorer la qualité de vie des citoyens, sont des environnements favorables à l'utilisation de ces systèmes sophistiqués. En analysant les données de santé des patients à l'aide d'algorithmes, ces systèmes sont capables de recommander des professionnels de santé similaires en fonction de leur maladie ou de leur état de santé. L'objectif des systèmes de recommandation dans le domaine de la santé est de permettre aux patients de bénéficier d'un accès rapide et efficace aux soins médicaux dont ils ont besoin, tout en contribuant à réduire les coûts et à améliorer l'efficacité globale du système de santé. Avec l'avancée rapide des technologies, ces systèmes sont de plus en plus appelés à jouer un rôle majeur dans la prise en charge de la santé des individus dans les villes intelligentes.

1.2 Motivation

Trouver rapidement un médecin spécialisé peut devenir un véritable défi lorsque l'on tombe malade en visitant une ville. Les visiteurs ont besoin de trouver dans les meilleurs délais un médecin proche de leur emplacement, qui puisse traiter leur maladie tout en prenant en compte les compétences des médecins. Cette tâche peut être encore plus difficile dans un pays étranger, où les visiteurs peuvent être confrontés à des barrières de langue et à une incompré-

hension du système de santé local. Ils peuvent être confrontés à des difficultés pour comprendre comment accéder aux services médicaux, quels médecins sont spécialisés dans leur condition médicale, quelles sont les cliniques ou hôpitaux réputés, et comment obtenir des informations fiables sur les médecins et leurs compétences. De plus, en raison de l'urgence de la situation, il est important de trouver un médecin qui puisse fournir un traitement adéquat dans les meilleurs délais. Cependant, la recherche manuelle de médecins spécialisés, en particulier dans un pays étranger, peut prendre beaucoup de temps et d'efforts, ce qui peut entraîner des retards dans l'obtention des soins médicaux nécessaires.. Pour avoir un aperçu plus précis de cette situation, **la Figure 1.1** donne un aperçu sur les défis rencontrés lors de la recherche d'un médecin spécialisé en voyage.



Figure 1.1 – Ville intelligente

L'objectif principal de ce mémoire est de présenter une nouvelle approche de recommandation dédiée au domaine de la santé dans les villes intelligentes. Cette approche est basée sur l'utilisation d'un outil logiciel spécifiquement développé pour faciliter la recherche rapide des meilleurs professionnels de santé disponibles dans la région, tout en réduisant leur stress et leur anxiété lorsqu'ils sont malades dans une ville étrangère. Par conséquent, cette application vise à garantir aux visiteurs un séjour agréable tout en leur offrant un accès facile aux meilleurs soins de santé disponibles.

1.3 Contributions

Les contributions principales de ce mémoire sont multiples :

- Établissement d'un état de l'art des travaux majeurs liés aux systèmes de recommandation en relation avec notre projet d'étude.
- Création de deux datasets spécifiques permettant de réaliser notre étude.
- Proposition d'un nouveau modèle de recommandation basé sur le profil de l'utilisateur.
- Exploitation de l'historique des informations de l'utilisateur afin d'accélérer le processus de recommandation et d'améliorer sa pertinence.
- Présentation d'une nouvelle approche permettant de sélectionner les meilleures recommandations parmi une liste de propositions.

1.4 Méthodologie de travail

Notre travail a été guidé par les nombreuses problématiques rencontrées par les touristes malades dans une ville intelligente. Face à la complexité de trouver des médecins compétents et adaptés à leur situation, nous avons consacré notre temps à proposer un cadre méthodologique complet et efficace pour orienter au mieux les patients dans leur processus. Nous avons ainsi suivi les étapes suivantes :

- Recherche et analyse : Nous avons effectué une recherche approfondie de l'état de l'art des différents systèmes de recommandation (SRs) proposés par des chercheurs dans le domaine de la santé, en comparant les avantages et inconvénients de chaque système.
- Identification du problème et de la solution : Nous avons défini la problématique spécifique des touristes malades dans une ville intelligente et proposé une solution efficace en développant un système de recommandation adapté.
- Implémentation et expérimentation du système proposé : Nous avons mis en évidence le fonctionnement du SR que nous avons développé, ainsi que son intérêt pour les patients. Nous avons réalisé des expérimentations pour valider l'efficacité de notre solution et son utilité pour les utilisateurs.

Notre démarche méthodologique complète nous a permis de proposer une solution concrète et efficace pour aider les patients à trouver des médecins compétents et adaptés à leur situation dans une ville intelligente.

1.5 Organisation du mémoire

Le reste du mémoire est organisé comme suit :

- Chapitre 2 : Ce chapitre présente un aperçu général en deux parties sur les SRs et les villes intelligentes. La première partie présente la définition des SRs et de leurs différents types, les avantages et inconvénients de chaque type, ainsi que les défis liés aux SRs. La deuxième partie définit les villes intelligentes, leurs caractéristiques, dimensions et défis, et explore la relation entre les systèmes de recommandation et les villes intelligentes. Enfin, les différents domaines d'application des SRs sont présentés.

- Chapitre 3 : Ce chapitre Aborde l'état de l'art et les travaux déjà réalisés sur les SR en général, et plus précisément sur le domaine de la santé, en citant leurs principaux avantages et inconvénients.

- Chapitre 4 : Ce chapitre présente en détail notre approche, notamment l'architecture du système et les différentes étapes suivies pour construire notre système de recommandation destiné aux touristes malades en ville intelligente.

- Chapitre 5 : Ce chapitre détaille les aspects techniques liées à notre application, ainsi que les environnements logiciels et matériels utilisés lors de ce projet. Une évaluation du système est également présentée dans ce chapitre, ainsi que les différentes interfaces de l'application.

- Chapitre 6 : Enfin, ce mémoire est clôturé par ce chapitre qui donne les conclusions et perspectives de ce travail. Ce chapitre propose une synthèse et un bilan du travail effectué tout au long du mémoire ainsi qu'un ensemble de perspectives liées notamment à la poursuite de ce travail ainsi qu'aux nouveaux thèmes de recherche qui nous paraissent les plus pertinents.

Chapitre 2

Notions fondamentales

2.1 Introduction

Avec la disponibilité croissante des informations en ligne, les besoins des utilisateurs en matière de communication, d'échange d'idées et de partage d'informations sont devenus plus complexes. C'est pourquoi les SRs et les villes intelligentes sont deux concepts interconnectés qui ont le potentiel de transformer la façon dont nous vivons et interagissons avec notre environnement. Les SRs, sont des algorithmes qui proposent des suggestions personnalisées de produits, de services ou d'informations en fonction des préférences et des habitudes des utilisateurs. D'autre part, les villes intelligentes sont des villes qui utilisent des technologies de pointe pour améliorer la qualité de vie de leurs citoyens, optimiser les services publics, réduire les coûts et résoudre des problèmes complexes. Les SRs ont un rôle clé à jouer dans la réalisation de ces objectifs en proposant des recommandations personnalisées pour les services publics, les restaurants, les magasins et les services de transport, entre autres.

Dans ce chapitre, nous explorerons les généralités des SRs et des villes intelligentes, leurs avantages, leurs défis et les possibilités d'amélioration de la vie urbaine. Nous aborderons également la relation entre les SRs et les villes intelligentes.

2.2 Systèmes de recommandation

Les systèmes de recommandation ont été définis de plusieurs manières :

Les SRs peuvent être définis comme des programmes qui tentent de recommander les éléments (produits ou services) les plus appropriés à des utilisateurs spécifiques (individus ou entreprises) en mesurant l'intérêt de l'utilisateur pour

un élément sur la base d'informations connexes concernant les éléments, les utilisateurs [1] et les interactions entre eux.

En extrayant les données et les services les plus pertinents d'un ensemble de données, les SRs sont développés dans le but de réduire la surcharge d'informations et de fournir des services individualisés. La capacité d'un SR à créer des suggestions personnalisées en déduisant les préférences et les intérêts de l'utilisateur à partir des données de la personne et les intérêts de l'utilisateur à partir de l'activité de cette personne et/ou de celle d'autres utilisateurs est sa composante la plus cruciale [2].

La définition la plus courante et la plus large : « Un système qui peut fournir des recommandations personnalisées ou aider l'utilisateur à trouver des ressources intéressantes ou utiles dans un grand espace de données [3] ».

Un système de recommandation se compose de deux éléments clés : **les utilisateurs** et **les articles**.

- L'utilisateur est celui qui interagit avec le système, à qui le système recommande des éléments et qui, à son tour, donne un retour d'information sur ces éléments.

- L'article est le terme général utilisé pour décrire la ressource proposée par le système aux utilisateurs.

Le système fait des recommandations aux utilisateurs, qui fournissent ensuite un retour d'information sous la forme d'évaluations ou de notes. Ces évaluations sont enregistrées sous la forme d'un triplé (utilisateur, élément, note) et sont utilisées pour créer la "matrice des scores", qui représente les interactions entre les utilisateurs et les éléments. Les évaluations peuvent prendre différentes formes, telles que des notes numériques sur une échelle de 1 à 5 ou des options binaires "j'aime / je n'aime pas".

La figure 2.1 fournit un aperçu visuel de ce processus, illustrant comment les évaluations des utilisateurs sont utilisées pour alimenter la matrice des scores et influencer les recommandations ultérieures. Cela permet au système de fournir des suggestions personnalisées qui correspondent aux besoins spécifiques de chaque utilisateur.

2.3 Types de systèmes de recommandation

Le SR est l'un des outils les plus puissants et les plus pratiques du monde numérique actuel [4]. Il justifie souvent les suggestions qu'il fait afin d'aider les

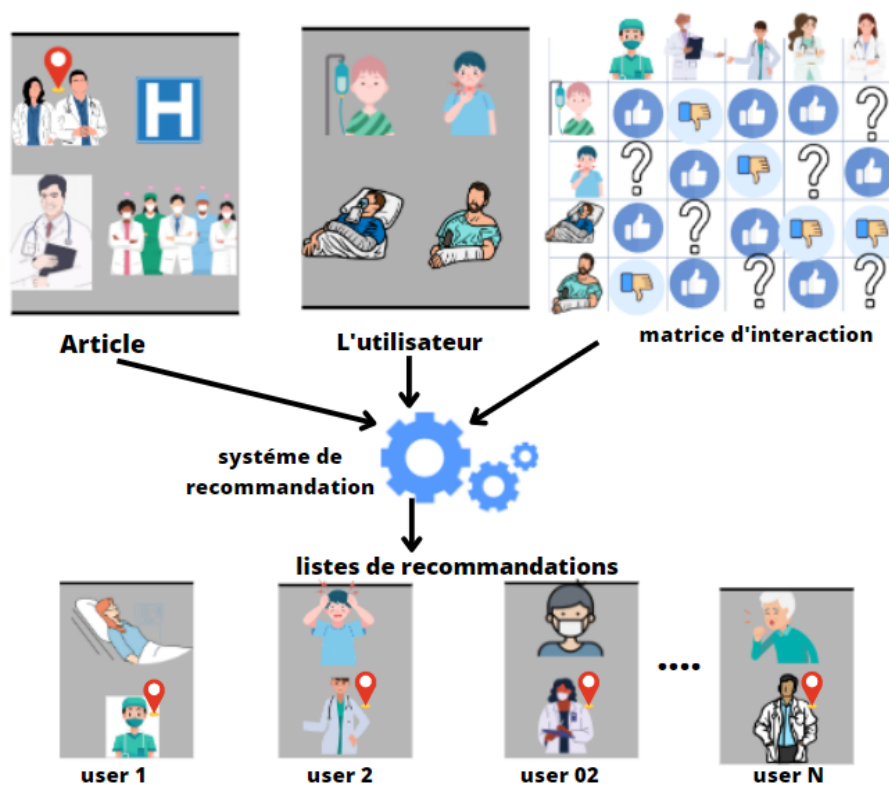


Figure 2.1 – Aperçu d’un système de recommandation

utilisateurs en ligne à localiser leurs biens, leurs amis, leurs activités... Les SRs peuvent être classés en trois types de méthodes : le filtrage basé sur le contenu (FBC), le filtrage collaboratif(FC) et le filtrage hybride [5]. Ces stratégies de recommandation doivent être étudiées afin de fournir au consommateur final les meilleures recommandations en fonction de ses intérêts.

2.3.1 Filtrage basé sur le contenu

Plusieurs SR ont été développés entre 1994 et 1995, dont le système "GroupLens" pour la recommandation d'articles d'actualité et de films [6], ainsi que le système "Ringo" [7]. Ces systèmes utilisent des descriptions similaires pour les éléments recommandés à l'utilisateur à un moment donné et les éléments qu'il a aimé dans le passé. Ils prennent en compte plusieurs facteurs communs, tels que le genre, les acteurs, le réalisateur, la langue, etc. [5].

Différents types de modèles sont utilisés dans le (FBC) pour identifier les similitudes entre les articles et produire des recommandations exploitables. Pour illustrer la relation entre les différents composants d'un corpus, des modèles probabilistes tels que le classificateur Naive Bayes, les arbres de décision ou les réseaux neuronaux [8] peuvent être utilisés. Ces modèles sont entraînés à l'aide d'approches d'intelligence artificielle et d'indexation. **La Figure 2.2** montre

une illustration des concepts discutés précédemment.

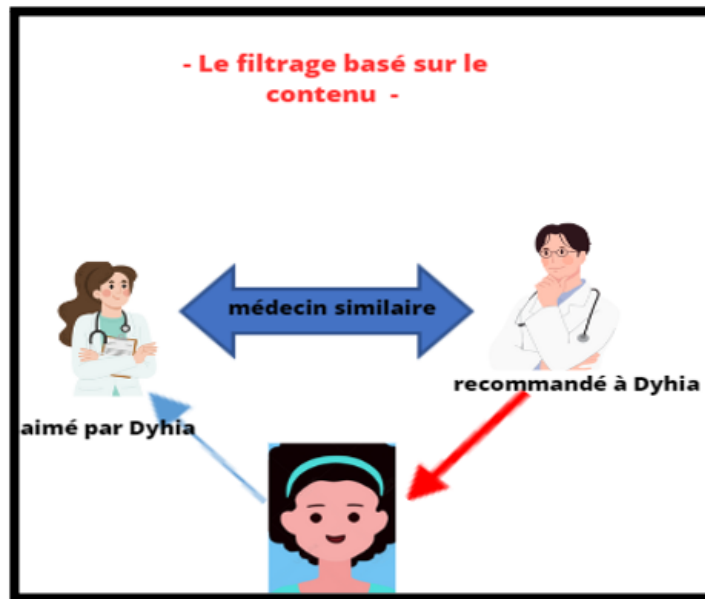


Figure 2.2 – Filtrage basé sur le contenu

2.3.2 Filtrage collaboratif

Introduit dans la littérature académique par le système Tapestry, il a été créé en 1992 [9] par le centre de recherche "Xerox" aux Etats-Unis. Ce type fonctionne en utilisant une base de données (user-item matrix) pour exploiter les évaluations faites par les utilisateurs sur certains documents, afin de recommander ces mêmes documents à d'autres utilisateurs, sans avoir analysé le contenu de ces documents.

Il existe généralement deux grandes sous-méthodes de FC : Les notes des utilisateurs stockées dans des bases de données sont utilisées par des méthodes basées sur la mémoire [5] pour faire des prédictions et des inférences. En créant une représentation simplifiée de la matrice des scores, les méthodes basées sur un modèle [10] rationalisent le processus de calcul et gèrent les résultats erronés. Avant de pouvoir être utilisé pour fournir des recommandations, le modèle doit passer par une phase d'apprentissage. Parmi les algorithmes les plus populaires de cette méthodologie, on trouve les approches basées sur le regroupement et les approches probabilistes. **La figure 2.3** présente une illustration des concepts discutés précédemment.

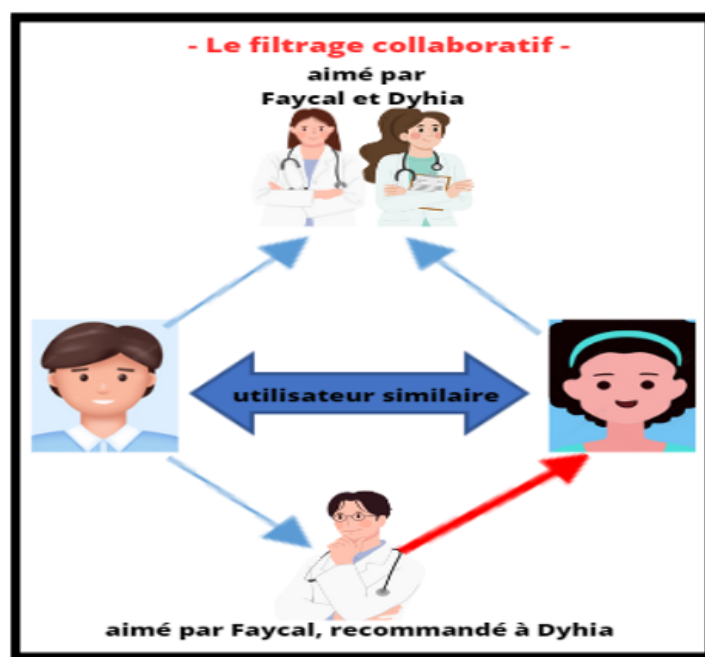


Figure 2.3 – Filtrage collaboratif

2.3.3 Approche hybride

Un SR est dit hybride quand il combine deux ou plusieurs approches de recommandation différentes. En raison de leur capacité à améliorer les performances, les systèmes hybrides sont de plus en plus courants. Ils utilisent un certain nombre de techniques de recommandation pour produire des prévisions et des suggestions.

Le FC, par exemple, peut être utilisé en plus de méthodes telles que Naive Bayes et k-plus-Voisin [11]. En outre, il est possible de combiner les deux techniques de filtrage collaboratif basées sur la mémoire et le modèle. Sept nouvelles approches, dont la réflexion, la commutation de méthodes mixtes, la combinaison de caractéristiques, l'amélioration des caractéristiques et le méta-niveau - ont été créées à la suite de la combinaison de plusieurs méthodologies. L'utilisation de ces approches hybrides [12]. vise à résoudre les problèmes posés par les SRs, tels que le manque d'informations au démarrage et l'incapacité des systèmes à s'adapter à l'évolution des profils des utilisateurs. **La Figure 2.4** est consacrée à l'illustration des concepts discutés précédemment

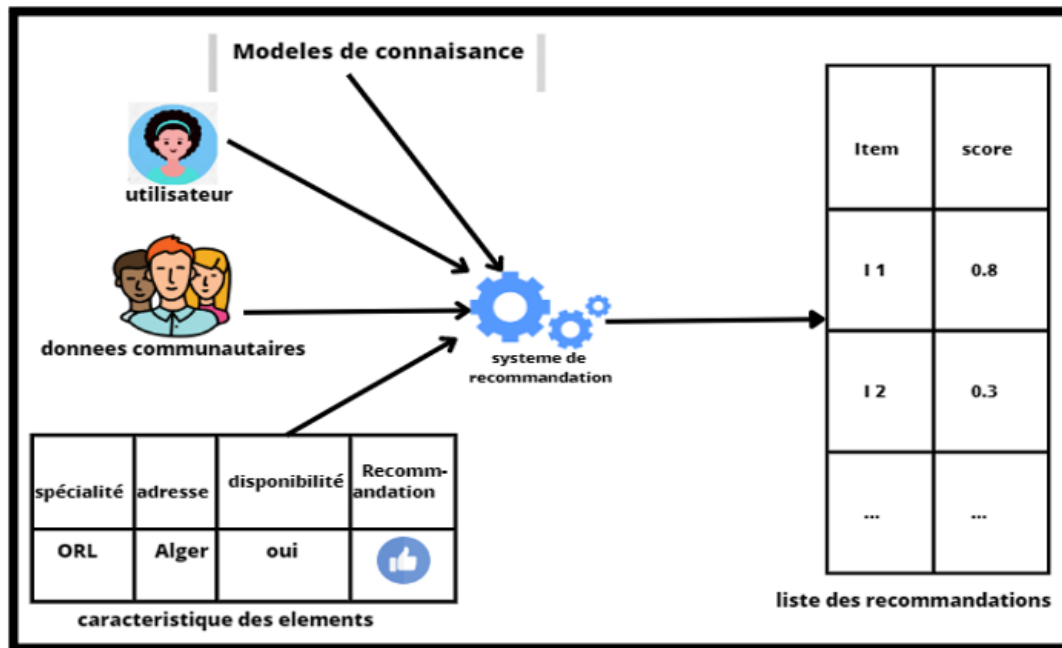


Figure 2.4 – Filtrage hybride

2.3.4 Autres types de recommandations

- **Recommandations basées sur les connaissances**

Les recommandations basées sur la connaissance utilisent des informations spécifiques pour identifier les articles qui correspondent aux préférences de l'utilisateur. Lorsque les données sont limitées, les systèmes basés sur la connaissance sont généralement fiables que les autres types de recommandations qui dépendent de l'historique de l'utilisateur. Toutefois, si le système de connaissance n'est pas en mesure d'apprendre des notes ou des actions de l'utilisateur, il risque de ne pas fournir de recommandations personnalisées. Il est important de noter que les systèmes de recommandation basés sur la connaissance sont souvent utilisés en conjonction avec d'autres méthodes pour garantir une expérience optimale à l'utilisateur [13].

- **Recommandation démographique**

La recommandation démographique est une méthode simple qui suggère des articles en fonction du profil démographique de l'utilisateur. Cette approche consiste à répartir les utilisateurs en plusieurs classes en fonction de leurs caractéristiques démographiques telles que le sexe, l'âge, la localisation, la langue, le pays, etc. L'idée sous-jacente à cette méthode est que les utilisateurs qui ont évolué dans des environnements similaires auront plus de chances de trouver les bons articles [14].

Voici un tableau (**Table 2.1**) comparant les avantages et les inconvénients de certaines techniques populaires de SR [15].

Technique	Avantage	Inconvénient
Filtrage collaboratif	<ol style="list-style-type: none"> 1. Peut faire des recommandations personnalisées basées sur les préférences des autres utilisateurs. 2. Peut traiter de grandes quantités de données. 3. Peut fonctionner correctement même lorsque les informations disponibles sur les utilisateurs ou les articles sont limitées. 	<ol style="list-style-type: none"> 1. Peut ne pas fonctionner correctement en cas de problèmes de démarrage à froid (lorsqu'il y a peu d'informations disponibles sur un nouvel utilisateur ou un nouvel objet). 2. Peut produire des recommandations inexactes si les données sont biaisées ou s'il existe de grandes différences dans les préférences des utilisateurs. 3. Peut-être vulnérable à la manipulation si les utilisateurs peuvent artificiellement gonfler leurs préférences.
Filtrage basé sur le contenu	<ol style="list-style-type: none"> 1. Peut faire des recommandations basées sur les attributs ou les caractéristiques des articles. 2. Peut gérer les problèmes de démarrage à froid en utilisant les informations relatives au contenu pour formuler des recommandations. 3. Peut faire des recommandations plus précises lorsqu'il y a beaucoup d'informations disponibles sur les articles. 	<ol style="list-style-type: none"> 1. Peut ne pas tenir compte des préférences ou des opinions des autres utilisateurs. 2. Peut produire des recommandations qui se limitent à un ensemble particulier d'articles. 3. Peut-être vulnérable à la partialité si les informations utilisées pour décrire les articles sont biaisées.
Systèmes hybrides de recommandation	<ol style="list-style-type: none"> 1. Peut combiner les points forts de plusieurs techniques de SR. 2. Peut traiter différents types de données et différents types de recommandations. 3. Peut produire des recommandations plus précises et plus diversifiées. 	<ol style="list-style-type: none"> 1. Peut-être plus complexe et plus difficile à concevoir et à mettre en œuvre. 2. Peut produire des recommandations contradictoires si les différentes techniques ne sont pas d'accord. 3. Peut-être plus coûteux en termes de calcul.

Table 2.1 – Avantages et inconvénients des systèmes de recommandation

2.4 Défis des systèmes de recommandations

De nombreuses plateformes en ligne, y compris les sites de commerce électronique et les plateformes de partage de médias sociaux, dépendent fortement des SR à l'heure actuelle. Ces systèmes utilisent des algorithmes pour recommander aux utilisateurs des articles ou des produits en fonction de leur comportement antérieur et de leurs préférences. Les systèmes de recommandation continuent à faire face à un certain nombre de défis qui peuvent entraver leur efficacité malgré leur utilisation répandue. Nous nous pencherons sur ces défis dans la suite de ce chapitre [16].

- Démarrage à froid : il s'agit d'une situation dans laquelle le SR ne dispose pas de suffisamment d'informations sur un nouvel utilisateur ou un nouvel élément pour formuler des recommandations précises [15]. Ces problèmes peuvent également être résolus par le filtrage hybride.

- Évolutivité : capacité du système à gérer des quantités croissantes de données et d'utilisateurs tout en maintenant ses performances [17].

- Confidentialité : protection des informations sensibles de l'utilisateur, telles que ses préférences et ses comportements, qui sont utilisées par le SR pour générer des recommandations [18].

- Sparsité : dans les boutiques en ligne, nous constatons que le nombre d'articles candidats à la recommandation est souvent élevé et que les utilisateurs n'évaluent qu'un petit sous-ensemble d'entre eux. Il est donc difficile de déterminer les intérêts d'un utilisateur et cela peut être associé à un mauvais voisinage. Par conséquent, la matrice d'évaluation (interaction entre l'utilisateur et l'article) est une matrice creuse avec un taux élevé de valeurs manquantes [19].

- Sémantique : Si le SR n'est pas en mesure de gérer une telle situation, il peut générer une liste de suggestions contenant des éléments similaires, ce qui réduirait la qualité des recommandations. Par exemple, les expressions "book" (réservation) et "book" (manuscrit), qui ont des significations différentes selon le contexte, ne seront pas distinguées par un système de recommandation.

Ces contraintes constituent des obstacles importants à l'élaboration et à la mise en œuvre de SR efficaces, nécessitant un examen minutieux et une résolution méthodique tout au long de ces processus.

2.5 Villes intelligentes

Dans cette section, nous examinons les nombreuses idées et les principes de base associés aux villes intelligentes, ainsi que leurs caractéristiques, leurs besoins et leur architecture. Le terme "ville intelligente" est apparu au début des années 1990 [20], soulignant l'importance des technologies de l'information et de la communication dans la construction de l'infrastructure d'une ville moderne.

La signification du terme (smart) varie en fonction de son utilisation. Différents articles ont utilisé des termes différents, tels que ville intelligente, ville de la connaissance, ville numérique, pour décrire ce concept. Il existe une différence significative entre une ville intelligente et une ville numérique, qui sont souvent confondues ou utilisées comme synonymes, mais qui ne sont pas exactement les mêmes [21].

La ville numérique est basée sur les technologies de l'information et de la communication et prend de la valeur lorsque la technologie numérique est utilisée pour améliorer la ville intelligente et la population [22].

La ville intelligente utilise des outils numériques pour améliorer la qualité de vie des gens et utilise la technologie pour le développement urbain dans des domaines tels que la mobilité, l'environnement et la participation des citoyens. La ville intelligente découle souvent de la ville numérique pour améliorer la gestion urbaine [23].

Voyons maintenant les définitions pour mieux comprendre le concept de ville intelligente. La ville intelligente est une ville interconnectée, instrumentée et intelligente qui comprend des dispositifs et des outils de capture de données tels que des compteurs intelligents, des smartphones, l'internet, des réseaux sociaux et de nombreux autres systèmes de collecte de données.

Les villes intelligentes sont des villes avancées, hautement technologiques, qui connectent leurs habitants, leurs informations et leurs éléments urbains à l'aide de technologies avancées afin de créer une ville durable, plus respectueuse de l'environnement, propice à l'innovation et à la compétitivité des entreprises, et qui offre à ses habitants un niveau de vie plus élevé [24].

2.5.1 Technologies de l'information et de la communication (TIC) et les technologies liées aux villes intelligentes

Une stratégie pluridisciplinaire est nécessaire pour développer une ville intelligente, impliquant des parties prenantes issues de différents secteurs, notamment la politique, la finance, la gestion municipale et les technologies de l'information et de la communication. Dans le domaine des TIC, les principaux acteurs impliqués dans le développement des villes intelligentes sont les suivants :

Les capteurs : Les éléments clés de la gestion et de l'optimisation des ressources dans les villes intelligentes sont les capteurs, qui permettent la collecte et la transmission de données en temps réel sur une variété de sujets liés à la ville, y compris la détection des compteurs de gaz et d'électricité [25].

Smartphones : les smartphones sont des appareils électroniques portables qui permettent la collecte et la transmission de données en temps réel. Les smartphones comprennent des capteurs intégrés tels que le GPS, des accéléromètres et des boussoles qui peuvent être utilisés pour collecter des données de position et assigner des tâches à certaines personnes. Les smartphones peuvent également être utilisés pour accéder aux applications des villes intelligentes, telles que les systèmes de transport, les services médicaux, les programmes de suivi de la qualité de l'air, etc. [26]

2.5.2 Caractéristiques

Les villes intelligentes sont souvent définies par quatre thèmes principaux : la durabilité, la commodité, la qualité de vie et l'urbanisation intelligente. Chacun de ces thèmes est essentiel au développement d'une ville intelligente, car ils optimisent l'infrastructure, les services et les ressources de la ville pour répondre aux besoins des citoyens de manière efficace et durable [27].

La figure 2.5 illustre ces concepts.

Ces aspects peuvent être étroitement liés et interdépendants. Par exemple, dans le domaine des soins de santé, l'utilisation de technologies intelligentes pour surveiller la santé des patients peut améliorer leur qualité de vie tout en réduisant les coûts des soins de santé pour les patients et les prestataires de soins. De même, la promotion de pratiques durables dans les entreprises de soins de santé peut contribuer à la durabilité environnementale tout en améliorant la santé et le bien-être des patients et des communautés

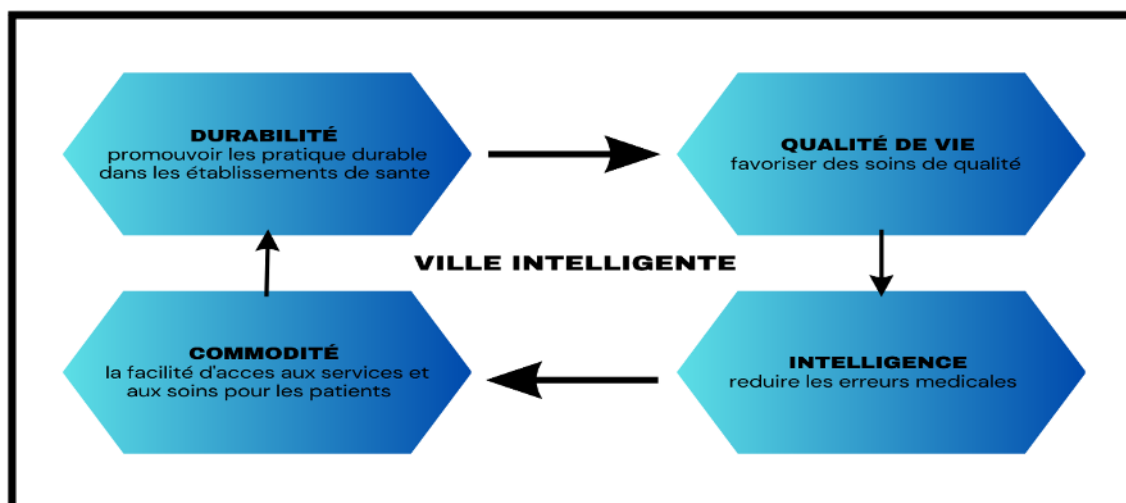


Figure 2.5 – Caractéristiques d'une ville intelligente

2.5.3 Dimensions

Les villes intelligentes sont basées sur l'intégration de différentes dimensions telles que l'économie intelligente, la mobilité intelligente, l'environnement intelligent, les personnes intelligentes, la vie intelligente et la gouvernance intelligente pour décrire les différentes dimensions des villes intelligentes [2].

La figure 2.6 illustre ces différentes dimensions.

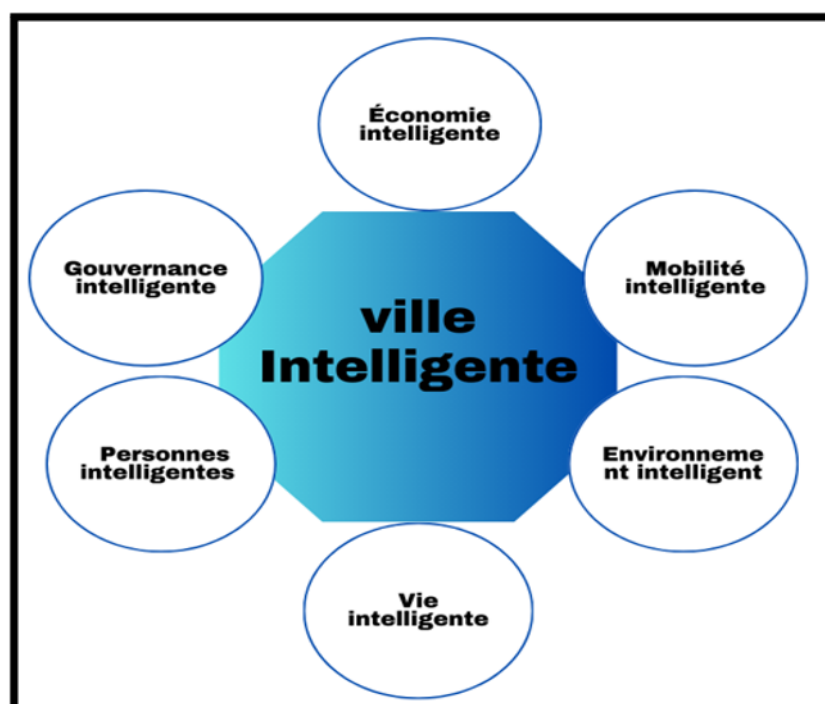


Figure 2.6 – Dimensions des villes intelligentes

- Économie intelligente : les villes intelligentes peuvent encourager l'innovation dans le domaine de la santé en soutenant les start-ups et les entreprises qui développent des technologies de pointe pour améliorer la santé des citoyens [28].
- Mobilité intelligente : Les villes intelligentes peuvent améliorer l'accessibilité aux soins de santé en utilisant des technologies de transport intelligentes pour faciliter le déplacement des patients et du personnel médical [29].
- Environnement intelligent : Les villes intelligentes peuvent réduire l'exposition des citoyens aux risques environnementaux en utilisant des technologies de surveillance pour détecter les polluants et les maladies liées à l'environnement [30].
- Personnes intelligentes : les villes intelligentes peuvent encourager des modes de vie plus sains en offrant des espaces publics agréables pour faire de l'exercice, en encourageant l'utilisation de modes de transport actifs et en fournissant des informations sur les choix alimentaires sains [31].
- Vie intelligente : les villes intelligentes peuvent fournir des soins de santé à domicile en utilisant des technologies intelligentes pour surveiller l'état de santé des personnes âgées [32].
- Gouvernance intelligente : les villes intelligentes peuvent améliorer la gestion des services de santé en utilisant la technologie pour surveiller les tendances en matière de santé et améliorer la planification et la fourniture des services [33].

En intégrant ces dimensions de la santé, les villes intelligentes peuvent améliorer la qualité de vie des citoyens et fournir un accès plus facile et plus efficace aux services de santé.

2.5.4 Systèmes de recommandation et villes intelligentes

La relation entre les systèmes de recommandation et les villes intelligentes est cruciale, car les villes intelligentes offrent une variété de services intelligents parmi lesquels les utilisateurs peuvent avoir du mal à choisir. Les SRs peuvent aider les utilisateurs en leur fournissant des recommandations personnalisées de services et de produits qui répondent à leurs besoins et à leurs préférences, facilitant ainsi leur vie quotidienne. Les six dimensions des villes intelligentes, telles qu'elles sont définies par [34], comprennent l'économie intelligente, l'environnement, la mobilité, la gouvernance, la vie et les personnes, et les systèmes de recommandation examinés sont classés en fonction des actions et des objec-

tifs de la ville intelligente qu'ils ciblent pour chaque dimension.

Le tableau 2.2 résume les différents domaines d'application d'une ville intelligente dans lesquels les SRs peuvent jouer un rôle :

Domaine	Description du domaine	Exemple d'applications
L'économie intelligente	Optimiser l'utilisation des ressources et des services dans la ville pour créer une économie plus durable et plus efficace	Recommandations personnalisées pour des bâtiments et des systèmes de transport économes en énergie, optimisation des ressources et des services publics
Environnement intelligent	Améliorer la qualité de vie des citoyens et créer un environnement plus durable	Recommandations pour des produits et services respectueux de l'environnement, systèmes de gestion des déchets
Mobilité intelligente	Améliorer la mobilité et l'accessibilité des citoyens dans la ville	Recommandations d'options de transport basées sur les conditions de circulation en temps réel et les préférences personnelles
Gouvernance intelligente	Améliorer l'efficacité et la transparence de l'administration de la ville et des processus de prise de décision	Recommandations pour des initiatives et des services à l'échelle de la ville, basées sur le retour d'information et les préférences des citoyens
Vie intelligente	Améliorer la qualité de vie et le bien-être des citoyens	Recommandations concernant les activités de loisirs, les services de santé et les autres nécessités de la vie quotidienne
Les personnes intelligentes	Donner aux citoyens les moyens de participer à l'élaboration et à la mise en œuvre d'initiatives en matière de villes intelligentes	Recommandations de programmes et d'initiatives d'engagement des citoyens en fonction des intérêts et des préférences de chacun

Table 2.2 – Domaines clés des villes intelligentes, descriptions et exemples d'applications

2.5.5 Conclusion

Dans ce deuxième chapitre, nous avons abordé les notions de SR et de villes intelligentes, en fournissant une définition claire de ces concepts ainsi qu'une présentation des différents types, défis et problèmes qu'ils impliquent. Nous avons également mis en évidence la relation entre les SRs et les villes intelligentes en présentant des exemples concrets.

Dans le chapitre suivant, nous approfondirons notre analyse en examinant l'état de l'art des travaux relatifs aux SRs, en nous concentrant sur les différentes approches utilisées. Nous discuterons des avancées récentes dans le domaine et évaluerons les avantages et les limites de ces approches. Cette analyse nous permettra de mieux comprendre les problèmes et les défis actuels des SRs dans le contexte des villes intelligentes.

Chapitre 3

L'état de l'art

3.1 Introduction

Le thème de la santé est en pleine expansion dans l'industrie des services de santé numériques. Les groupes internationaux affirment que l'utilisation de technologies adaptées aux difficultés complexes de l'environnement est cruciale pour améliorer la qualité des systèmes de santé, en particulier dans les pays en voie de développement. La construction d'un système d'information de santé intégré dans une approche systémique est rendue possible, entre autres, par les progrès des TIC. Les problèmes auxquels sont confrontés les services de santé sont de plus en plus considérés comme ayant une relation directe avec la santé. Les systèmes de soins de santé doivent analyser un grand nombre de données relatives aux patients afin de fournir des informations et des prévisions sur diverses maladies. Ces systèmes doivent être intelligents et capables de prédire l'état des patients en fonction de leur mode de vie, de leur état de santé physique et de leurs activités sociales. Comme les gens ont tendance à utiliser les TIC pour en savoir plus sur leur santé, il est important de mettre en œuvre des systèmes de santé intelligents afin de formuler des recommandations en matière de diagnostic, d'assurance maladie et de médecine alternative sur la base de l'état de santé des patients.

Dans ce chapitre, nous décrivons la situation actuelle des nombreux travaux réalisés dans le domaine de la santé. Nous effectuerons une comparaison de ces contributions afin de mettre en évidence les avantages et les inconvénients des différentes stratégies proposées.

3.2 Travaux connexes

Ce passage met en évidence le fait que de nombreux systèmes de villes intelligentes proposent des recommandations en matière de soins de santé à leurs utilisateurs, et quelques-uns de ces systèmes ont été sélectionnés pour une discussion et une évaluation plus approfondies. Les recommandations sont regroupées en catégories à des fins d'analyse et d'évaluation :

3.2.1 Filtrage collaboratif

Dans le domaine médical, le FC peut être utilisé pour recommander des traitements en examinant les antécédents médicaux, l'état de santé actuel et les réactions du patient à divers traitements. En leur fournissant des informations sur les choix de traitement qui se sont avérés efficaces pour des patients présentant des pathologies similaires, les systèmes de filtrage collaboratif peuvent aider les médecins à prendre des décisions éclairées sur les soins à prodiguer aux patients. Ils ne doivent pas être utilisés comme unique source d'information sur la santé, mais plutôt comme un outil supplémentaire dans le processus de prise de décision. Nous passerons en revue les projets qui ont été proposés dans ce domaine et découvrirons quelles stratégies sont souvent employées pour obtenir des résultats positifs :

Leanza et al. [35], ont mis au point de déployer des SRs en temps réel couplés à l'infrastructure de capteurs des villes intelligentes afin de fournir aux résidents des suggestions d'itinéraires qui tiennent compte de leur état de santé et de leurs préférences. Les composants, l'architecture et le fonctionnement du système proposé ont tous été décrits par les auteurs. En outre, des expériences ont été menées et une application pour smartphone appelée « SmartRoute » a été développée en utilisant des données réelles fournies par l'infrastructure IoT de la ville intelligente ainsi que des contributions participatives d'utilisateurs qui ont testé le système dans des circonstances réelles. Les résultats expérimentaux démontrent que la stratégie est réalisable et qu'elle peut offrir aux citoyens des recommandations judicieuses, même en présence d'une quantité importante de données inconnues. L'étude s'inscrit dans le cadre d'une discussion plus large sur les investissements respectueux de l'environnement dans les villes intelligentes.

Erdeniz S.P.[36], s'inscrit dans un contexte de recherche plus large sur l'uti-

lisation des SRs pour améliorer l'état de santé des individus. Dans ce domaine de recherche, de nombreux travaux ont examiné les avantages et les applications des technologies de recommandation dans les systèmes de santé mobiles basés sur l'Internet des objets (IdO) (m-health).

Dans ce contexte, l'article fait référence à une étude antérieure qui propose deux nouveaux recommandeurs pour un projet IoT afin de fournir de nouvelles applications de santé, des appareils et des plans d'activité physique pour les patients. Les auteurs ont utilisé des techniques de filtrage collaboratif et de filtrage basé sur le contenu pour recommander aux patients des applications et des appareils de santé mobiles personnalisés en fonction de leur profil de santé.

Erdeniz S.P s'appuie sur cette étude précédente en proposant une approche plus complète pour les SRs de santé mobile basés sur l'IdO. Les auteurs utilisent des techniques de filtrage collaboratif basées sur la factorisation matricielle et l'analyse de grappes, ainsi que des techniques de filtrage basées sur le contenu et le traitement automatique du langage naturel (NLP) pour analyser les informations sur la santé des utilisateurs. Les résultats de l'évaluation montrent que l'approche proposée est efficace pour recommander des applications de santé mobiles personnalisées.

L'article présente également des exemples d'utilisation de SRs dans l'application pilote "Quantified-Self" du projet AGILE, qui utilise l'IdO pour collecter des données sur la santé et le bien-être des utilisateurs. Cette application permet aux utilisateurs de suivre leur santé et leur activité physique, de recevoir des recommandations personnalisées et de participer à des programmes de prévention de la santé.

Saman et al. [37], ont étudié l'utilisation des SRs dans les appareils IoT (Internet des objets) et ont recueilli des données auprès d'entreprises telles que Telus, Libelium et BlueRover. Sur la base d'une enquête menée auprès de 1 875 utilisateurs, le SR propose des options aux utilisateurs en fonction des appareils IoT populaires (IOTPO), des services IoT populaires (IOTPS) et de la similarité des profils (IOTSRS). Les résultats montrent que la plus grande précision est obtenue en recommandant des services basés sur les profils des utilisateurs, ce qui indique l'importance des profils des utilisateurs dans la détermination de leurs intérêts et de leurs préférences pour les appareils IoT. La précision du SR augmente avec le nombre d'utilisateurs, ce qui en fait un outil efficace pour analyser les préférences des utilisateurs d'appareils IoT.

Sandeep Kumar et al. [38] ont proposés un SR pour fournir aux gens des

recommandations d'itinéraires d'exercice personnalisés en fonction de leur état de santé, en utilisant des données en temps réel pour améliorer leur mode de vie et améliorer leurs activités de santé quotidiennes. Le système utilise des algorithmes de FC basés sur l'apprentissage profond pour recommander des itinéraires d'exercice appropriés en fonction des données de santé de l'utilisateur. Le système a été évalué avec succès sur un ensemble de données de test, montrant une grande précision des recommandations et une amélioration significative de la santé et du bien-être de l'utilisateur.

3.2.2 Filtrage hybride

Les systèmes de filtrage hybride sont une combinaison de filtrage collaboratif et de filtrage basé sur le contenu qui utilise à la fois les données de l'utilisateur et du produit pour recommander des produits ou des services. Dans le domaine des soins de santé, ils peuvent recommander des traitements en fonction des antécédents médicaux, des préférences et des directives. Les systèmes hybrides offrent des recommandations plus précises et plus personnalisées que l'une ou l'autre de ces approches, mais ils peuvent être complexes à mettre en place et nécessiter beaucoup de données.

Jabeen et al. [39], ont présenté un SR hybride basé sur IdO pour diagnostiquer les symptômes des maladies cardiaques et fournir des recommandations physiques et diététiques personnalisées.

Le système a été conçu pour collecter les données des patients à l'aide de capteurs IoT, qui recueillent des informations importantes telles que la pression artérielle, la fréquence cardiaque et le niveau d'activité physique. Ensuite, un modèle de prédiction des maladies cardiaques a été utilisé pour diagnostiquer les maladies cardiovasculaires et les classer en fonction du sexe et de l'âge des patients.

Le SRs hybride proposé par les auteurs combine deux approches de recommandation : le filtrage collaboratif et le filtrage basé sur le contenu. La première approche recommande des articles similaires à ceux que le patient a aimés, en utilisant les données de patients présentant des caractéristiques similaires. La seconde approche recommande des articles en fonction de leur contenu et des informations relatives au patient, telles que l'âge, le sexe, les antécédents médicaux, etc.

Les résultats ont montré que le SR hybride avait une précision de 87,2% pour la recommandation de régimes et de 91,5 % pour la recommandation d'exer-

cices. Les auteurs ont également comparé leur SR à d'autres SRs traditionnels, tels que le filtrage collaboratif et le filtrage de contenu, et ont montré que leur approche hybride était plus efficace.

Senthilselvan [40] présente un SR personnalisé appelé ProTrip qui aide les voyageurs en générant des suggestions basées sur leurs intérêts, leurs préférences, leur séquence de voyage, leurs activités, leurs motivations, leurs opinions et leurs informations démographiques. ProTrip est un système axé sur la santé qui suggère des aliments en fonction des caractéristiques climatiques, des choix personnels et de la valeur nutritionnelle. Il peut également aider les voyageurs souffrant de maladies chroniques ou suivant des régimes stricts. ProTrip s'appuie sur des connaissances ontologiques et des mécanismes de filtrage personnalisés, et a amélioré la précision et l'efficacité par rapport aux modèles existants.

Le système ProTrip est capable de personnaliser les recommandations en fonction des préférences et des besoins individuels de chaque voyageur. Les données démographiques telles que l'âge, le sexe et la profession sont également prises en compte pour fournir des recommandations pertinentes. Le système est capable de prendre en compte l'état de santé du voyageur et de lui suggérer des aliments adaptés à son état, en fonction de son régime alimentaire et de ses restrictions.

Le système ProTrip utilise des ontologies pour représenter les connaissances et les relations entre les différents domaines d'expertise, tels que la nutrition, les activités de voyage et les conditions météorologiques. Ces ontologies sont utilisées pour guider le processus de recommandation en identifiant les préférences et les besoins des voyageurs.

Les résultats de l'évaluation de ProTrip ont montré que le système est capable de fournir des recommandations pertinentes et précises en termes d'alimentation et d'activités de voyage pour les voyageurs, avec une précision de recommandation de plus de 80 %. Le système a également fait preuve d'une efficacité supérieure à celle des modèles de recommandation existants, tels que le filtrage collaboratif et la factorisation matricielle.

L'article **Zhang et al.** [41] présente un état des lieux des SRs médicale et décrit un nouveau système appelé iDoctor qui est basé sur des méthodes hybrides de factorisation matricielle. iDoctor diffère des SRs existants en utilisant l'analyse des sentiments pour comprendre l'influence des émotions sur les opinions des utilisateurs, et en incorporant les préférences des utilisateurs et les caractéristiques des médecins dans le SR.

Les résultats expérimentaux montrent qu'iDoctor fournit des prédictions plus exactes et une plus grande précision dans les recommandations de soins de santé. Les auteurs suggèrent que iDoctor pourrait avoir des applications pratiques dans le secteur des soins de santé, notamment en fournissant des recommandations de soins de santé personnalisées aux patients. L'article souligne l'importance de l'analyse des sentiments et des préférences des utilisateurs dans la conception des SRs de soins de santé.

Farman [42] présente une approche basée sur l'IdO qui permet aux patients de recevoir des recommandations diététiques à l'aide de systèmes basés sur l'ontologie. L'article aborde également l'utilisation de capteurs pour collecter des informations sur la santé des patients et l'utilisation de réfrigérateurs intelligents et de boîtes à pharmacie pour fournir des recommandations en matière de régime alimentaire et de médicaments. L'objectif d'ape est de réduire le fardeau que représentent les patients chroniques pour les hôpitaux et de permettre aux patients de recevoir des soins à distance.

Han et al. [43], étudient la manière dont les SRs peuvent être utilisés pour faciliter l'adéquation entre les patients et les médecins dans le domaine des soins primaires.

L'article propose un SR hybride qui utilise à la fois des approches basées sur le contenu et des approches collaboratives. Les auteurs ont recueilli des données sur les patients et les médecins dans un contexte de soins primaires, telles que des informations démographiques, des antécédents médicaux, des évaluations de patients, des notes de médecins, etc. Ils ont ensuite utilisé ces données pour créer des profils de patients et de médecins.

Ils ont ensuite utilisé ces données pour créer des profils de patients et de médecins. Ensuite, les auteurs ont utilisé des techniques d'apprentissage automatique pour former un modèle de recommandation hybride qui peut prédire la probabilité qu'un patient et un médecin soient bien assortis. Le modèle tient compte à la fois des similitudes entre les profils des patients et des médecins et des avis et évaluations donnés par d'autres patients.

Les auteurs ont testé leur modèle sur un ensemble de données réelles de patients et de médecins et ont constaté qu'il était capable de fournir des recommandations précises et utiles. Ils ont également comparé leur approche hybride à des approches basées uniquement sur le contenu ou sur la collaboration, et ont constaté que leur approche hybride était plus performante.

Ben Abdesslem Karaa et al. [44] ont développé un système appelé RecSPSC

qui utilise les tweets pour fournir des recommandations pour les projets de démarrage dans les domaines de la ville intelligente et de la santé intelligente. Le système utilise l'apprentissage automatique et l'algorithme Word2Vec pour analyser les tweets, ainsi qu'une méthode de recommandation basée sur l'ontologie pour améliorer la précision des recommandations. Les résultats montrent que RecSPSC surpasse les approches de recommandation traditionnelles en termes de précision. L'objectif de ce travail est d'améliorer la qualité de vie en fournissant des recommandations pour les projets de démarrage dans les villes intelligentes, ce qui peut avoir un impact positif sur le développement économique et social d'un pays.

3.2.3 Filtrage de contenu

Le filtrage de contenu est un SR utilisé dans le domaine des soins de santé, qui émet des recommandations basées sur les informations contenues dans les données de santé. Il peut être utilisé à des fins de diagnostic, pour permettre aux patients de trouver des informations sur leur état de santé, ou aux entreprises de soins de santé d'améliorer les soins prodigués aux patients

Les auteurs **Vairale et al.** [45] expliquent que les patients atteints de la thyroïde ont des restrictions alimentaires spécifiques qui peuvent aider à gérer leur état de santé. Cependant, il peut être difficile pour les patients de trouver des aliments adaptés à leur régime alimentaire. Dans cet article, les auteurs proposent une méthode basée sur le contenu pour recommander des aliments aux patients atteints de la thyroïde.

La méthode de recommandation est basée sur l'algorithme KNN (K-Nearest Neighbors) qui est un algorithme de classification utilisé dans l'apprentissage automatique. Dans cette méthode, les aliments sont représentés sous forme de vecteurs de caractéristiques, qui sont utilisés pour calculer la similitude entre les différents aliments. Les aliments les plus similaires sont ensuite recommandés aux patients en fonction de leur régime alimentaire.

Les auteurs ont évalué la méthode de recommandation en utilisant un ensemble de données de 30 aliments différents et ont obtenu des résultats satisfaisants. La méthode a été en mesure de recommander des aliments adaptés au régime alimentaire des patients atteints de la thyroïde en utilisant une méthode basée sur le contenu.

3.3 Tableaux comparatifs

Nous avons présenté les principales approches des recommandations dans le domaine des soins de santé.

Dans les tableaux qui suivent, nous présentons l'état de l'art de certaines contributions relatives aux SR dans le domaine des soins de santé. Les tableaux se composent de 7 colonnes où chaque colonne indique des critères de comparaison comme suit :

- La colonne "**Catégorie**" indique l'approche en question.
- La colonne "**Référence**" indique la source de l'article.
- La colonne "**Algorithme de recommandation**" indique le type d'algorithme utilisé dans les SR pour suggérer des articles à un utilisateur en fonction de son historique et de ses préférences
- La colonne "**Modèle sémantique**" décrit formellement la connaissance de l'entreprise.
- La colonne "**ensemble de données/base de données**" regroupe un ensemble de données cohérentes qui peuvent être présentées sous différents formats.
- La colonne "**Résultats**" indique ce qui est produit par une activité consciente orientée vers une fin.
- La colonne "**Évaluation**" indique les métriques utilisées pour l'évaluation d'un modèle en fonction du type de problème à résoudre.

Catégories	Référence	Algorithme de recommandation	modèle sémantique	ensemble de données	Résultats	Évaluation
filtrage collaboratif	[35]	K plus proches voisins		Dossiers médicaux électroniques, préférences des utilisateurs	Pratique et bonne recommandation	Erreur absolue moyenne (MAE)
	[36]	K plus proches voisins		appareils comme (compteurs de pas, bracelets)	application m-santé	
	[37]	K plus proches voisins		Collecte de données auprès d'entreprises telles que Telus, Libelium et BlueRover.	Recommande des appareils IoT aux utilisateurs en fonction de leurs intérêts et de leurs besoins.	Précision
	[38]	k- plus proche voisin		Dossier médical électronique	fournit aux personnes des recommandations d'itinéraires d'entraînement personnalisés en fonction de leur état de santé	Erreur absolue moyenne (MAE)

Table 3.1 – État des lieux de certaines contributions relatives aux systèmes d'orientation des soins de santé (filtrage collaboratif).

Catégories	Référence	Algorithme de recommandation	modèle sémantique	ensemble de données	Résultats	Évaluation
filtrage hybride	[39]	K plus proches voisins		DHôpital de renom	Une bonne classification des maladies cardiovasculaires à l'aide de capteurs de détection de données	Précision Et Recall et Erreur absolue moyenne (MAE)
	[40]	K-Means	Ontologie	Alimentation Informations Utilisateur	Plans diététiques pour les voyageurs souffrant de maladies	Précision
	[41]	K plus proches voisins		Dossier médical électronique	iDoctor fournit des recommandations plus personnalisées, plus professionnelles et plus objectives	
	[42]	K plus proches voisins	Ontologie	Médicaments	réduire le fardeau que représentent les malades chroniques pour les hôpitaux recommander des médicaments pour les malades -proposer un régime alimentaire efficace pour les malades	Précision Et Recall
	[43]	K-Means		Sur la base de 4,5 millions de rendez-vous de patients avec plus de 100 000 spécialistes médicaux	Recommandation efficace	validation croisée
	[44]	K plus proches voisins et K-Means		Médias sociaux (twitter)	Precision and recall, F-measure	

Table 3.2 – État des lieux de certaines contributions relatives aux systèmes d'orientation des soins de santé (filtrage hybride).

Catégories	Référence	Algorithme de recommandation	modèle sémantique	ensemble de données	Résultats	Évaluation
filtrage de contenu	[45]	K plus proches voisins		les bases de données 30 aliments différents	Sites web indiens d'alimentation	

Table 3.3 – État des lieux de certaines contributions relatives aux systèmes d'orientation des soins de santé (filtrage de contenu).

3.4 Discussion et comparaison

Notre étude nous a permis d'évaluer l'efficacité de diverses approches et méthodes de création de SR en examinant les résultats en termes d'exactitude, de rappel et de précision.

Sur la base des études examinées ci-dessus, nous avons constaté que la majorité des projets de recommandation dans le domaine des services de santé préfèrent utiliser une approche de recommandation collaborative.

Toutefois, l'utilisation de la fonction de FC peut entraîner des problèmes de démarrage pour les utilisateurs qui viennent de s'inscrire. En effet, pour générer des recommandations personnalisées, ces systèmes s'appuient généralement sur les évaluations des utilisateurs. Par conséquent, parce qu'ils n'ont pas fourni suffisamment d'informations pour que le système puisse les comprendre, les nouveaux utilisateurs qui n'ont pas encore donné de commentaires ou d'évaluations ne peuvent pas recevoir de recommandations personnalisées.

Nous avons examiné différentes approches en partant d'articles déjà lus et cités. Nous avons également examiné les articles lus mais non cités. Les tableaux ci-dessus montre une comparaison partielle.

Cette étude donne un aperçu complet des différentes façons dont les SR sont utilisées dans le domaine de la santé. Nous avons parlé des différents sous-domaines des applications de SR, des différentes méthodes de recommandation employées, des différentes manières dont elles sont évaluées et, enfin, de la manière dont elles présentent les recommandations aux utilisateurs. En outre, la majorité des approches basées sur le profil de l'utilisateur qui ont été mentionnées ci-dessus n'ont pas abordé de manière adéquate la notion qui conduit à des conflits entre les préférences de l'utilisateur pour sa localisation et son état de santé et de maladie, par exemple.

Lorsque l'on parle de SR, il est habituel d'énumérer les différentes approches

employées.

3.4.1 Mesure de la similarité

est une technique utilisée pour quantifier le degré de similarité entre deux objets (utilisateur-utilisateur), ou (article-élément), ou (utilisateur-élément), fréquemment dans le cadre de l'analyse de données ou de l'apprentissage automatique : Cosinus Similarity, Jaccard's Coefficient, Euclidean distance... par exemple, dans cet article [42], il utilise la distance euclidienne pour calculer la similarité entre les préférences des utilisateurs et les articles recommandés dans un système basé sur l'IoT pour recommander des services de soins de santé.

3.4.2 Matégorisation des utilisateurs ou des éléments

Cette méthode consiste à regrouper les utilisateurs ou les éléments en différentes catégories à l'aide de méthodes de catégorisation supervisées ou non supervisées. Par exemple, dans l'article [37], les auteurs ont utilisé la classification des éléments pour regrouper les éléments en fonction de caractéristiques similaires, puis ils ont utilisé la classification des utilisateurs pour diviser les utilisateurs en groupes en fonction de leurs préférences et de leurs caractéristiques démographiques. Ensuite, en utilisant les préférences des utilisateurs et leur groupe, ils ont combiné ces méthodes pour recommander des articles pertinents pour eux.

3.4.3 Techniques de pondération des termes (TF-IDF)

Cette méthode consiste à localiser les termes les plus significatifs dans les résumés d'articles et à leur attribuer une pondération plus élevée. Par exemple, dans l'article [44], les méthodes de pondération TF-IDF ont été utilisées pour accroître la précision du SR dans le domaine de la santé. Ils ont discuté de l'utilisation de cette méthode pour évaluer l'importance des termes dans un document ou un groupe de documents et ont proposé d'attribuer un poids à chaque terme afin d'évaluer la similarité des documents et de suggérer des documents connexes.

3.4.4 Représentations de réseaux neuronaux

Il s'agit d'une méthode d'apprentissage automatique composée de couches de réseaux neuronaux artificiels qui peuvent reconnaître des structures de modèles complexes dans les données. Par exemple, l'auteur de l'article [39] a pro-

posé un SR basé sur un réseau neuronal pour prédire les maladies cardiovasculaires. Le système a été formé sur un ensemble de données de patients comprenant des caractéristiques telles que l'âge et le sexe. Le réseau neuronal a été créé pour apprendre à extraire les caractéristiques les plus significatives des données des patients et à prédire les maladies cardiovasculaires.

Sur la base des résultats de cette analyse, nous avons déterminé que les SR sont des systèmes informatisés qui utilisent des algorithmes pour recommander aux utilisateurs des mesures de santé visant à améliorer leur bien-être.

Nous allons maintenant analyser un élément particulier présenté dans le tableau en identifiant ses forces et ses avantages. Cependant, nous examinerons également leurs lacunes et leurs limites afin de formuler une critique utile qui pourrait contribuer à améliorer ces éléments.

Les auteurs présents [36] une analyse approfondie des avantages et des inconvénients des SRs pour les applications mobiles de santé basées sur l'IdO. (Internet des objets). Les avantages de ces systèmes sont nombreux et comprennent une personnalisation accrue des recommandations par l'utilisateur en fonction de ses préférences et de ses besoins en matière de soins de santé, une meilleure gestion des données basées sur l'IdO et un engagement accru de l'utilisateur en termes de santé grâce à la fourniture de recommandations personnalisées en temps réel. L'article souligne également certains inconvénients, tels que les dangers pour la confidentialité et la sécurité des données de santé collectées et analysées par les systèmes IoT, les dangers de recommandations biaisées, et le manque de normes et de réglementations pour garantir la fiabilité et la sécurité des applications mobiles de santé IoT. Il est donc essentiel de tenir compte de ces facteurs avant d'utiliser des SRs IoT pour les applications de santé mobiles afin de garantir une utilisation efficace et sûre de ces technologies dans le secteur de la santé. Les résultats de cette étude peuvent aider à mieux comprendre les avantages et les défis des SRs pour les applications mobiles de santé IoT et à développer des stratégies pour maximiser les avantages tout en minimisant les risques potentiels.

Les auteurs [38] décrivent un système de filtrage collaboratif et de recommandations de santé basées sur l'apprentissage profond. Les avantages de l'article comprennent l'utilisation de l'apprentissage en profondeur pour identifier les relations complexes entre les données de santé, fournissant ainsi aux utilisateurs des recommandations plus spécifiques et mieux adaptées. En outre, le filtrage collaboratif permet de recommander des biens ou des services en fonction

des préférences de l'utilisateur, ce qui accroît son bonheur et améliore son expérience. En outre, le SR DeepReco peut être intégré à divers systèmes de soins de santé afin de fournir des recommandations aux patients ou aux professionnels de santé. L'article présente toutefois quelques inconvénients, notamment l'impact potentiel des données de santé collectées sur la qualité des recommandations, la nécessité de disposer d'une importante collection de données de haute qualité pour piloter le modèle d'apprentissage en profondeur, ce qui nécessite des investissements en temps et en ressources, et la nécessité de protéger la vie privée des patients lors de la collecte, du stockage et de l'utilisation des données de santé.

Dans leur publication [39], les auteurs proposent une analyse détaillée d'une recommandation combinée pour la prévention des maladies cardiovasculaires basée sur l'Internet des objets. Les avantages d'un tel système comprennent la collecte de données plus précises et en temps réel sur la santé des patients, une meilleure personnalisation des recommandations pour les utilisateurs et une efficacité accrue dans la prévention des maladies cardiovasculaires. Cependant, l'article souligne également les risques potentiels tels que la confidentialité et la sécurité des données de santé collectées et analysées par le système IoT, les risques de biais dans les recommandations en raison de l'analyse limitée des données disponibles, et le coût potentiellement élevé de la mise en place et de la maintenance du SR IoT. L'article souligne l'importance de prendre en compte ces avantages et inconvénients lors de la mise en place de systèmes de recommandation IoT pour la prévention des maladies cardiovasculaires afin de maximiser les avantages pour les patients tout en minimisant les risques potentiels pour leur santé et leur vie privée.

Les avantages et les inconvénients d'un SR médicale personnalisé basé sur la factorisation hybride des matrices sont examinés par les auteurs [41] Le système iDoctor analyse les informations et les préférences de l'utilisateur en matière de santé afin de lui recommander des médicaments et des traitements. Ce système présente un certain nombre d'avantages, notamment des recommandations personnalisées, l'implication du patient dans le processus de guérison et une efficacité accrue de la méthode de recommandation. Cependant, il existe certains inconvénients potentiels, notamment la disponibilité limitée des données des patients, les problèmes de sécurité et de confidentialité des données, et les contraintes de généralisation du système. En conclusion, l'article souligne l'importance de peser soigneusement les avantages et les inconvénients des systèmes de recommandation médicale personnelle comme iDoctor, ainsi que la

nécessité de maintenir les systèmes à jour et de s'assurer qu'ils sont adaptés à une variété de populations de patients

Les avantages pour le secteur des soins de santé sont considérables, selon l'essai de l'auteur [42]. Le système permet une gestion préemptive de la santé des patients en collectant et en analysant les données de santé en temps réel à l'aide de l'internet des objets, ce qui a le potentiel d'améliorer considérablement les résultats des soins médicaux. L'utilisation d'une ontologie floue de type 2 permet une modélisation plus précise et plus souple des idées médicales complexes, ce qui améliore la spécificité des recommandations proposées. En outre, le SR offre aux patients des recommandations personnalisées basées sur leurs informations de santé et leurs préférences, ce qui augmente l'efficacité des soins médicaux en proposant des traitements adaptés à chaque patient. L'article met également en évidence certains inconvénients susceptibles de limiter l'adoption de ce type de système dans les établissements de santé, tels que le coût élevé de l'infrastructure IdO et la complexité de la mise en œuvre. De plus, la précision des recommandations peut être impactée par la qualité des données recueillies et les incertitudes associées à l'ontologie de floutage de type 2, ce qui peut compromettre l'efficacité du système. Enfin, lors de la mise en place d'un tel système, il est crucial de prendre en considération la confidentialité des données des patients. Malgré ces inconvénients, l'article de Farman Ali propose une nouvelle approche pour améliorer la qualité des services de santé et la satisfaction des patients, ce qui ouvre des perspectives prometteuses pour l'avenir des soins de santé.

Les auteurs dans leur article [43] présentent un nouveau SR mixte pour l'appariement précis des patients et des médecins dans les soins primaires. Le système fournit des résultats plus précis et plus complets en combinant plusieurs méthodes de recommandation, ce qui permet d'augmenter le niveau de satisfaction des clients et d'améliorer la qualité des soins. L'utilisation des informations sur la santé des patients et des préférences des médecins permet de personnaliser les recommandations et d'optimiser le jumelage, ce qui se traduit par des soins plus adaptés et plus appropriés. Toutefois, l'article met également en évidence certains inconvénients, tels que la nécessité de disposer de données de haute qualité et de critères de recommandation pertinents afin de garantir la précision du système. En outre, la mise en œuvre d'un tel système peut nécessiter des investissements en temps et en ressources pour la collecte et l'analyse des données, ce qui pourrait constituer un obstacle à son acceptation. Enfin, le SR peut être influencé par des biais inconscients ou des idées préconçues lors du

choix des critères de recommandation, ce qui souligne la nécessité de contrôler et d'évaluer en permanence les performances du système. Malgré ces obstacles, SR mixte décrit dans cet article a un potentiel significatif pour améliorer l'efficacité et la qualité des services de santé en mettant en relation les patients avec les médecins les mieux adaptés à leurs besoins et à leurs préférences.

L'étude des auteurs [44] met en évidence les avantages de l'utilisation d'un nombre limité d'abus de médias sociaux pour recommander des projets de start-up intelligents dans les domaines des villes intelligentes et de la santé intelligente. En effet, ce système permet de recueillir efficacement et en temps réel des données sur les demandes du marché et les tendances émergentes, ce qui améliore le taux de réussite des entreprises en identifiant les projets les plus pertinents. En outre, la personnalisation des recommandations en fonction des préférences de l'utilisateur peut accroître leur pertinence et leur utilité. Toutefois, la qualité des données recueillies sur les plateformes de médias sociaux, qui peuvent contenir des erreurs, des préjugés et des informations inexacts, peut avoir un impact sur la qualité des recommandations. Lors de la collecte, du stockage et de l'utilisation des données provenant des médias sociaux, la confidentialité des informations des utilisateurs doit également être prise en considération.

Notre travail s'inspire de diverses approches basées sur les profils des utilisateurs et d'une variété de méthodologies pour finalement atteindre notre objectif. Notre objectif est d'améliorer ces travaux existants et d'offrir des recommandations rapides et précises. Dans ce contexte, nous proposons un SR qui utilise les informations personnelles des profils d'utilisateurs obtenues. Ce système recommande des spécialistes en fonction des informations relatives au patient, notamment son âge, sa maladie et ses symptômes et l'emplacement géographique du patient.

Le profil de l'utilisateur est établi à partir de ses données personnelles et de son état de santé actuel. Ce profil change en fonction des événements observés, ce qui permet de fournir des recommandations d'experts adaptées à l'état de santé actuel de l'utilisateur dans une ville intelligente.

3.5 Conclusion

Dans ce chapitre, nous avons présenté l'état de l'art des contributeurs importants aux recommandations en matière de santé et analysé la manière dont les RS ont été utilisées dans les applications de santé. Nous avons examiné les différents travaux présentés par catégorie. Pour étudier ces catégories, nous avons d'abord identifié les critères qui les caractérisent. Enfin, nous avons utilisé ces critères pour évaluer les approches examinées. La discussion sur l'état actuel des lignes directrices a montré la nécessité de disposer de lignes directrices efficaces et plus claires. Dans le chapitre suivant, nous décrivons notre approche en détail.

Chapitre 4

Contributions

4.1 Introduction

Des centaines de millions de personnes tombent malades chaque année, et lorsqu'elles sont confrontées à un problème de santé, la première chose qu'elles recherchent est un médecin qui possède les meilleures compétences pour résoudre leur problème de santé. Ce besoin est encore plus crucial lorsqu'un touriste tombe malade lors de son séjour dans une ville étrangère. Le touriste se retrouve confronté à la barrière de la langue, à l'absence de connaissances locales et à la difficulté de trouver un médecin compétent dans un environnement inconnu.

Dans ce mémoire, nous présentons un SR capable de suggérer les médecins qui correspondent le mieux au profil d'un patient spécifique et qui possèdent les compétences et l'expertise nécessaires pour résoudre ce problème. Cependant, compte tenu du nombre de professionnels de santé disponibles et des différentes spécialités médicales, il peut être extrêmement difficile pour un patient, en particulier un touriste, de choisir le praticien ou l'établissement qui répond le mieux à ses besoins dans une ville inconnue.

C'est pourquoi nous concevons un SR utilisant la technique du FC, basée sur l'utilisateur, pour les profils qui sont similaires aux profils déjà existants dans la base de données, en utilisant la similarité de Jaccard [46]. De plus, nous intégrerons également le filtrage basé sur les connaissances dans le cas d'un démarrage à froid, afin de proposer des recommandations même en l'absence de données utilisateur préalables dans une nouvelle ville.

Ce chapitre expose donc l'approche proposée, nous présentons également l'architecture générale du système, accompagnée de diagrammes utiles pour la

conception, ainsi que les différentes étapes de la construction de notre système de recommandation.

4.2 Approche proposée

Dans cette section, nous décrivons en détail notre approche pour développer notre SR basé sur le filtrage collaboratif (FC) par les utilisateurs et le filtrage basé sur la connaissance[47]. Notre SR est relatif au domaine de la santé, avec pour objectif d'assister les touristes qui tombent malades pendant leur séjour à choisir un médecin de manière plus éclairée.

L'architecture du système est présentée dans **la Figure 4.1** et constitue les étapes suivantes :

- Récupération des informations de l'utilisateur et création de son profil : Cette étape consiste à collecter les informations fournies par l'utilisateur afin de créer son profil.
- Classification de l'utilisateur en fonction de son historique et de ses antécédents médicaux : La deuxième étape implique la classification de l'utilisateur en se basant sur son historique médical et ses antécédents.
- Recherche d'un profil similaire en utilisant la similarité de Jaccard : Cette étape vise à trouver des profils similaires en comparant le profil créé lors de la première étape avec les profils des médecins. L'objectif est de rechercher des recommandations pertinentes.
- Regroupement des profils similaires et des médecins spécialisés : Cette étape regroupe les profils similaires ainsi que les médecins spécialisés afin d'accélérer le processus de recommandation.
- Affichage de la liste des médecins similaires et trouver le médecin le plus proche à l'aide de l'algorithme Astar [48].
- L'utilisateur attribue une note au médecin qui lui a été recommandé.
- Recommandations : Cette dernière étape représente le processus complet de recommandation, où des suggestions spécifiques sont générées en fonction des profils et des médecins regroupés précédemment.

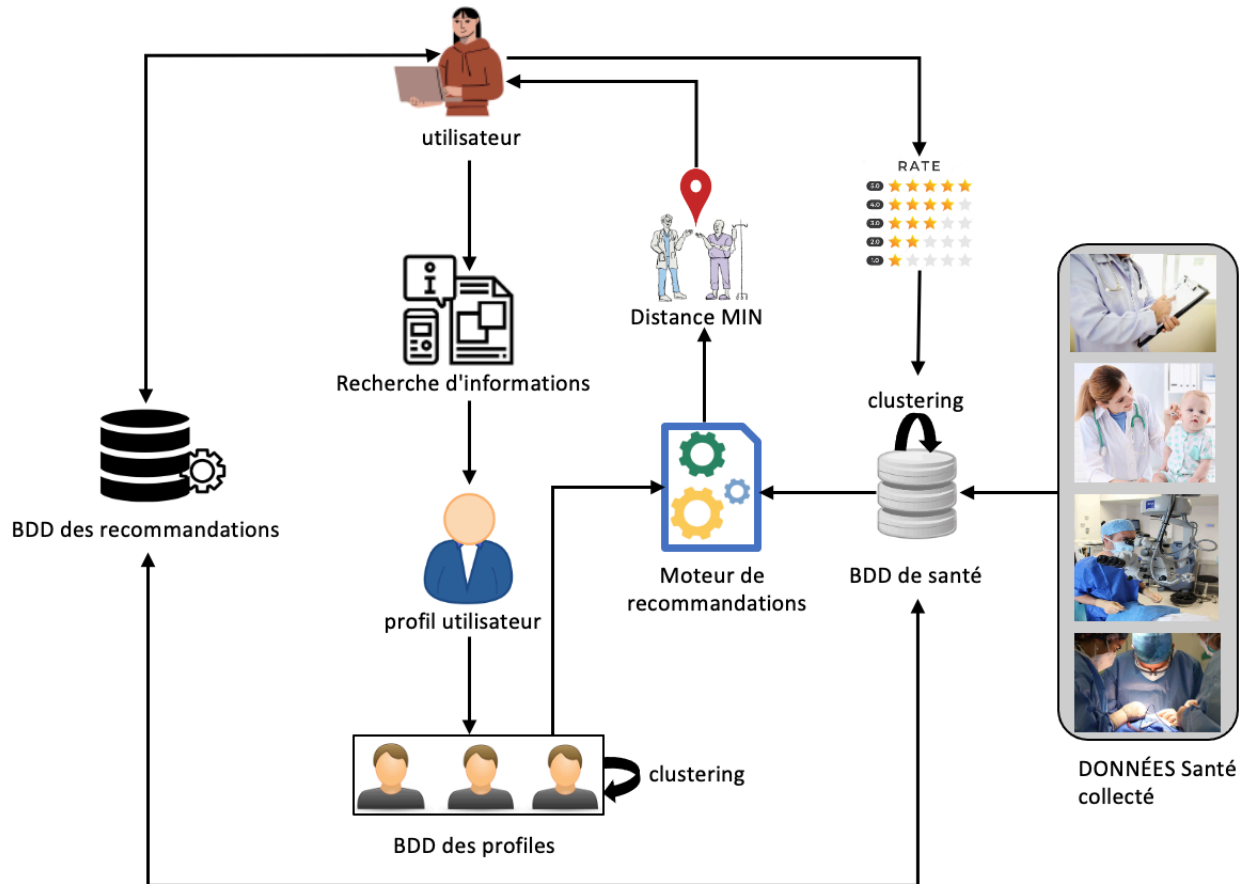


Figure 4.1 – Approche proposée

Pour atteindre cet objectif, nous avons développé une approche soutenue par un outil logiciel qui permet de comprendre les différentes étapes de notre approche. **L’algorithme 1** décrit la séquence des différentes étapes du processus de recommandation.

Entrée :

- Données du patient et les antécédents médicaux

Sortie :

- Recommandation des médecins spécialiste

Debut

1. Collecter les données de l'utilisateur, notamment son nom, son âge, ses antécédents médicaux et ses symptômes.
2. Créer le profil de l'utilisateur en utilisant les données collectées.
3. Comparer le profil de l'utilisateur avec les profils existants des médecins spécialistes disponibles pour traiter sa maladie.
4. Si un profil de médecin correspondant est trouvé, aller à l'étape 6.
5. Sinon, effectuer une recherche supplémentaire pour trouver des médecins spécialistes qui traitent spécifiquement la maladie de l'utilisateur.
6. Procéder à la recommandation du médecin spécialiste le mieux adapté à l'utilisateur, en tenant compte de facteurs tels que la réputation du médecin la distance géographique, la disponibilité
7. Si le système signale un problème de démarrage à froid (difficulté à trouver un médecin) dû au manque d'informations, passer à l'étape 8. Sinon, passer à l'étape 9.
8. Utiliser une approche basée sur les connaissances pour résoudre le problème de démarrage à froid en utilisant des connaissances médicales et des ressources supplémentaires
9. Délivrer les résultats de la recommandation au patient, y compris la disponibilité, l'emplacement géographique, les coordonnées du médecin spécialiste recommandé.

Fin

Algorithm 1 : Conception du SR

Dans ce qui suit, nous allons résumer et présenter les étapes de notre approche en détails :

4.2.1 Création du profil utilisateur

Pour créer le profil utilisateur, plusieurs méthodes sont utilisées en se basant sur les informations collectées. Les données sont recueillies de deux manières

différentes : les données explicites, qui sont basées sur les réponses fournies par l'utilisateur lors de la saisie de son profil, et les données implicites, qui sont captées à partir de dispositifs IdO (Internet des Objets). Ces données incluent des informations physiologiques telles que la taille, la fièvre, le poids, la fréquence cardiaque ect. En utilisant les dispositifs IdO sensoriels, nous sommes en mesure d'obtenir des données réelles, ce qui permet d'améliorer l'efficacité des recommandations.

Le processus de création du profil utilisateur commence par la collecte des informations à partir d'un formulaire d'inscription. Ce formulaire comprend des champs tels que le nom, le prénom, l'âge et d'autres détails pertinents tels que les antécédents médicaux, les préférences personnelles et les coordonnées du patient. Une fois que toutes ces informations sont fournies, le profil utilisateur est créé. Ce profil est ensuite enregistré dans la base de données des utilisateurs, ce qui permet de conserver toutes les informations pertinentes pour les recommandations ultérieures.

4.2.2 Recherche de profils similaires

Notre approche vise à fournir des recommandations personnalisées à tous les utilisateurs, à condition qu'un processus de recommandation ait déjà été effectué et que les profils des médecins soient enregistrés dans une base de données dédiée.

L'objectif principal de cette étape est de trouver un médecin similaire en utilisant les informations pertinentes du profil utilisateur, telles que l'âge, le sexe, le diagnostic, les antécédents médicaux et les symptômes sélectionnés.

Cela implique d'utiliser ces informations du profil utilisateur pour effectuer une correspondance avec les profils des médecins dans la base de données. Les profils des médecins devraient contenir des détails tels que leur spécialité, les symptômes spécifiques à leurs spécialités, leurs disponibilités, leur expérience, etc.

Pour évaluer la similarité entre le profil utilisateur et les profils des médecins, nous utilisons une mesure adaptée, telle que la similarité de Jaccard, qui nous permet d'estimer la proximité entre les caractéristiques et les attributs du profil utilisateur et ceux des profils des médecins, afin de déterminer les corres-

pondances les plus pertinentes.

Nous avons choisi la similarité de Jaccard pour notre système en raison de plusieurs avantages clés qu'elle offre. Tout d'abord, sa simplicité en fait une mesure facile à comprendre et à implémenter. De plus, la similarité de Jaccard prend en compte à la fois les éléments communs et les éléments uniques dans les ensembles comparés. Cela signifie qu'elle peut capturer à la fois les similitudes et les différences entre les textes, les phrases ou les mots. De plus la similarité de Jaccard est largement utilisée dans le domaine du traitement du langage naturel en raison de sa capacité à mesurer efficacement la similarité entre des textes.

Dans le contexte de la recommandation de médecins, la similarité de Jaccard peut être utilisée pour mesurer la similitude entre les ensembles de symptômes, d'antécédents médicaux ou de spécialités des médecins consultés. En calculant la similarité de Jaccard entre les ensembles, nous pouvons évaluer à quel point ils sont similaires. Plus la similarité de Jaccard est élevée entre deux utilisateurs, plus ils ont de caractéristiques communes. Par conséquent, il est plus probable qu'ils partagent les mêmes besoins en matière de soins de santé.

La similarité de Jaccard est définie comme le nombre d'éléments communs entre les deux ensembles divisé par le nombre total d'éléments uniques dans les deux ensembles.

La formule de **la similarité de Jaccard** est :

$$\text{Sim (A, B)} = |A \cap B| / |A \cup B|$$

A et B sont deux ensembles, $|A|$ et $|B|$ représentent le nombre d'éléments dans chaque ensemble, et $|A \cap B|$ représente le nombre d'éléments communs à A et B.

Le résultat de la formule est un nombre compris entre 0 et 1. Un résultat de 0 signifie qu'il n'y a pas d'élément en commun entre les deux ensembles, tandis qu'un résultat de 1 signifie que les deux ensembles sont identiques.

La complexité de la similarité de Jaccard dépend de la taille des ensembles A et B, représentées respectivement par n et m. Dans le pire des cas, où les ensembles sont représentés sous forme de listes non triées, le calcul de l'intersection nécessite une comparaison de chaque élément de l'ensemble A avec chaque élément de l'ensemble B. Cela entraîne une complexité de $O(n*m)$.

Plus précisément, on peut construire un ensemble de symptômes associés

à la maladie du patient, puis comparer cet ensemble avec les ensembles de symptômes correspondant aux domaines d'expertise des différents médecins. La similitude de Jaccard peut ensuite être utilisée pour classer les médecins en fonction de leur similarité avec le patient et recommander les médecins les plus similaires en premier. La **figure 4.2** suivante montre un schéma de calcul de prédiction.

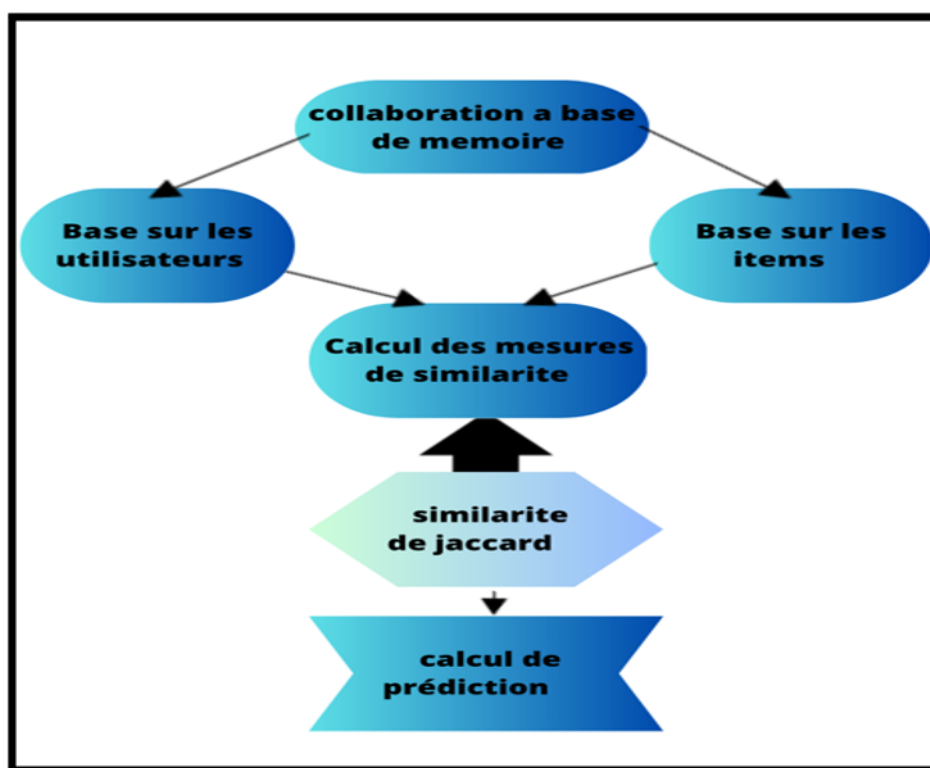


Figure 4.2 – Calcul de prédiction

L'algorithme 2 représente les différentes étapes par lesquels passe la similarité de Jaccard pour trouver les profils similaires.

Entrée :

- P : Profil de l'utilisateur (informations sur la maladie, les antécédents médicaux, etc.)
- BDDP : Base de données des profils de médecins (informations sur les spécialités, les expériences, les domaines d'expertise, etc.)
- S : Seuil de confiance pour le calcul de la similarité de Jaccard.

Sortie :

- Taux de similarité de Jaccard entre le profil de l'utilisateur et les profils de médecins similaires.

Début

1. Obtenir le profil de l'utilisateur.
2. Pour chaque profil de médecin i dans BDDP faire
 - a) Récupérer les informations de spécialité et les ensembles de symptôme qu'ils ont traités dans le passé, d'expérience et de domaine d'expertise du médecin.
 - b) Calculer la similarité de Jaccard entre le profil de l'utilisateur et le profil du médecin en utilisant les informations de spécialité, d'expérience et de domaine d'expertise.
 - c) Si la similarité de Jaccard est supérieure ou égale au seuil S , ajouter le médecin au cluster des médecins similaires.
3. Retourner le taux de similarité de Jaccard entre le profil de l'utilisateur et les profils des médecins similaires.

Fin

Algorithm 2 : Comparaison des profils

Dans ce qui suit nous expliquant notre approche à l'aide d'un exemple :

Exemple 1

Pour illustrer notre approche, prenons l'exemple d'un touriste qui tombe malade lors de son voyage, présentant les symptômes suivants représentés par l'ensemble A :

$A =$ fièvre, toux, maux de tête, fatigue, Difficulté à respirer.

Nous avons une base de données de 5 médecins spécialistes (M_1 , M_2 , M_3 , M_4 et M_5) et les ensembles de symptômes qu'ils ont traités dans le passé sont

les suivants :

- M1 : fièvre, toux, maux de tête, fatigue
- M2 : toux, maux de tête, fatigue, diarrhée, Difficulté à respirer
- M3 : fièvre, toux, maux de tête, Difficulté à respirer
- M4 : fièvre, toux, fatigue
- M5 : maux de tête, fatigue, Difficulté à respirer, diarrhée

Le calcul de la similarité de Jaccard entre les ensembles de symptômes du patient (A) et ceux des médecins spécialistes (M1-M5) se fait comme suit :

- $\text{Sim}(A, M1) = |A \cap M1| / |A \cup M1| = 4 / 5 = 0.8$
- $\text{Sim}(A, M2) = |A \cap M2| / |A \cup M2| = 4 / 6 = 0.67$
- $\text{Sim}(A, M3) = |A \cap M3| / |A \cup M3| = 4 / 5 = 0.8$
- $\text{Sim}(A, M4) = |A \cap M4| / |A \cup M4| = 3 / 5 = 0.6$
- $\text{Sim}(A, M5) = |A \cap M5| / |A \cup M5| = 3 / 6 = 0.5$

Dans cet exemple, les médecins M1 et M3 présentent la plus haute similarité de Jaccard avec le touriste malade, obtenant une valeur de 0.8. Cette constatation met en évidence que les symptômes spécifiques liés à leurs spécialités sont ceux qui correspondent le plus étroitement aux symptômes sélectionnés par le malade, ce qui les rend particulièrement adaptés pour traiter ce patient spécifique.

L'utilisation de la méthode de similarité de Jaccard nous permet ainsi d'identifier de manière précise et fiable les médecins les plus pertinents pour le touriste malade. Cette méthode nous offre une approche objective pour évaluer la similitude des ensembles de symptômes spécifiques à chaque spécialité médicale. Elle s'avère donc essentielle dans le processus de sélection des médecins les mieux qualifiés pour répondre aux besoins de chaque patient.

Afin de mieux illustrer cette similarité, nous avons représenté les médecins spécialistes (M1, M2, M3, M4 et M5) sous forme de barres dans un graphique. Les barres représentent la similarité de Jaccard entre l'ensemble de symptômes du patient (A) et ceux des médecins spécialistes.

Chaque médecin spécialiste est associé à une barre, dont la hauteur est proportionnelle à la similarité de Jaccard calculée précédemment. Ainsi, plus la barre est haute, plus la similarité entre les ensembles de symptômes du malade et du médecin spécialiste est élevée.

Dans notre exemple, les barres correspondantes aux médecins M1, M2, M3, M4 et M5 ont des hauteurs respectives de 0.8, 0.67, 0.8, 0.6 et 0.5. Cela nous permet de visualiser rapidement les médecins qui présentent la plus haute similarité avec le malade. **La figure 4.3** représente le graphe qui illustre la similarité de Jaccard pour chaque médecin :

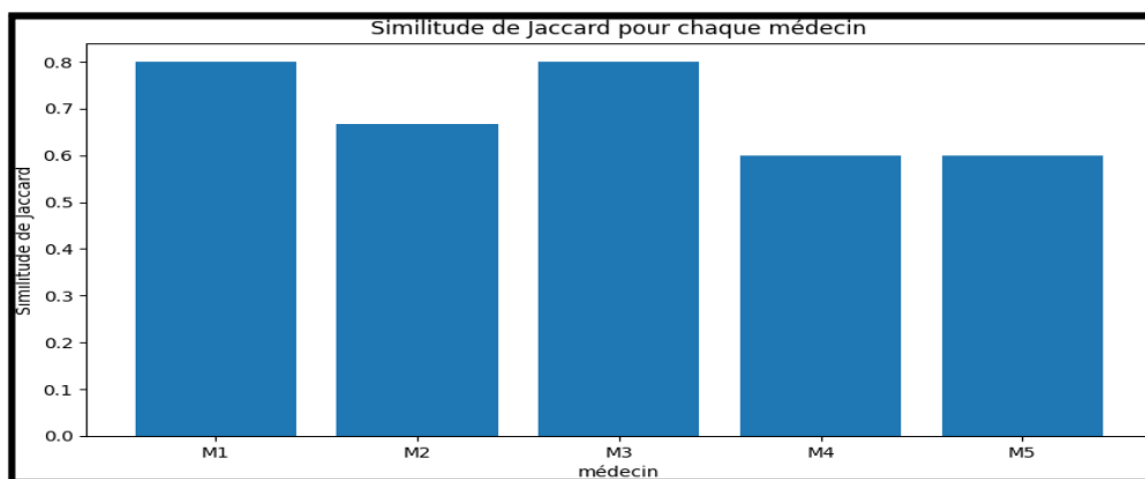


Figure 4.3 – Similitude de jaccard pour chaque médecin

Cette représentation graphique des barres facilite la comparaison visuelle entre les médecins spécialistes et leur similarité respective avec le malade en question. Elle offre une vue d'ensemble claire et concise de la pertinence des médecins dans le contexte donné, en mettant en évidence ceux dont les ensembles de symptômes se rapprochent le plus de ceux du patient.

Dans cet exemple les médecins M1 et M3 présentent la plus haute similarité de Jaccard avec le touriste malade, atteignant une valeur de 0,8.

4.2.3 Clustering

L'apprentissage non supervisé est un paradigme d'apprentissage automatique (ou Machine Learning) qui consiste à extraire des motifs ou des structures à partir de données non étiquetées, c'est-à-dire des données qui ne sont pas associées à des étiquettes ou à des sorties connues. Il existe plusieurs techniques d'apprentissage non supervisé, telles que le clustering, la réduction de dimensionnalité, la détection d'anomalies et la génération de données.... L'apprentissage non supervisé est utilisé dans une variété de domaines, notamment la reconnaissance de formes, l'analyse de données, la bio-informatique et le domaine médical...

Dans notre mémoire, nous appliquons l'algorithme de clustering K-means [49] et une approche de l'algorithme Particle Swarm Optimization (PSO) [50].

Le clustering est une méthode d'apprentissage automatique qui permet de regrouper des points de données en des groupes homogènes appelés clusters, de manière à ce que les éléments d'un même cluster soient similaires. L'algorithme K-Means est utilisé dans notre mémoire pour réaliser le clustering.

Dans le contexte du clustering, nous utilisons une approche combinée de l'algorithme Particle Swarm Optimization (PSO) et de l'algorithme K-Means pour améliorer la précision et l'efficacité du processus. La figure suivante montre les résultats d'application de la méthode de clustering K-means et PSO.

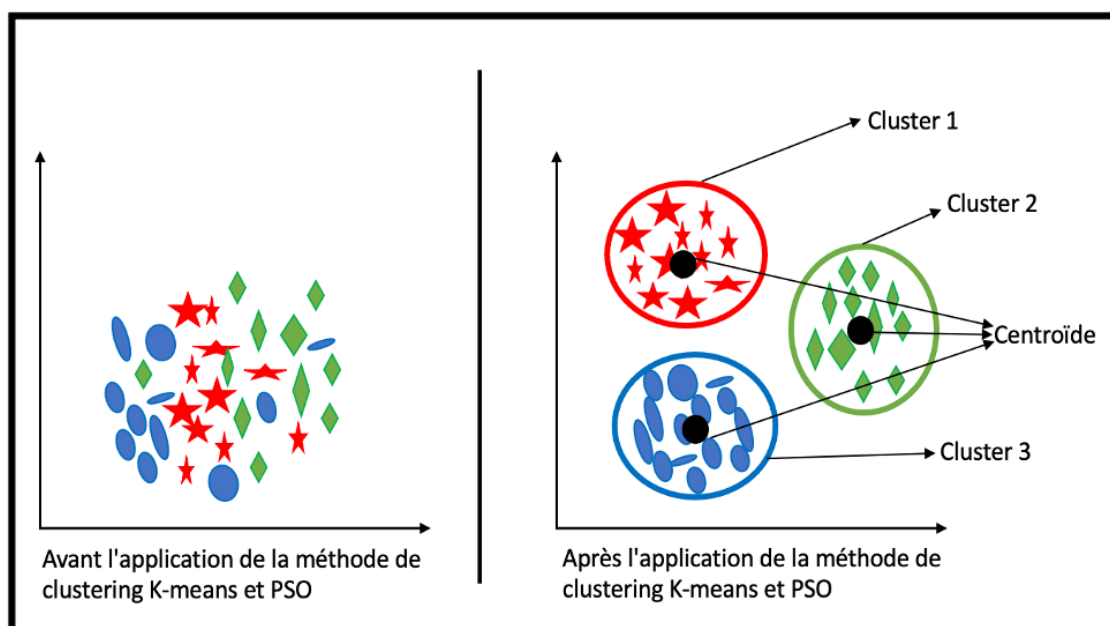


Figure 4.4 – Initialisation optimale des K-means à l'aide de Particle Swarm Optimization

L'algorithme PSO, inspiré du comportement collectif des oiseaux et des poissons, est utilisé pour optimiser la recherche de la meilleure solution dans un espace de recherche multidimensionnel. Il fonctionne en créant un ensemble de particules représentant des solutions potentielles, qui se déplacent et ajustent leur position et leur vitesse en fonction de leur expérience individuelle et des informations partagées avec les autres particules. Les particules communiquent pour partager les meilleures positions trouvées, ce qui guide leur exploration coordonnée de l'espace de recherche. L'objectif est de trouver la meilleure solution en optimisant les positions des particules, en utilisant à la fois l'expérience individuelle et la connaissance collective du groupe. L'algorithme PSO est large-

ment utilisé dans divers domaines pour résoudre des problèmes d'optimisation, en raison de sa simplicité et de son efficacité pour trouver des solutions dans des espaces de recherche complexes.

K-Means et PSO

Dans le domaine de la santé, cette approche combinée de K-Means et PSO peut être utilisée pour recommander des médecins spécialistes aux patients en fonction de leurs symptômes, de leur profil de santé et de leurs préférences de traitement.

Tout d'abord, l'algorithme K-Means est appliqué sur un ensemble de données contenant des informations sur les médecins, telles que leurs compétences, leurs expériences et leurs spécialités. K-Means divise ces données en K clusters, où K est le nombre prédéfini de spécialités médicales. Chaque cluster représente un groupe de médecins partageant des caractéristiques similaires.

Ensuite, pour chaque cluster de médecins, l'algorithme PSO est utilisé pour sélectionner les symptômes les plus informatifs à considérer dans ce groupe spécifique. Les particules dans PSO représentent différentes combinaisons de symptômes potentiels. En ajustant les positions et les vitesses des particules dans l'espace de recherche, le PSO cherche à trouver la combinaison optimale de symptômes qui maximise la similarité entre les symptômes des patients et les critères des médecins spécialistes de ce cluster.

Les particules communiquent entre elles pour partager les meilleures positions trouvées jusqu'à présent, permettant une exploration coordonnée de l'espace de recherche. Cette exploration guidée par le PSO permet de sélectionner les symptômes les plus informatifs et pertinents pour chaque spécialité médicale, en tenant compte des critères spécifiques des médecins du cluster.

En utilisant cette approche combinée de K-Means et PSO, on peut regrouper les médecins selon leurs spécialités et sélectionner les symptômes les plus informatifs au sein de chaque cluster de médecins. Cela permet d'améliorer la recommandation médicale en tenant compte à la fois des spécialités des médecins et des symptômes pertinents pour chaque groupe de spécialités.

4.2.4 Recommandation

Le moteur de recommandation passe principalement par deux cas (**voir L'algorithme 3**) :

Entrée :

- IPU : Informations du profil de l'utilisateur (nom, prénom, âge, sexe, type de maladie)
- T : Taux de comparaison
- Cordonnées géographique du patient

Sortie :

- MR : Médecins recommandés

Debut**1. Si T est proche de 1 alors :**

- a. Récupérer l'ancienne recommandation pour le même profil dans la base de données des recommandations médicales.
- b. Renvoyer cette recommandation au patient.

2. Sinon :

- a. Récupérer les IPU depuis le profil patient.
- b. Utiliser l'algorithme de recommandation spécialisé pour trouver les médecins spécialistes de la maladie.
- c. Pour chaque médecin M :
 1. Calculer la similarité de Jaccard entre les symptômes du patient et ceux de M.
 2. Si la similarité est supérieure ou égale à T alors :
 - Sélectionner le chemin le plus courts entre le patient et M à l'aide de A*.
 - Ajouter M à la liste des médecins recommandés avec sa distance géographique.
- e. Faire la recommandation et retourner les MR.

Fin

Algorithm 3 : Recommandation

Lorsque le système de recommandation n'a pas suffisamment de données utilisateur pour fournir des recommandations personnalisées, il peut utiliser une approche de filtrage basée sur les connaissances pour suggérer des médecins spécialistes appropriés. Cette approche est basée sur les connaissances médicales et les compétences des médecins, ainsi que sur d'autres informations per-

tinentes telles que les spécialités, les qualifications, les évaluations des patients et les informations sur les établissements de santé.

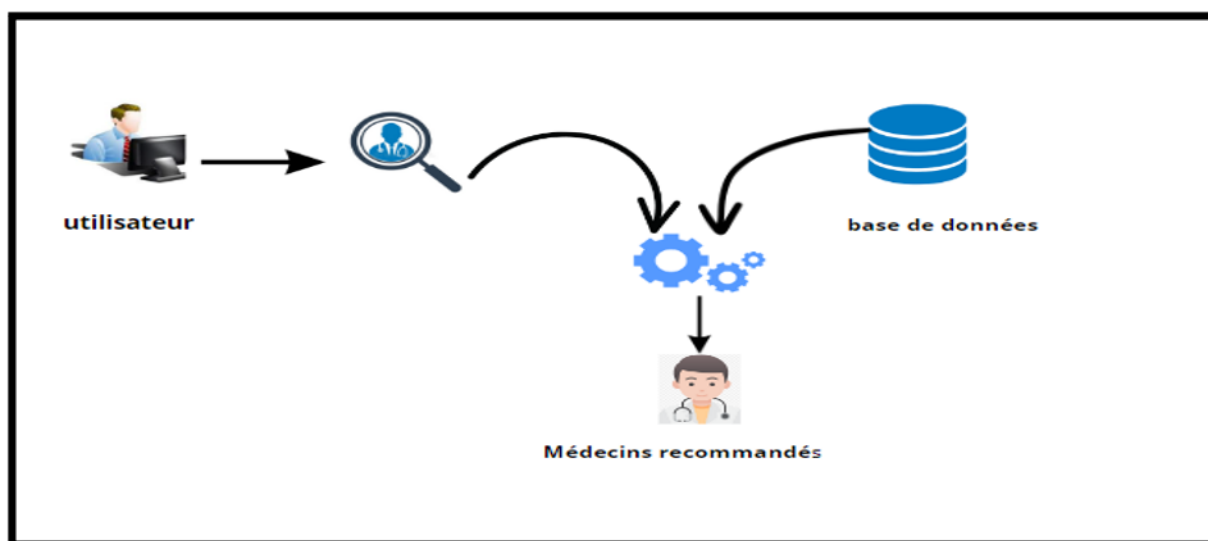


Figure 4.5 – Système de recommandation à base de connaissance

En utilisant cette approche, le système de recommandation peut fournir des recommandations pertinentes et fiables, même en l'absence de données utilisateur. Par exemple, si un nouveau patient s'inscrit sur le système et ne fournit pas suffisamment d'informations sur son profil de santé, le système peut utiliser les connaissances médicales et les compétences des médecins pour suggérer des médecins.

Cette approche de filtrage basée sur les connaissances peut également être utile dans les cas où les données utilisateur sont limitées ou non disponibles. Par exemple, si un patient recherche des médecins spécialisés dans une maladie rare, le système peut utiliser les connaissances médicales pour suggérer des médecins.

Dans la suite, nous allons poursuivre l'explication de **l'exemple 1** de notre approche.

Afin d'optimiser davantage la recommandation, nous proposons de calculer la distance la plus courte entre la localisation géographique du malade et celle des médecins similaires. Après avoir obtenu la liste des médecins similaires en utilisant la similarité de Jaccard, cette considération de la proximité géographique permet d'optimiser la prise en charge du malade en minimisant la distance à

parcourir. Cela permet au malade de bénéficier d'une prise en charge plus rapide et d'une meilleure réactivité médicale.

Afin d'ajouter la localisation à cette démarche, il est nécessaire de disposer des informations de localisation du touriste malade ainsi que des médecins spécialistes afin d'intégrer ce critère dans le processus de recommandation. Le graphe de la **figure 4.6** illustre cette situation :

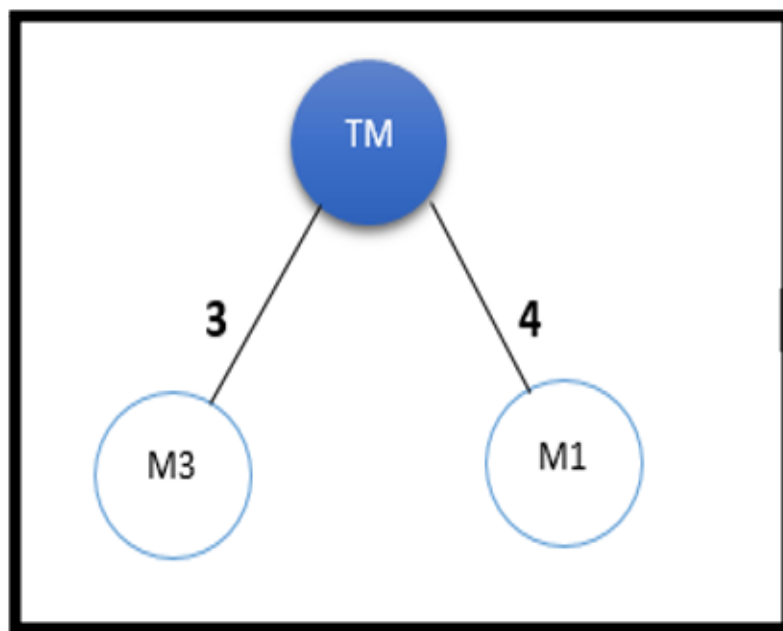


Figure 4.6 – Graphe distance

Dans le graphe suivant :

- "TM" désigne le touriste malade,
- "M1" correspond au médecin 1.
- "M3" correspond au médecin 3.

Les nombres le long des arêtes indiquent les distances nécessaires pour se déplacer d'un nœud à un autre :

- L'arête reliant "TM" à "M1" a une valeur de 4.
- L'arête reliant "TM" à "M3" a une valeur de 3.

Ces distances s

Entrée :

- Nœud de départ (TM), Nœud de destination (M1 ou M3)

Sortie :

- Chemin optimal du nœud de départ au nœud de destination

Debut

1. Définir le nœud de départ (TM) et le nœud de destination (M1 ou M3).
2. Initialiser la liste des nœuds ouverts avec le nœud de départ (TM) et un coût initial de 0.
3. Initialiser la liste des nœuds fermés comme étant vide.
4. Tant que la liste des nœuds ouverts n'est pas vide :
 - a. Sélectionner le nœud avec le coût total le plus bas parmi les nœuds ouverts.
 - b. **Si** le nœud actuel est le nœud de destination (M1 ou M3), terminer l'algorithme.
 - c. **Sinon**, déplacer le nœud actuel de la liste des nœuds ouverts vers la liste des nœuds fermés.
 - d. Pour chaque voisin du nœud actuel :
 - Calculer le coût pour atteindre ce voisin depuis le nœud de départ (TM).
 - Calculer l'estimation heuristique du coût restant pour atteindre le nœud de destination (M1 ou M3).
 - Calculer le coût total en additionnant le coût réel et l'estimation heuristique.
 - Si le voisin est déjà dans la liste des nœuds ouverts avec un coût total inférieur, ignorer ce voisin.
 - Sinon, mettre à jour le voisin avec le nouveau coût total et le nœud actuel comme parent.
 - e. Ajouter le nœud actuel à la liste des nœuds fermés.
5. Si la liste des nœuds ouverts est vide et le nœud de destination n'a pas été atteint, il n'existe pas de chemin possible.

Fin

Algorithm 4 : A*MD

Ces distances sont essentielles pour évaluer les chemins possibles et sélec-

tionner le chemin le plus court en fonction de ces coûts à l'aide de l'algorithme A* Multiple Destinations ou "A* MD" (A-star Multiple Destinations).

Pour calculer le chemin le plus court entre le touriste malade (TM) et les médecins (M1 et M3) dans le graphe donné, nous allons utiliser l'algorithme A* MD. En utilisant les valeurs de distance fournies. Voici les étapes détaillées :

Donc en utilisant l'algorithme A* avec les valeurs données, le touriste malade (TM) va choisir le médecin M3 comme destination car le coût total (3) pour atteindre M3 est inférieur au coût total (4) pour atteindre M1.

4.3 Diagramme d'activité

A la fin le schéma ci-dessous la figure 4.7 présente le diagramme d'activité de notre système de recommandation, qui décrit les différentes étapes que l'utilisateur devra suivre et que le système exécutera afin d'obtenir une recommandation.

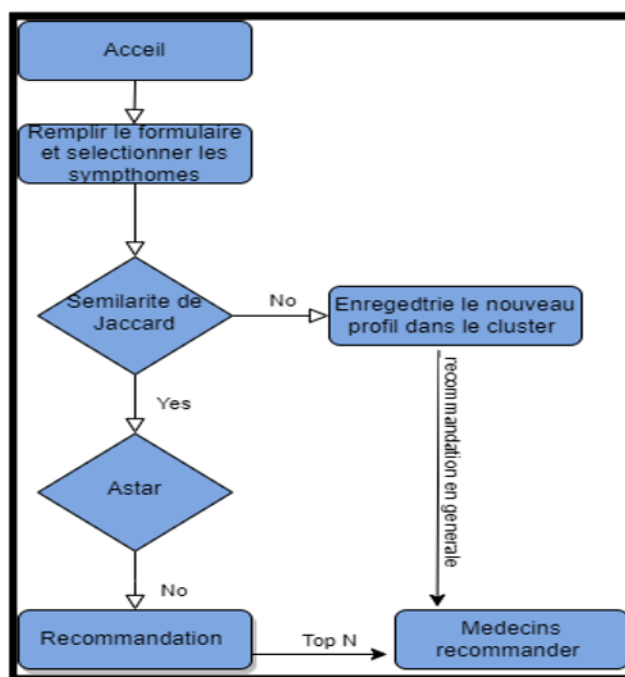


Figure 4.7 – Diagramme d'activité

- **Accueil** : L'accueil représente une interface conviviale qui accueille l'utilisateur sur notre site. Il fournit des informations sur les services ou les

fonctionnalités offertes par le site. Depuis l'accueil, l'utilisateur peut accéder et remplir le formulaire utilisateur.

- **Remplir le formulaire et sélectionner les symptômes** : Dans cette étape, l'utilisateur accède au formulaire et le remplit avec toutes ses informations personnelles. Le formulaire peut contenir des champs tels que le nom, l'âge, le sexe, les antécédents médicaux, etc. Une fois que l'utilisateur a saisi toutes les informations requises, il procède à la sélection des symptômes parmi les options proposées.
- **La similarité de Jaccard** : Cette méthode est utilisée pour comparer les ensembles de symptômes de l'utilisateur avec les ensembles de symptômes clés de chaque spécialité médicale.
- **Astar** : si les symptômes de l'utilisateur existent dans la base de données, le système utilise l'algorithme Astar pour calculer la distance géographique optimale entre l'utilisateur et les médecins disponibles, afin de fournir les N meilleures recommandations basées sur cette proximité géographique.
- **Enregistrer le nouveau profil dans le cluster** : lorsque les symptômes de l'utilisateur ne correspondent à aucune spécialité spécifique ou que les informations fournies sont insuffisantes, le système enregistre le nouveau profil dans un cluster dédié et propose une liste générale de médecins pour aider l'utilisateur dans sa recherche médicale.
- **Recommandation** : Le système de recommandation proposera à l'utilisateur des médecins en fonction de sa maladie et de ses critères de recherche.

4.4 Conclusion

Dans ce chapitre, nous avons présenté notre approche de recommandation pour les villes intelligentes, basée sur le filtrage collaboratif. Nous avons décrit les étapes clés de la modélisation de notre système et exposé en détail son fonctionnement.

Notre système de recommandation apporte une contribution significative dans le contexte des villes intelligentes. Il vise à améliorer l'accès aux services de santé en recommandant les professionnels les mieux adaptés aux besoins de chaque individu.

Dans le prochain chapitre, nous fournirons une description détaillée des aspects techniques liés à l'implémentation et à la mise en œuvre de notre système.

Nous expliquerons comment nous avons concrètement mis en place notre système en présentant les choix d'architecture, les outils de développement utilisés, les différentes étapes de déploiement, ainsi que les tests et évaluations effectués pour évaluer les performances de l'approche proposée.

Chapitre 5

Expérimentation

5.1 Introduction

Dans le cadre de notre recherche, nous avons choisi d'utiliser l'apprentissage non supervisé pour recommander des médecins spécialistes aux touristes atteints d'une maladie pendant leur séjour. Le processus de classification sera effectué sur un ensemble de touristes malades afin de les regrouper en différents types de maladies.

Dans ce chapitre, nous allons décrire le jeu de données utilisées pour former notre système, ainsi que les outils matériels et logiciels utilisés lors du développement de notre application. Nous présenterons également le langage de programmation utilisé.

Enfin, nous exposerons et expliquerons les différentes interfaces de notre application web appelée "DocFast", ainsi que la métrique de classification utilisée pour évaluer les performances de notre système.

5.2 Description du Dataset

Le terme "Dataset" est souvent utilisé dans le domaine de la science des données pour désigner une source de données qui contient des lignes et des colonnes de données. Les lignes représentent des points de données individuels et les colonnes représentent les caractéristiques de ces points de données. Les ensembles de données peuvent être stockés dans différents formats tels que CSV, TSV, Excel, une table dans une base de données ou un output d'un service Web.

Lors de l'utilisation d'un ensemble de données pour entraîner un modèle, il est important de déterminer les caractéristiques les plus importantes et celles qui

peuvent être ignorées en toute sécurité. Cela peut être une tâche difficile dans les grands ensembles de données qui peuvent contenir des milliers de caractéristiques et de lignes. Dans de tels cas, il est souvent nécessaire de faire appel à un expert du domaine qui comprend l'importance des caractéristiques, leurs interdépendances et la validité des valeurs de données pour les caractéristiques.

Pour rendre la recommandation pertinente, nous avons créé deux jeux de données.

Le premier jeu de données comprend environ 118 médecins spécialisés dans différents domaines tels que l'ORL, la chirurgie urologique, l'ophtalmologie, la cardiologie, la pneumologie, etc. Les informations ont été collectées à partir du site web de « la Direction de la Santé et de la Population de la Wilaya de Bejaia ».

Nous avons soigneusement sélectionné les médecins en fonction de leur disponibilité et de leur adresse de localisation. De plus, nous avons inclus des informations telles que leur numéro de téléphone dans ce premier jeu de données.

Le 1er jeu de données est organisé en une table **figure 5.1** avec les suivantes colonnes :

1. **Spécialité** : la spécialité du médecin (par exemple : pédiatre).
2. **Nom et prénom** : le nom de famille du médecin (par exemple : ZERATI Mouloud).
3. **Commune** : la localité où le médecin exerce sa profession (Bejaia).
4. **Adresse** : l'adresse du cabinet médical où le médecin exerce (par exemple : 9D, rue Ahcène DEHAS).
5. **Numéro de téléphone** : le numéro de téléphone du cabinet médical où le médecin exerce (par exemple : + 213 55 93 45 67 89).
6. **Disponibilité** : indique si le médecin est disponible pour prendre de nouveaux patients (oui ou non).

specialites	medecins	Adresse	Ville	Telephone	Disponible
ORL	• ABBACI A/Madjid	• 14 Rue Ouabdelkader	• BEJAIA	• 034.21.15.33	• True
	• HADJI Soraya	• Cite 300 logts BT C8 n 71hadadene	• BEJAIA	• /	• False
	• HARRAR Small	• Cite Rabea BT A 1 escalier 01 n 02	• BEJAIA	• 034.20.11.13	• True
	• IOUKNANE Djamel	• Cite aouchiche Forum de l'Universite route de BOUKHIAMA	• BEJAIA	• 034.22.68.38	• True
	• MEDJMADJ Brahim	• Cite des 96 logts BT A n 1 EL HAMMADIA	• BEJAIA	• /	• True
	• OUYAHIA Fouad	• Cite EDIMCO rue tekamira bloc n 8	• BEJAIA	• /	• True
	• HAMMOUDI Djamilia	• 08 rue frere BAOUZIA	• BEJAIA	• /	• True
Chirurgie-urologique	• ACHOUR Mouloud	• Residence la Grande Baie n 4 cite Tobbal	• BEJAIA	• 034.21.23.95	• False
	• BOUCHEFA Mokhtar	• Cite des HAMMADITES residence 111 SIDI- AHMED	• BEJAIA	• 034.21.96.74	• True
	• HOCINI Fares	• Bt n 08 Cite DJAMA Abdelkader RDC Madala	• BEJAIA	• /	• False
	• DAHMANI Karim Omar Kheirou	• 21, cite zerara nouvelle	• BEJAIA	• 034.21.44.92	• True
Rhumatologie	• ADMAM Bouzid	• 24 chemin des cretes	• BEJAIA	• 034.20.78.18	• True
	• BAKOURI Mahmoud	• Cite 1000 logts bt E 6 n 521 ihadaden	• BEJAIA	• 034 12 72 59	• True
	• GADOUCHE Abdelmadjid	• Cite CNEP les 100 logts BT B5 n 84	• BEJAIA	• 034.12.55.00	• True
	• HAMID Saïd	• 600 logts ihaddaden Bt B 27 n 532 cite	• BEJAIA	• 034.21.75.48	• True
Medecin-generaliste	• AIT-HAMOUDA	• Tour Nabila DJAHNINE app n 1 rue de la liberte	• BEJAIA	• 034.12.70.81	• True
	• BENBELKACEM Saïd	• 60 Bvd KRIM Belkacem 1er etage ihadaden	• BEJAIA	• 034.12.71.04	• True
	• CHAFA Hassane	• Rue HASSIBA BEN BOUALI route de l'universite	• BEJAIA	• 034.21.29.19	• True
	• HAMDJ Athmane	• Lot Djama n 1 Targa-Ouzemour	• BEJAIA	• 034.21.43.09	• True
	• SMADHI Djedjiga	• 17 rue DIDOUCHE Mourad cite smina	• BEJAIA	• 034.21.03.91	• True
	• DEBIANE Abdellah	• 17 rue DIDOUCHE Mourad cite smina	• BEJAIA	• 034.10 03 19	• True
	• ZERROUGUI Djamel	• Cite Aouchiche route de Boukhiana	• BEJAIA	• /	• True
	• MEKHFY Makhrouf	• 08 rue Ouyougout El Khemis • Tazeboudit route de Boukhiana promotion imm. Aouchiche	• BEJAIA	• / • 034.21.69.94	• True • True

Figure 5.1 – Dataset medecins

Notre deuxième jeu de données comprend environ 15 spécialités. Nous avons créé un ensemble de spécialités ainsi qu'un ensemble de symptômes clés associés à chaque spécialité. Les informations ont été collectées à partir de diverses sources, y compris Google.

Le 2 eme jeu de données est organisé en une table **Figure 5.2** avec les suivantes colonnes :

1. **Spécialité** : la spécialité du médecin (par exemple : pédiatre).
2. **Symptôme** : symptôme clé associé à chaque spécialité.
3. **Score** : représente le poids attribué à chaque symptôme clé.

En utilisant ces jeux de données, nous avons pu développer un système de recommandation précis et efficace pour aider les patients à trouver le médecin spécialiste adapté à leurs besoins de santé.

specialites	symptomes	score
ORL	• Perte-d'audition	• 0.8
	• Ecoulement-nasal-persistant	• 0.6
	• Maux-de-gorge	• 0.7
	• Vertiges	• 0.5
	• Otite	• 0.4
	• Douleur-auriculaire	• 0.3
Chirurgie-urologique	• Douleur-lors-de-la-miction	• 0.9
	• Sang-dans-l'urine	• 0.7
	• Difficulte-a-uriner	• 0.8
	• Douleur-dans-la-region-pelvienne	• 0.6
	• Incontinence-urinaire	• 0.5
	• Gonflement-testiculaire	• 0.4
Rhumatologie	• Douleurs-articulaires	• 0.7
	• Fatigue	• 0.5
	• Gonflement-articulaire	• 0.6
	• Douleurs-musculaires	• 0.4
	• Raideur-matinal	• 0.3
	• Limitation-mouvement	• 0.2
Medecin-generaliste	• Mal-de-gorge	• 0.6
	• Enrouement	• 0.4
	• Toux	• 0.5
	• Gonflement	• 0.3
	• Diarrhees	• 0.2
	• Maux-de-tete	• 0.4
	• Fievre	• 0.3
	• Maux-d'estomac	• 0.2

Figure 5.2 – Dataset specialite

5.3 Environnement de développement

5.3.1 Environnement matériel

DELL

Processeur : Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80 GHz

Mémoire installée(RAM) :8,00 Go.

Type de Système Système d'exploitation : 64 bits, processeur x64.

Windows 10 Professionnel N.

macOS Big Sur

MacBook Pro (Retina, 15 pouces, fin 2013)

Processeur : 2GHZ Intel Core i7 quatre cœurs

Mémoire 8GO 1600 MHz DDR3

Graphisme Intel Iris Pro 1536 Mo

Numéro de série C02LV6WXFD56

5.3.2 Environnement logiciel

Pour mener à bien notre projet informatique, nous avons d'abord dû réfléchir à l'outil logiciel le plus approprié à utiliser. En effet, il est important de choisir un langage de programmation qui convient le mieux à la tâche à accomplir plutôt que de considérer qu'un langage est plus puissant qu'un autre.

Nous avons donc opté pour l'implémentation de notre application web en utilisant le langage de programmation Python et son framework Django. Et pour la partie Front-end en React .js. Cette décision a été prise après avoir pris en compte plusieurs facteurs tels que la compatibilité avec les bibliothèques tierces nécessaires, la rapidité de développement et la facilité de maintenance du code. En ce qui concerne la configuration de notre base de données, nous avons utilisé tiny db.

React

React [51] est une bibliothèque JavaScript open-source développée par Facebook. Elle est largement utilisée pour la création d'interfaces utilisateur interactives et réactives dans le développement web. React utilise une approche basée sur des composants, ce qui signifie que vous pouvez diviser votre application en petits morceaux réutilisables appelés composants. Ces composants gèrent leur propre état et peuvent être mis à jour de manière efficace lors de modifications spécifiques, ce qui permet d'obtenir une interface utilisateur rapide et fluide. React utilise également la notion de virtual DOM (Document Object Model) pour optimiser les performances en mettant à jour uniquement les parties de l'interface utilisateur qui ont changé, au lieu de recharger toute la page. Cela permet de créer des applications réactives et performantes.

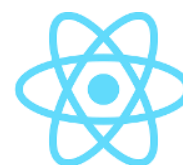


Figure 5.3 – React

Axios

Axios [52] est une bibliothèque JavaScript populaire utilisée pour effectuer des requêtes HTTP à partir de navigateurs web et de Node.js. Elle offre une API simple et élégante qui prend en charge des fonctionnalités telles que les requêtes basées sur les promesses, les intercepteurs pour la gestion des requêtes et des réponses, la transformation automatique des données au format JSON



Figure 5.4 – Axios

CSS

CSS (Cascading Style Sheets) [53] est un langage informatique utilisé pour décrire la présentation et la mise en forme des pages web. Il est souvent utilisé conjointement avec HTML pour séparer la structure d'une page web de sa présentation. Avec CSS, on peut appliquer des styles tels que les couleurs, les polices de caractères, les marges, les bordures et les espaces entre les éléments HTML. Cela permet une plus grande flexibilité dans la conception des pages web et facilite la mise à jour de l'apparence du site en modifiant simplement les règles CSS plutôt que de modifier le code HTML de chaque page individuellement.



Figure 5.5 – CSS

TinyDB

TinyDB [54] est une base de données NoSQL légère et simple à utiliser, écrite en Python et qui stocke les données au format JSON. Elle est idéale pour les applications légères et les petits projets, offrant une syntaxe similaire à celle d'un dictionnaire Python pour stocker et récupérer les données. TinyDB propose des fonctionnalités de recherche, de filtrage et de manipulation des données, tout en restant facilement intégrable dans des applications Python. Elle convient particulièrement aux besoins de stockage de petites quantités de données dans des scripts, des applications de bureau ou des projets personnels.



Figure 5.6 – TinyDB

VS Code

VS Code [55] est un éditeur de code source développé par Microsoft. C'est un logiciel que les programmeurs utilisent pour écrire, modifier et organiser leur code. Il offre des fonctionnalités pratiques telles que la coloration syntaxique, la complétion automatique du code, le débogage et l'intégration avec des outils de contrôle de version comme Git. VS Code est apprécié pour sa facilité d'utilisation, sa flexibilité et sa capacité à être personnalisé avec des extensions pour répondre aux besoins spécifiques des développeurs.



Figure 5.7 – VS code

Django

Django [56] est un framework Python très populaire dans le monde du développement web. Il est utilisé par de grandes entreprises telles qu'Instagram, Pinterest et la NASA pour créer des applications web performantes et efficaces. La philosophie de Django est basée sur le modèle MVT (Model-View-Template).

La Template est une couche de présentation qui gère entièrement la partie interface utilisateur. Les vues sont utilisées pour exécuter la logique métier et interagir avec un modèle pour transporter des données et restituer un résultat. Bien que Django suive le modèle MVC, il conserve ses propres conventions intactes. Par conséquent, le contrôle est géré par le framework lui-même.

La **figure 5.8** illustre l'interaction entre les différents composants de l'architecture MVT de Django pour répondre aux demandes des utilisateurs. Il convient de noter que le contrôleur ne correspond pas au contrôleur du MVC, mais plutôt à Django lui-même, qui gère toutes les opérations internes liées au choix de la vue à laquelle envoyer la requête HTTP et autres opérations connexes.

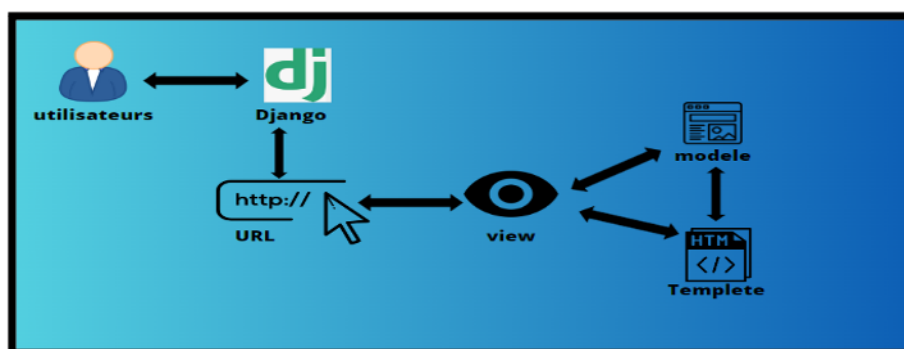


Figure 5.8 – Flux de contrôle Django

De plus, Django offre de nombreux avantages **la figure 5.9** résume les différents avantages :

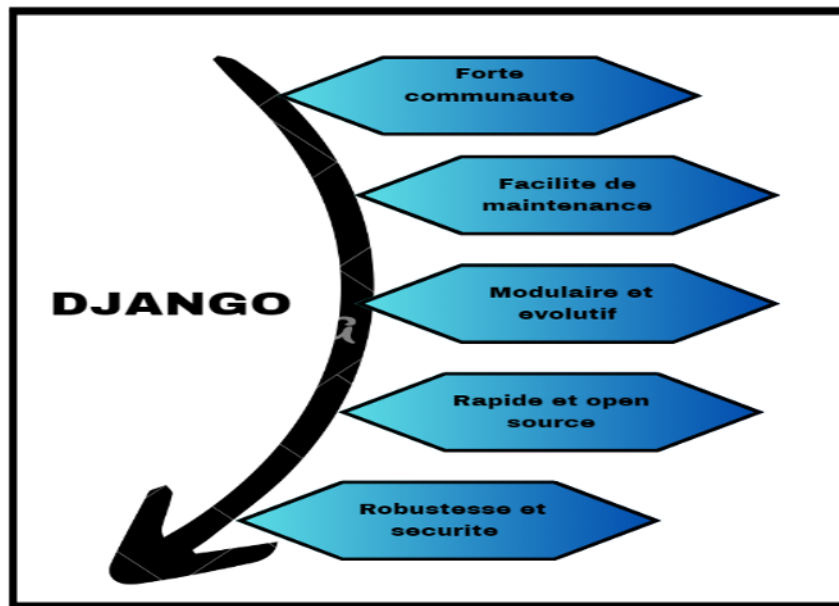


Figure 5.9 – Avantages de Django

Python

Python est un langage de programmation populaire, créé par Guido van Rossum en 1991. Il est open source, de haut niveau et à usage général. Python prend en charge les paradigmes de développement orienté objet, impératif, fonctionnel et procédural, et est particulièrement utilisé dans le domaine de la programmation d'apprentissage automatique. Il a été l'un des premiers langages à offrir des bibliothèques et des outils pour l'apprentissage automatique.

Parmi les bibliothèques les plus utilisées dans notre implémentation, on trouve :

- **Scikit-learn** : qui permet de gérer les algorithmes de machine learning de base tels que le clustering, les régressions linéaires et logistiques, la régression, la classification, et bien d'autres encore.
- **NumPy** : est une bibliothèque Python qui fournit des structures de données et des fonctions pour le calcul scientifique et numérique. Elle offre des opérations mathématiques et statistiques avancées. NumPy est largement utilisé dans des domaines tels que l'apprentissage automatique, la science des données, pour ses performances élevées et sa facilité d'utilisation.
- **Matplotlib** : est une bibliothèque Python populaire pour la création de graphiques et de visualisations de données. Avec Matplotlib, les utilisateurs peuvent personnaliser tous les aspects de leurs graphiques, tels que les axes, les étiquettes, les légendes, les couleurs et les styles de ligne. Cette flexibilité en fait un outil polyvalent pour représenter visuellement

des données dans divers domaines tels que la science des données, la recherche et la visualisation de données.

5.4 implémentation

Dans cette section, nous allons présenter le schéma de navigation de notre application ainsi que ses différentes interfaces, tout en détaillant les fonctionnalités offertes par chacune d'entre elles. Pour le processus de recommandation,

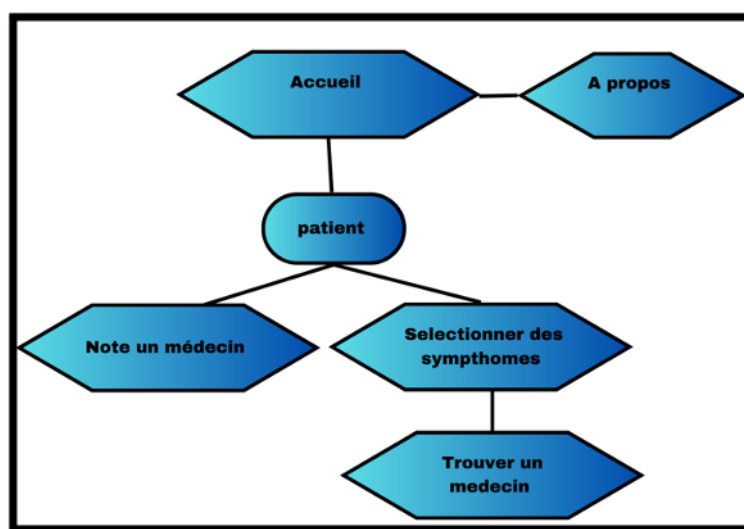


Figure 5.10 – Schéma de navigateur

dans son ensemble, est modélisé dans un diagramme de séquence exprimé dans la norme UML (Unified Modeling Language) représenté dans la figure 26.

Lorsque l'utilisateur accède au formulaire, il le remplit avec toutes ses informations personnelles. Une fois le formulaire rempli, l'utilisateur le valide. Le système reçoit ensuite le formulaire validé et demande à l'utilisateur de sélectionner ses symptômes parmi les options proposées. L'utilisateur choisit les symptômes pertinents. Ensuite, le système vérifie si les symptômes sélectionnés correspondent à des symptômes clés d'une spécialité médicale. Si c'est le cas, le système utilise le moteur de recommandation pour obtenir les N meilleures recommandations spécifiques à cette spécialité. Les recommandations spécifiques sont ensuite retournées à l'utilisateur via l'interface utilisateur. En revanche, si les symptômes ne correspondent à aucune spécialité spécifique, le moteur de recommandation utilise la base de données pour obtenir les N meilleures recommandations générales. Ces recommandations générales sont ensuite transmises à l'utilisateur via l'interface utilisateur.

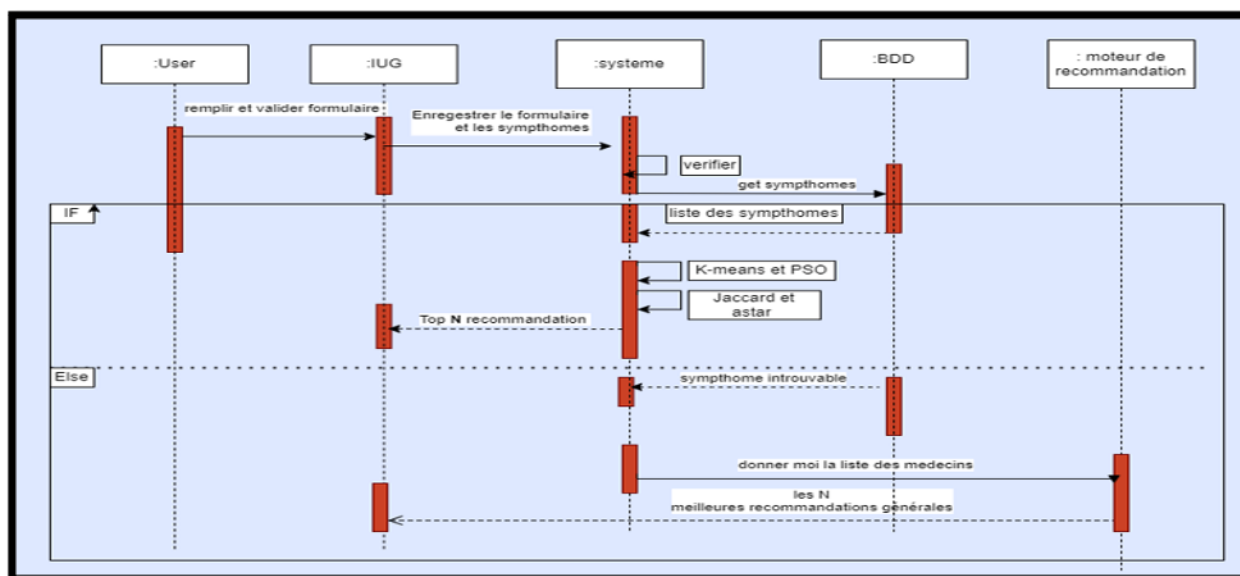


Figure 5.11 – Diagramme de sequence

IUG : Interface Utilisateur Graphique

5.4.1 Page d'accueil

La page d'accueil joue un rôle essentiel en expliquant l'intérêt de notre application DocFast et en guidant les visiteurs malades sur la manière de l'utiliser efficacement. En bas de la page, à gauche, se trouve un court texte explicatif mettant en évidence les avantages et les fonctionnalités clés de notre application. À droite, un formulaire stratégiquement placé permet de collecter les informations nécessaires sur l'utilisateur. Le formulaire commence par la saisie du prénom et du nom du patient, suivi de son adresse e-mail. Cela nous permet de lui envoyer un courrier de satisfaction pour enrichir notre système de recommandation. De plus, la saisie de l'adresse actuelle du patient nous permet de proposer le médecin le plus proche de sa localisation, garantissant ainsi un service de soins de santé pratique et personnalisé.

Une fois que vous avez rempli ces informations nécessaires, vous pouvez passer à l'étape suivante en cliquant sur le bouton Sélectionner les symptômes. Cette fonctionnalité permet de progresser efficacement dans le processus de recherche en vous permettant de spécifier les symptômes que Vous rencontrez.

DocFast
Trouver votre médecin plus vite

DocFast

Comment notre site vous aidera ?

Pour vous aider à trouver rapidement des professionnels de santé qualifiés et compétents, notre site de recommandation de médecins utilise des algorithmes sophistiqués pour trier et recommander les médecins les mieux adaptés à votre situation. En entrant vos symptômes et votre emplacement, notre site recherche des médecins dans votre région qui ont de l'expérience dans le traitement de votre maladie ou condition spécifique.

Notre site utilise également une technologie de géolocalisation pour vous aider à trouver des médecins proches de chez vous. Vous pouvez également affiner votre recherche en fonction de la spécialité du médecin, de l'assurance maladie acceptée et de la disponibilité pour prendre un rendez-vous. Tout cela rend le processus de recherche de médecin plus facile et plus efficace pour les patients.

Notre site de recommandation de médecins est donc un outil précieux pour aider les patients à trouver rapidement des professionnels de santé qualifiés et compétents, tout en leur offrant un moyen facile de filtrer les résultats en fonction de leurs besoins spécifiques.

Veuillez remplir le formulaire suivant:

Nom: Prénom:

Date de naissance :

Mail :

Veuillez entrer votre localisation actuelle

N Rue: Rue:

Ville: Adresse postale:

Equipe :
Dyha TAGZRT
Faycel BOURDJOUA

Nos Adresses
Makouda, Tip Ouzou, Algere
Arizour ,Bejaia, Algere

Nos Telephones
+213 6 98 18 38 40
+213 5 82 05 39 49

Nos mails
dyhastagzrt2022@gmail.com
faycelbourdjoua@gmail.com

© 2023 Dyha Tagirt & Faycel Bourdjoua. Tous droits réservés.

Figure 5.12 – Accueil

5.4.2 Trouvez un médecin

Dans la page Trouvez un médecin, vous trouverez un ensemble de symptômes parmi lesquels l'utilisateur peut sélectionner ceux qui lui correspondent. Une fois les symptômes choisis, l'utilisateur peut cliquer sur le bouton **Trouvez un médecin** pour afficher la liste des médecins similaire.

De plus, si l'utilisateur souhaite modifier les informations saisies dans le formulaire, il peut simplement cliquer sur le bouton **Modifier le formulaire**. Cela le ramènera à la page d'accueil où il pourra mettre à jour les informations personnelles.

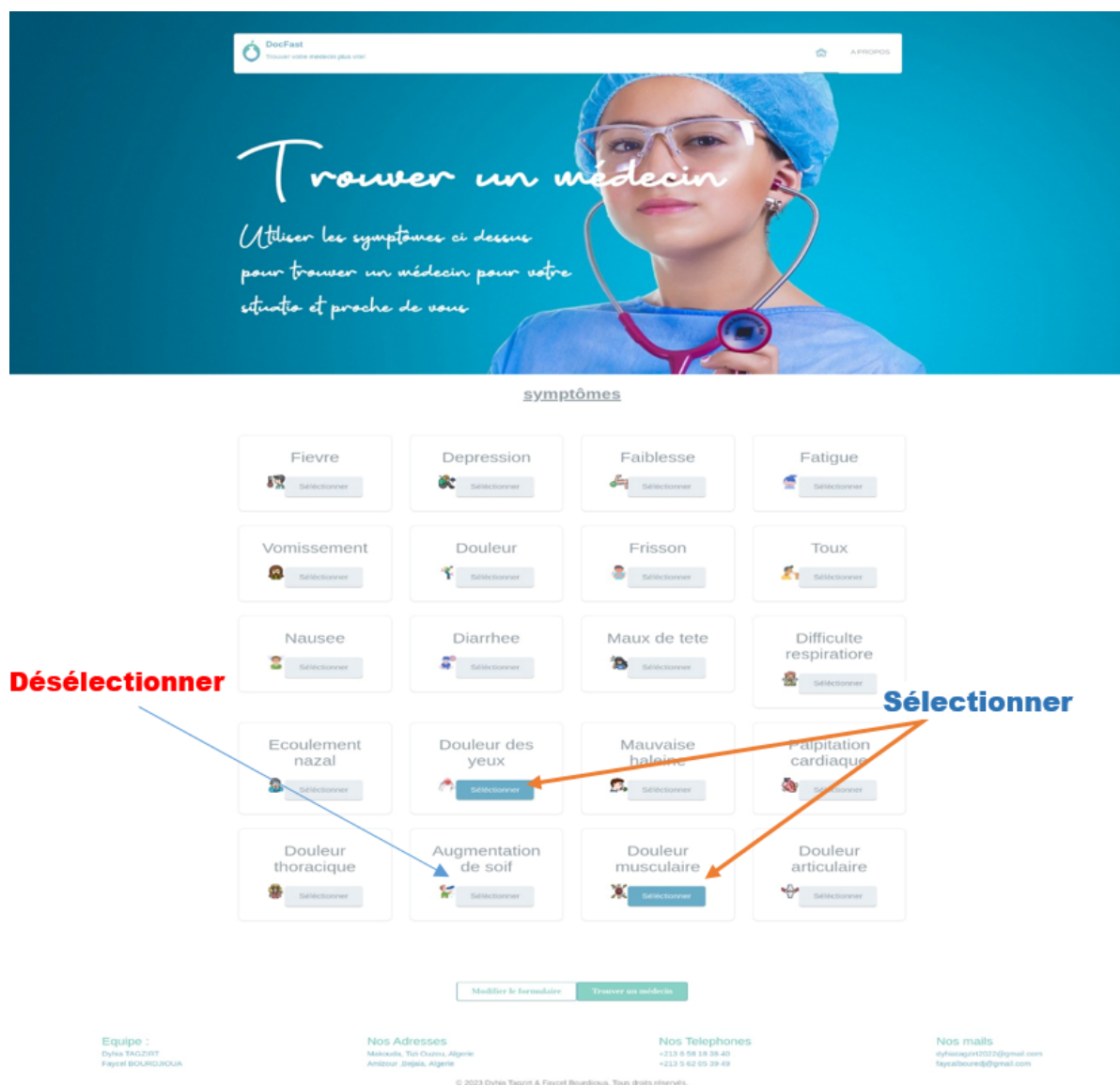


Figure 5.13 – Trouver un médecin

5.4.3 Choisissez un médecin

À partir des informations saisies par l'utilisateur, notre algorithme recommande les médecins qui correspondent aux symptômes et à la localisation du patient. Ces médecins sont triés en fonction de leur disponibilité et dans l'ordre décroissant, du meilleur recommandé au moins recommandé. Ensuite, pour les médecins qui ont la même recommandation, Ils sont triés en fonction de leur proximité géographique par rapport au patient.

Lorsque le patient clique sur le bouton **Consulter**, il choisit ainsi un médecin et le rôle de notre application prend fin. Le but de ce bouton est d'envoyer un message de satisfaction au patient 24 heures plus tard, afin de lui permettre de

noter le médecin consulté. Cette évaluation contribue à l'amélioration continue de notre service et aide les autres utilisateurs à faire des choix éclairés lors de leur recherche de médecins.

Le bouton **Modifier les symptômes** nous permet de retourner à la page précédente, où l'utilisateur peut sélectionner ou modifier ses symptômes pour affiner la recherche de médecins.

DocFast
Trouver votre médecin plus vite

Choisissez un medecin
Choisissez un medecin qui vous convient

Les Medecins

ACHOUR Moubout	Chirurgie-urologique	Residence la Grande Boulevard de la Grande BEJAJA	034.21.23.95	☆☆☆☆☆	Indisponible	2368 Km	Consulter
HOCINE Fares	Chirurgie-urologique	Bt n° 05 Cité DJAMA Abdelkader RDC Madela BEJAJA	/	☆☆☆☆☆	Indisponible	3800440.8 Km	Consulter
BOUCHEFA MONTAZ	Chirurgie-urologique	Cité des HAMMADITES résidence 111 SIDI- AHMED BEJAJA	034.21.96.74	☆☆☆☆☆	Disponible	228922.1 Km	Consulter
DAHMANI Karim Omar Khetou	Chirurgie-urologique	71, cité zeyna nouvelle BEJAJA	034.21.44.92	☆☆☆☆☆	Disponible	2199933.5 Km	Consulter

[Modifier les symptômes](#)

Equipe :
Dyha TAGDIRT
Fayçal BOURDJOUA

Nos Adresses
Makouda, 102 Ouzou, Algérie
Ameour ,Beyas, Algérie

Nos Telephones
+213 5 58 10 38 40
+213 5 62 05 38 49

Nos mails
dyha.tagdir1202@gmail.com
faycalbourd@gmail.com

© 2023 Dyha Tagdir & Fayçal Bourdjoua. Tous droits réservés.

Figure 5.14 – Liste des médecins similaire

5.4.4 Notez votre médecin

Après 24 heures, l'application envoie automatiquement un lien à l'utilisateur. En cliquant sur ce lien, une page intitulée « Noter votre médecin » s'ouvre, lui demandant de donner son avis sur le médecin consulté.

Le bouton **Retourner vers l'accueil** permet de revenir à la page d'accueil de l'application.

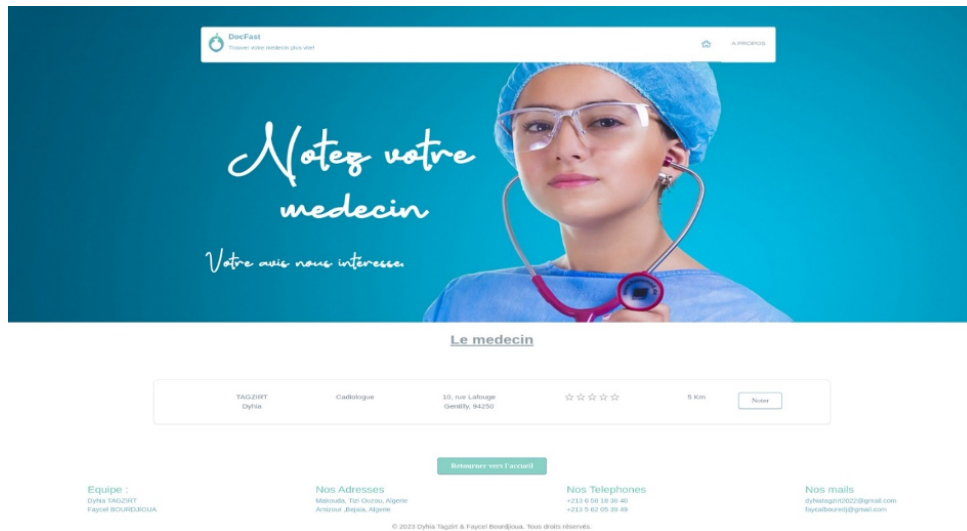


Figure 5.15 – Noter un médecin

5.4.5 A propos

La page ”À propos” représente la présentation de notre équipe de développement et de ses services, ainsi que le but de notre application. Elle est soigneusement conçue pour offrir aux visiteurs une vue d’ensemble claire et concise de ce que nous faisons et de ce que nous proposons. (figure 5.16)

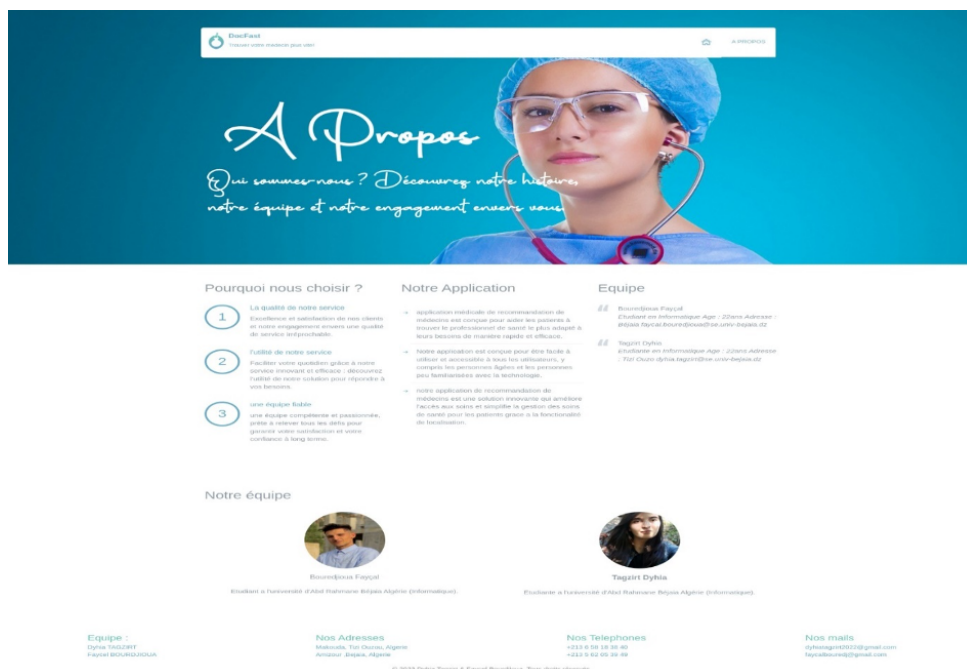


Figure 5.16 – A propos

5.5 Evaluations

5.5.1 Evaluation des algorithmes (K-means, PSO)

La figure 5.17 suivante représente le regroupement des médecins après l'application de l'algorithme de K-means sur le jeu de données des médecins ainsi que celui des spécialités, ce regroupement permet d'identifier des similarités et des relations entre les médecins en fonction de leurs spécialités

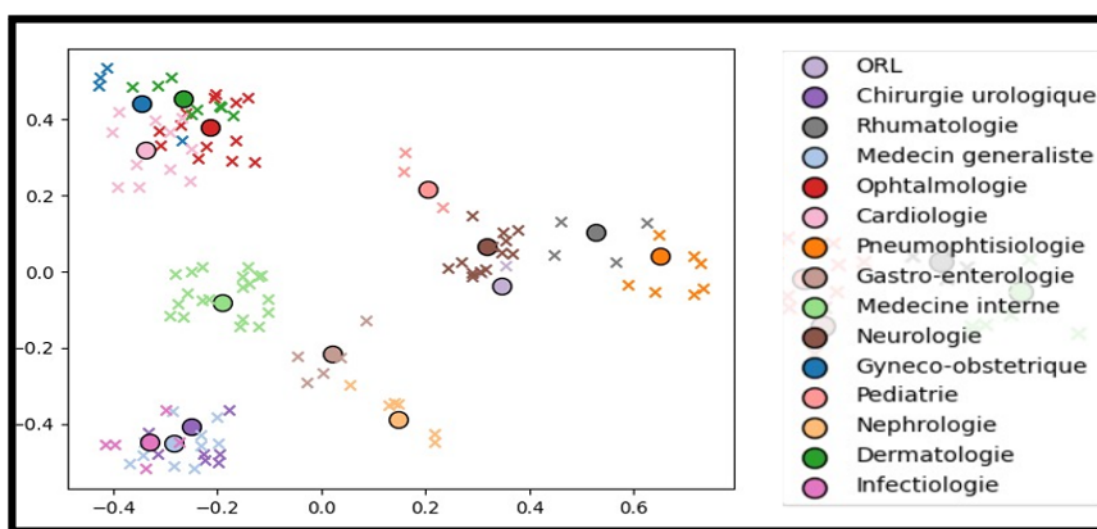


Figure 5.17 – Regroupement des médecins selon leur spécialités

Pour recommander un médecin à un utilisateur, nous supposons que les symptômes de l'utilisateur sont les suivants :

- Diarrhees
- Fatigue
- Difficulte-a-uriner
- Douleur-lors-de-la-miction
- Sang-dans-l'urine

Après l'application de la similarité de Jaccard, le système nous a recommandé d'ajouter le l'utilisateur au cluster de la spécialité **Chirurgie urologique**. La figure 5.18 illustre cette recommandation.

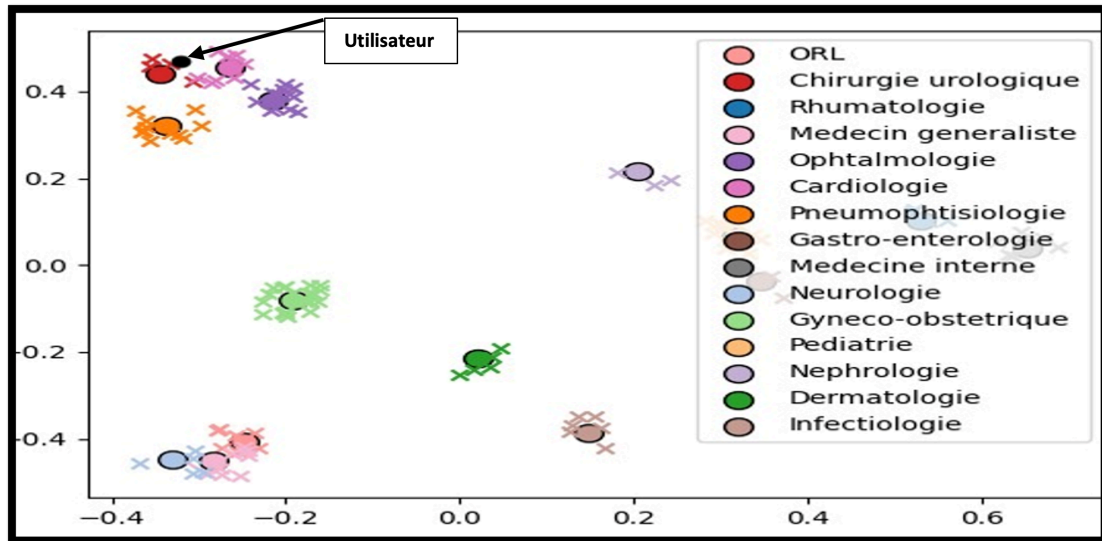


Figure 5.18 – Regroupement des médecins selon leur spécialités

Dans notre système, nous avons appliqué deux algorithmes différents. Dans ce qui suit, nous allons discuter des résultats obtenus pour la similarité de Jaccard en utilisant le K-means seul et en utilisant le K-means avec PSO.

Dans le premier, nous avons utilisé l’algorithme de clustering pour regrouper les médecins en différentes spécialités. Ensuite, nous avons calculé la similarité de Jaccard entre les symptômes du patient et les symptômes clés des spécialités, et nous avons obtenu un score de similarité de 27% indique le degré de correspondance entre les symptômes du patient et les symptômes clés de la spécialité **la figure 5.19** montre les résultats obtenue :

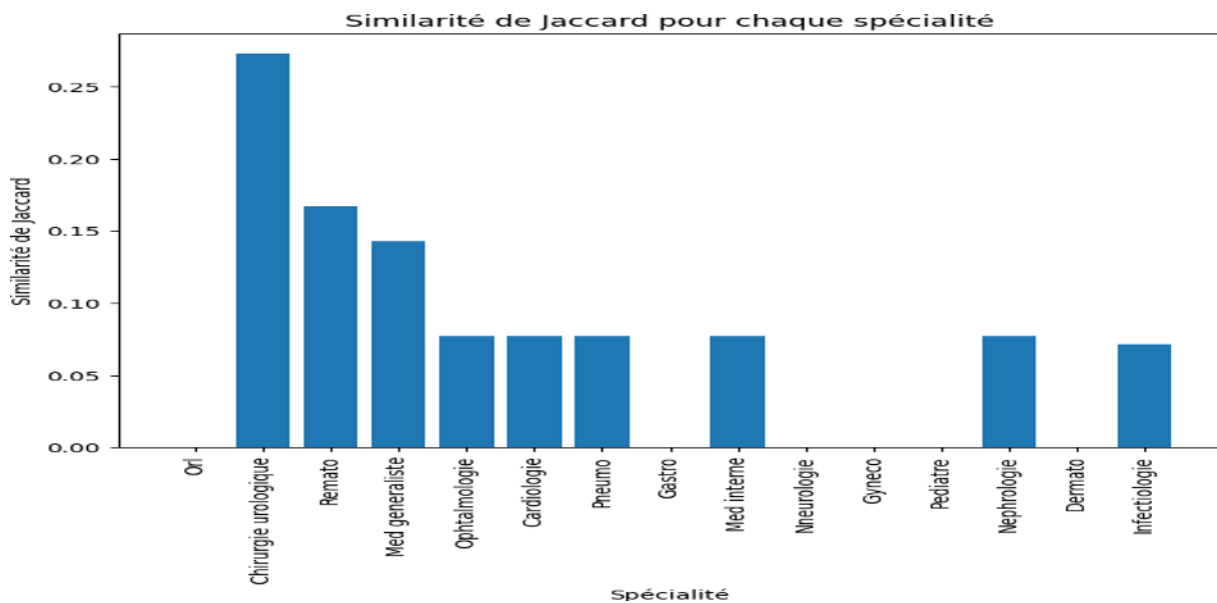


Figure 5.19 – Similarité de Jaccard pour chaque médecin en utilisant l’algorithme K-means

Pour le deuxième algorithme, nous avons intégré le PSO au K-means afin de

sélectionner les symptômes les plus pertinents dans la spécialité donnée. Et le rôle du PSO dans ce cas était de guider la recherche des meilleurs symptômes à considérer pour la recommandation. En ajustant les positions et les vitesses des particules dans l'espace de recherche, le PSO a permis de trouver une combinaison optimale de symptômes qui maximise la similarité entre les symptômes du patient et les critères des médecins spécialistes.

Ensuite, nous avons appliqué la similarité de Jaccard entre ces symptômes sélectionnés et les symptômes du patient, et nous avons obtenu un score de similarité de 37,5% indique le degré de similarité entre ces symptômes sélectionnés et les symptômes du patient. Dans notre système, les résultats obtenus sont affichés dans **la figure 5.20** :

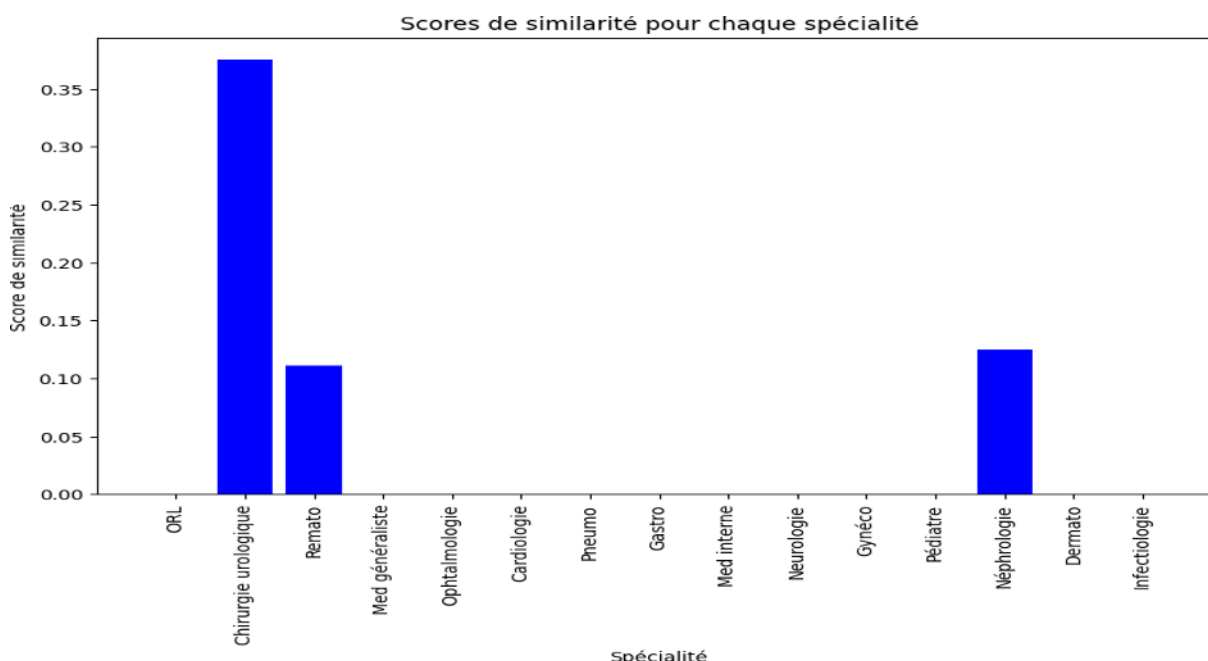


Figure 5.20 – Similarité de Jaccard pour chaque médecin en utilisant l'algorithme K-means et PSO

L'utilisation du K-means avec PSO a permis d'améliorer la pertinence de la recommandation en identifiant les symptômes les plus importants pour chaque spécialité, ce qui a conduit à un score de similarité de Jaccard plus élevé par rapport à l'utilisation du K-means seul.

Le PSO a contribué à optimiser la sélection des symptômes en recherchant les combinaisons les plus pertinentes, ce qui a renforcé la correspondance entre les symptômes du patient et les symptômes clés de la spécialité recommandée. Cette amélioration démontre l'efficacité de l'intégration du PSO dans le processus de recommandation médicale.

5.5.2 Évaluation des performances du modèle

Nous utilisons les paramètres de précision, de rappel et de mesure pour évaluer les résultats du modèle proposé.

La métrique de précision (équation A) donne un aperçu de l'exactitude avec laquelle un modèle peut prédire l'appartenance à une classe. Par conséquent, elle mesure la qualité des résultats de la classification. Elle est définie par le nombre de TPs divisé par la somme des TPs et des FPs.

$$Precision = \frac{TP}{TP + FP} \dots \dots \dots (A)$$

Cependant, la métrique du rappel (équation B) indique dans quelle mesure le modèle peut prédire l'appartenance à une classe et couvre donc l'aspect quantitatif du succès de la classification. Elle est définie par la fraction d'éléments qui sont recommandés parmi tous les éléments pertinents. Également appelée sensibilité.

$$Rappel = \frac{TP}{TP + FN} \dots \dots \dots (B)$$

Où :

- TP : nombre d'aliments pertinents recommandés.
- FN : nombre d'aliments pertinents non recommandés.
- FP : nombre d'aliments non pertinents recommandés.

La mesure F est l'une des mesures les plus populaires utilisées pour mesurer la précision des clusters, elle est basée sur la précision et la valeur de rappel (équation C) et peut être calculée par la formule suivante :

$$F - mesure = 2 \cdot \frac{Precision \cdot Rappel}{Precision + Rappel} \dots \dots \dots (C)$$

Nous effectuons une recommandation manuelle pour servir de point de com-

paraison dans le processus de recommandation. Nous sélectionnons les profils des utilisateurs et les spécialités des médecins.

La figure 5.21 montre les résultats de performance obtenus sur 38 patients en utilisant K-means, tandis que **La figure 5.22** montre les résultats de performance obtenus sur les mêmes 38 patients en utilisant K-means avec PSO.

Les valeurs d'évaluation de la précision et de la F-mesure sont identiques car les valeurs de vrais positifs, de faux positifs et de faux négatifs sont identiques pour le modèle donné.

D'après les résultats obtenus, nous avons constaté qu'en utilisant K-means en combinaison avec PSO, la performance en termes de précision, de rappel et de F-mesure s'est significativement améliorée. Avec K-means seul, nous avons obtenu un score de performance de 86,84%, tandis qu'en utilisant K-means avec PSO, nous avons obtenu un score de performance de 92,10%.

Ces résultats suggèrent que l'ajout de PSO dans le processus de K-means a permis d'améliorer de manière significative la précision, le rappel et la F-mesure du clustering par rapport à l'utilisation de K-means seul.

K-means			
patient	precision	recall	F_measure
1	1.0	1.0	1.0
2	1.0	1.0	1.0
3	1.0	1.0	1.0
4	1.0	1.0	1.0
5	1.0	1.0	1.0
6	1.0	1.0	1.0
7	1.0	1.0	1.0
8	1.0	1.0	1.0
9	1.0	1.0	1.0
10	1.0	1.0	1.0
11	1.0	1.0	1.0
12	1.0	1.0	1.0
13	0.0	0.0	0
14	1.0	1.0	1.0
15	1.0	1.0	1.0
16	0.0	0.0	0
17	1.0	1.0	1.0
18	1.0	1.0	1.0
19	1.0	1.0	1.0
20	0.0	0.0	0
21	1.0	1.0	1.0
22	1.0	1.0	1.0
23	1.0	1.0	1.0
24	1.0	1.0	1.0
25	1.0	1.0	1.0
26	1.0	1.0	1.0
1 / 3			
resultat_test.md			
patient	precision	recall	F_measure
27	1.0	1.0	1.0
28	1.0	1.0	1.0
29	1.0	1.0	1.0
30	1.0	1.0	1.0
31	0.0	0.0	0
32	1.0	1.0	1.0
33	1.0	1.0	1.0
34	1.0	1.0	1.0
35	0.0	0.0	0
36	1.0	1.0	1.0
37	1.0	1.0	1.0
38	1.0	1.0	1.0
moyenne %	86.84210526315789	86.84210526315789	86.84210526315789

Figure 5.21 – Résultat avec k-means

K-means et PSO			
patient	precision	recall	F_measure
1	1.0	1.0	1.0
2	1.0	1.0	1.0
3	1.0	1.0	1.0
4	1.0	1.0	1.0
5	1.0	1.0	1.0
6	1.0	1.0	1.0
7	1.0	1.0	1.0
8	1.0	1.0	1.0
9	1.0	1.0	1.0
10	1.0	1.0	1.0
11	1.0	1.0	1.0
12	1.0	1.0	1.0
13	1.0	1.0	1.0
14	1.0	1.0	1.0
2 / 3			
resultat_test.md			
patient	precision	recall	F_measure
15	1.0	1.0	1.0
16	1.0	1.0	1.0
17	1.0	1.0	1.0
18	0.0	0.0	0
19	1.0	1.0	1.0
20	0.0	0.0	0
21	1.0	1.0	1.0
22	1.0	1.0	1.0
23	1.0	1.0	1.0
24	1.0	1.0	1.0
25	1.0	1.0	1.0
26	1.0	1.0	1.0
27	1.0	1.0	1.0
28	1.0	1.0	1.0
29	1.0	1.0	1.0
30	1.0	1.0	1.0
31	0.0	0.0	0
32	1.0	1.0	1.0
33	1.0	1.0	1.0
34	1.0	1.0	1.0
35	1.0	1.0	1.0
36	1.0	1.0	1.0
37	1.0	1.0	1.0
38	1.0	1.0	1.0
moyenne %	92.10526315789474	92.10526315789474	92.10526315789474

Figure 5.22 – Résultat avec k-means et PSO

5.6 Conclusion

Dans ce chapitre, nous avons évoqué les principales étapes de notre implémentation. En premier lieu, nous avons défini les jeux de données en général, et présenté les datasets que nous avons utilisés pour former notre système, illustrés des captures d'écran de quelques affichages obtenus. Ensuite, nous avons donné l'environnement et les outils utilisés pour construire un système solide permettant de recommander des médecins aux utilisateurs. Enfin, nous avons expliqué l'utilité et l'objectif des métriques utilisées pour les calculs et l'évaluation des performances de notre système.

Dans le dernier chapitre, nous clôturons notre mémoire avec une conclusion générale comportant une idée générale du plan suivi lors de chaque chapitre, en discutant vers la fin sur nos travaux futurs et leurs objectifs.

Chapitre 6

Conclusion générale et perspectives

6.1 Introduction

Ce mémoire explore les SRs et examine les différentes approches qu'ils offrent, ainsi que leur pertinence dans le monde moderne. Ce domaine est vaste et interdisciplinaire, englobant des domaines tels que le filtrage, le traitement des données, la recherche, et bien d'autres.

Ce mémoire se concentre sur l'introduction d'une approche de recommandation novatrice spécifiquement conçue pour le domaine de la santé dans les villes intelligentes. L'objectif principal est de présenter une solution logicielle qui facilite la recherche rapide des meilleurs professionnels de santé disponibles dans la région, tout en réduisant le stress et l'anxiété des patients lorsqu'ils sont malades dans une ville étrangère.

Dans le chapitre 2, nous avons examiné trois approches distinctes de filtrage : les systèmes de filtrage collaboratif (FC), les systèmes de filtrage basés sur le contenu (FBC) et les systèmes hybrides. L'analyse des travaux sur les SR a révélé plusieurs limitations, telles que le démarrage à froid, la rareté des données et le manque d'informations de la part des utilisateurs, qui sont des défis rencontrés lors de leurs premières utilisations.

Ce travail s'est concentré sur la création d'un SR ayant le moins de limitations possible lors de son utilisation. Pour cela, nous avons utilisé l'approche du FC et ses deux sous-approches basées sur l'utilisateur et la connaissance. Nous avons personnalisé ces approches pour les rendre plus efficaces dans notre contexte spécifique.

Dans le chapitre 3, nous avons effectué une revue approfondie de l'état de

l'art sur les SRs et les recherches menées dans le domaine de la santé des utilisateurs. Cette analyse nous a permis de mieux comprendre la problématique et de restreindre notre champ de recherche. Nous avons accordé une attention particulière à l'objectif du SR, qui vise à aider les touristes malades se trouvant dans une ville étrangère à trouver un médecin.

En nous basant sur les principales recherches étudiées et traitées, nous avons conçu et développé notre application « FastDoc » dédiée à la recommandation de médecins pour les touristes malades en fonction de leurs symptômes, de leur antécédent médical, de leur âge, localisation géographique... En y apportant notre expertise personnelle, nous avons créé un système de recommandation qui répond précisément aux besoins des utilisateurs.

Dans le chapitre 4, nous avons présenté en détail notre approche novatrice visant à développer un SR basé sur le FC par les utilisateurs et le filtrage basé sur la connaissance. Notre approche utilise des méthodes avancées de machine learning ainsi que les algorithmes de K-means et de PSO pour améliorer la précision et la pertinence des recommandations fournies par le SR.

Notre SR est spécifiquement conçu pour répondre aux besoins du domaine de la santé, en fournissant une assistance éclairée aux touristes confrontés à des problèmes de santé pendant leur séjour. L'objectif principal de notre système est d'aider ces touristes à choisir un médecin de manière plus informée et judicieuse.

Pour atteindre cet objectif, nous avons utilisé des données provenant de différentes sources, telles que les profils des utilisateurs, les évaluations des médecins, les spécialités médicales et les symptômes des patients.

En utilisant l'algorithme de K-means, nous avons regroupé les médecins en clusters en fonction de ces caractéristiques, ce qui nous a permis de fournir des recommandations plus personnalisées et ciblées. De plus, nous avons utilisé l'algorithme de PSO pour optimiser les paramètres du modèle de recommandation, améliorant ainsi la précision des recommandations et la satisfaction des utilisateurs.

En appliquant la similarité de Jaccard, nous avons pu identifier les similarités entre les symptômes des utilisateurs et les symptômes clés de chaque spécialité médicale, afin de recommander des médecins appropriés.

En complément, nous avons utilisé le filtrage basé sur la connaissance pour prendre en compte le problème de démarrage à froid. Lorsque le système de recommandation ne dispose pas de suffisamment de données, il peut fournir des recommandations générales pour l'utilisateur.

Nous avons fourni une explication exhaustive de notre proposition, laquelle a été publiée dans le cadre d'une conférence prestigieuse MoMLLeTamp ;DS-2023. Cette publication détaille de manière précise et méthodique les différentes étapes que nous avons suivies pour développer notre SR, mettant en évidence l'approche novatrice et efficace que nous avons adoptée en combinant le filtrage collaboratif par les utilisateurs avec le filtrage basé sur la connaissance.

Dans le chapitre 5, nous abordons les éléments clés de la mise en œuvre de notre approche. Nous discutons du choix du langage de programmation, des jeux de données utilisés et des différents aspects impliqués dans l'implémentation de notre système.

La mise en œuvre de notre système au sein d'un système de gestion des informations de santé personnelles, ainsi que son évaluation, nous a permis de mesurer ses performances dans le domaine des recommandations de médecins. Les résultats expérimentaux ont confirmé l'efficacité de notre approche proposée.

6.2 Perspectives et travaux futurs

Les recherches futures ouvriront de nouvelles possibilités pour une meilleure catégorisation des médecins en fonction de leurs spécialités en utilisant les symptômes de l'utilisateur, ce qui permettra de générer des recommandations plus précises.

Cependant, malgré nos efforts, nous n'avons pas pu traiter la récupération des informations implicites de l'utilisateur pour la création de son profil à l'aide des dispositifs IdO en raison d'un manque de matériel adéquat (exemple : détecteur de température, diabète ...). Par conséquent, notre première perspective pour les travaux futurs consiste à compléter notre implémentation afin de pouvoir récupérer ces informations implicitement.

Deuxièmement, nous avons l'intention de permettre aux médecins de créer leur propre profil et aux clients d'effectuer des recherches sur les médecins en fonction de leur nom ou de leurs spécialités. Nous proposons également d'intégrer une fonctionnalité de chat entre l'utilisateur et le spécialiste, permettant une communication directe et instantanée pour faciliter les échanges et les consultations.

Troisièmement, dans le cadre de nos travaux futurs, nous souhaitons ajouter une composante intelligente capable de proposer des soins primaires aux patients. Notre objectif est de fournir aux patients des informations précises sur les mesures à prendre pour prendre soin de leur santé et prévenir les maladies.

Enfin, notre dernière perspective de recherche future est d'explorer l'utilisation des algorithmes d'apprentissage profond pour construire un système de recommandation plus performant et plus efficace. En exploitant les capacités de l'apprentissage profond, nous pourrions analyser de manière plus approfondie les données des utilisateurs et les corrélations entre les symptômes, les antécédents médicaux et les spécialités médicales, ce qui améliorera la précision et la pertinence des recommandations. Les algorithmes d'apprentissage profond peuvent aider à découvrir des relations complexes et subtiles entre les données, ce qui peut conduire à des recommandations plus précises et personnalisées.

References

- [1] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, “Recommender systems survey,” *Knowledge-Based Systems*, vol. 46, pp. 109–132, 2013.
- [2] L. Quijano-Sánchez, I. Cantador, M. E. Cortés-Cediel, *et al.*, “Recommender systems for smart cities,” *Information Systems*, 2020.
- [3] R. Burke, “Hybrid recommender systems : survey and experiments,” *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331–370, 2002.
- [4] F. Ricci, L. Rokach, and B. Shapira, *Introduction to recommender systems handbook*. Springer, 2011.
- [5] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems : A survey of the state of the art and possible extensions,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.
- [6] P. Resnick and H. R. Varian, “Recommender systems,” *Communication of the ACM*, vol. 40, pp. 56–58, 1997.
- [7] P. Maes and U. Shardanand, “Social information filtering : algorithms for automating ’word of mouth’,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, (Denver, Colorado, United States), ACM Press/Addison-Wesley Publishing Co, 1995.
- [8] C. M. Bishop, *Pattern recognition and machine learning*. New York : Springer, 2006.
- [9] D. Goldberg, D. Nichols, B. Oki, and D. Terry, “Using collaborative filtering to weave an information tapestry,” *Communications of the ACM*, vol. 35, pp. 61–70, 1992.
- [10] C. Desrosiers and G. Karypis, “A comprehensive survey of neighborhood-based recommendation methods,” in *Recommender Systems Handbook*, Springer, 2011.

- [11] N. Rao and V. Talwar, "Application domain and functional classification of recommender systems : a survey," *Desidoc Journal of Library and Information Technology*, vol. 28, no. 3, pp. 17–36, 2008.
- [12] H. El Bouhissi, M. Adel, A. Ketam, *et al.*, "Towards an efficient knowledge-based recommendation system," in *CEUR Workshop Proceedings*, vol. 2853, pp. 38–49, 2021.
- [13] B. Towle and C. Quinn, "Knowledge based recommender systems using explicit user models," *Knowledge Creation Diffusion Utilization*, pp. 74–77, 2000.
- [14] Bouchindhomme and Rochlitz, "Dans le flou artistique. Éléments d'une théorie de la 'rationalité esthétique'," *Pages*, pp. 203–238, 2000.
- [15] H. M. Mohamed, Khafagy, and M. H. Ibrahim, "Recommender système challenges and solutions survey," in *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, IEEE, 2019.
- [16] H. El Bouhissi, *Encyclopedia of Data Science and Machine Learning*. 2023.
- [17] A. M. Rashid, I. Albert, D. Cosley, S. K. Lam, S. M. McNee, J. A. Konstan, and J. Riedl, "Beyond algorithms : An HCI perspective on recommender systems," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 9, no. 4, pp. 351–375, 2002.
- [18] L. Sharma and A. Gera, "A survey of recommendation system : Research challenges," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 4, no. 5, pp. 1989–1992, 2013.
- [19] H. Mahmoud, A. Hegazy, and M. H. Khafagy, "An approach for big data security based on hadoop distributed file system," in *2018 International Conference on Innovative Trends in Computer Engineering (ITCE)*, (Aswan), pp. 109–114, 2018.
- [20] V. Albino, U. Berardi, and R. M. Dangelico, "Smart cities : Definitions, dimensions, performance, and initiatives," *Journal of Urban Technology*, vol. 22, 2015.
- [21] M. Papagelis, D. Plexousakis, and T. Kutsuras, "Alleviating the sparsity problem of collaborative filtering using trust inferences," in *International Conference on Trust Management*, pp. 224–239, 2005.

- [22] S. Alawadhi, A. Aldama-Nalda, H. Chourabi, R. R. Gil-Garcia, S. Leung, S. Mellouli, T. Nam, T. A. Pardo, H. J. Scholl, and S. Walker, "Building understanding of smart city initiatives," in *International conference on electronic government*, pp. 40–53, Springer, 2012.
- [23] E. O'Dwyer, I. Y. Pan, S. Acha, and N. Shah, "Smart energy systems for sustainable smart cities : Current developments, trends and future directions," *Applied Energy*, vol. 237, pp. 581–597, 2019.
- [24] T. Bakıcı, E. Almirall, and J. Wareham, "A smart city initiative : The case of barcelona," *Journal of the Knowledge Economy*, vol. 2, no. 1, pp. 1–14, 2012.
- [25] R. Petrolo, V. Loscri, and N. Mitton, "Towards a smart city based on cloud of things, a survey on the smart city vision and paradigms," *Transactions on Emerging Telecommunications Technologies*, vol. 28, no. 1, p. e2931, 2017.
- [26] G. P. Hancke, B. d. Carvalho e Silva, and G. P. Hancke Jr., "The role of advanced sensing in smart cities," *Sensors*, vol. 13, no. 1, pp. 393–425, 2012.
- [27] S. P. Mohanty, U. Choppali, and E. Kougianos, "Everything you wanted to know about smart cities : The internet of things is the backbone," *IEEE Consumer Electronics Magazine*, vol. 5, no. 3, pp. 60–70, 2016.
- [28] A. M. Townsend, *Smart Cities : Big Data, Civic Hackers, and the Quest for a New Utopia*. 2013.
- [29] X. Zhang, G. Zhang, and X. Zhang, eds., *Smart Health : International Conference, ICSH 2013, Beijing, China, August 3-4, 2013. Proceedings*. 2013.
- [30] M. Weier and et al., "Smart cities and health : The potential for health promotion and disease prevention," 2015.
- [31] A. P. Rojas-Gualdrón and et al., "Smart cities and health promotion : A meta-analysis," 2019.
- [32] M. Khajah and et al., "Smart home-based healthcare technologies : A systematic review," 2019.
- [33] D. G. T. White and et al., "Smart governance for health and wellbeing : The evidence," 2019.

- [34] K. Keane and V. Nisi, "Experience prototyping," pp. 224–237, 2013.
- [35] E. Leanza and G. Carbonaro, "Attaining sustainable, smart investment : The smart city as a place-based capital allocation instrument," in *E-Planning and Collaboration : Concepts, Methodologies, Tools, and Applications*, pp. 179–204, IGI Global, 2018.
- [36] S. P. Erdeniz, "Recommender systems for iot enabled m-health applications," in *IFIP Advances in Information and Communication Technology*, 2018.
- [37] S. Forouzandeh, A. R. Aghdam, M. Barkhordari, S. A. Fahimi, M. K. Vayqan, S. Forouzandeh, and E. G. Khani, "Ijcsns international journal of computer science and network security," vol. 17, August 2017.
- [38] A. K. Sahoo, C. Pradhan, R. K. Barik, and H. Dubey, "Deepreco : Deep learning based health recommender system using collaborative filtering," *Computation*, 2019.
- [39] Jabeen, "An iot based efficient hybrid recommender system for cardiovascular disease," *Peer-to-Peer Networking and Applications*, vol. 12, pp. 1263–1276, 2019.
- [40] N. Senthilselvan, "An ontology-driven personalized food recommendation in iot-based healthcare system," *The Journal of Supercomputing*, vol. 75, pp. 3184–3216, 2018.
- [41] Y. Zhang and et al., "Future generation computer systems," vol. 66, 2017.
- [42] F. Ali, "Type-2 fuzzy ontologydaided recommendation systems for iotd-based healthcare," *Computer Communications*, 2017.
- [43] Anonymous, "A hybrid recommender system for patient-doctor matchmaking in primary care," *Unknown*, Unknown.
- [44] W. B. A. Karaa, E. Alkhamash, T. Slimani, and M. Hadjouni, "journal of healthcare engineering," vol. 2021, 2021.
- [45] V. S. Vairale and S. Shukla, "Recommendation of food items for thyroid patients using content-based knn method," in *Lecture Notes in Networks and Systems*, vol. 132, pp. 71–77, 2021.
- [46] S. Sun, Z. Zhang, X. Dong, H. Zhang, and T. Li, "Effects of antibiotic resistance genes (args) on bacterial community and args abundance during composting," *PloS One*, vol. 12, no. 8, p. 55, 2017.

- [47] H. El Bouhissi, D. Tagzirt, F. Bouredjioua, and O. Pavlova, “Recommender system for smart cities,” in *MoML&DS-2023*, 2023. In press.
- [48] T. Chen, G. Zhang, X. Hu, and J. Xiao, “Unmanned aerial vehicle route planning method based on a* algorithm,” in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, (Wuhan, China), pp. 1510–1514, 2018.
- [49] A. Subasi, “Clustering examples,” in *Practical Machine Learning for Data Analysis Using Python*, Elsevier, 2020.
- [50] G. Patel, V. Dabhi, and H. Prajapati, “An intelligent recommender system for web personalization using user behavior analysis,” *Journal of Intelligent Systems*, vol. 25, no. 3, pp. 315–327, 2016.
- [51] Mozilla, “React - documentation mozilla.” <https://fr.legacy.reactjs.org/>.
- [52] Mozilla Developer Network, “axios - documentation mozilla.” <https://axios-http.com/docs/intro>, dernier accès.
- [53] Mozilla, “Css - documentation mozilla.” <https://developer.mozilla.org/en-US/docs/Web/CSS>.
- [54] Mozilla, “Tinydb - documentation mozilla.” <https://tinydb.readthedocs.io/en/latest/>.
- [55] Mozilla, “Vs code - documentation mozilla.” <https://code.visualstudio.com/>.
- [56] Django, “Django.” <https://www.djangoproject.com/>.

Résumé

Les systèmes de recommandation sont utilisés pour fournir des items (par exemple : films, musiques, articles, medecins, etc.) adaptés aux préférences des utilisateurs. Ils estiment l'intérêt d'un utilisateur pour une ressource donnée en se basant sur certaines de ses informations et sur les choix des autres utilisateurs similaires.

Dans le cadre de ce mémoire, notre objectif est de développer un SR basé sur le filtrage collaboratif. Ce système permet de trouver les informations qui satisferont les utilisateurs en se basant sur les évaluations des autres utilisateurs. Nous nous concentrons spécifiquement sur la recommandation de médecins pertinents pour aider les touristes souffrant de maladies dans une ville intelligente.

Nous proposons donc "DocFast", un nouveau système de recommandation basé sur l'utilisateur et les connaissances qui résout le problème du démarrage à froid. Nous intégrons des techniques d'apprentissage automatique telles que K-means pour classer les médecins en fonction des symptômes clés de leur spécialité, et ajoutons au cluster les patients déjà traités. De plus, nous utilisons l'algorithme PSO avec K-means pour optimiser la sélection des caractéristiques dans le cluster.

Dans ce travail, nous implémentons "DocFast" et présentons son concept général en utilisant deux ensembles de données que nous avons créés. Nous les analysons afin d'obtenir des résultats précis et pertinents. L'évaluation de notre système a montré des résultats satisfaisants, avec une performance de 92%.

Mots clés :

Apprentissage automatique, cluster, filtrage collaboratif, item, K-means, PSO, SR, ville intelligente, démarrage à froid

Abstract

Recommender systems are used to provide users with tailored items (e.g., movies, music, articles, recipes, etc.) based on their preferences. These systems estimate a users interest in a particular resource by leveraging their information and the choices made by similar users.

In this thesis, our objective is to develop a collaborative filtering-based recommendation system. This system aims to find information that satisfies users by utilizing evaluations from other individuals. Specifically, we focus on recommending relevant doctors to assist tourists suffering from illnesses in a smart city.

Introducing "DocFast" a novel user- and knowledge-based RS designed to address the cold- start problem. It incorporates machine learning techniques such as K-means, enabling the classification of doctors based on key symptoms related to their specialization. Additionally, we employ the PSO algorithm alongside K-means to optimize feature selection within the cluster.

This study implements and presents the general idea of "DocFast" using two self-created datasets. Thorough analysis of these datasets allows us to obtain precise and relevant results. The evaluation of our system demonstrates fairly good performance, achieving an accuracy rate of 92%.

Keywords : Cluster, Collaborative filtering, Cold-start, Items, K-means, Machine learning, PSO, RS, Smart city.