

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة سعيدة د. مولاي الطاهر

كلية الرياضيات و الإعلام الآلي و الاتصالات السلكية و اللاسلكية

قسم: الإعلام الآلي



Mémoire de Master en informatique

Spécialité : Intelligence artificielle : Principes et Applications

Thème

Création d'un système d'aide à la décision
pour le diagnostic des maladies rénales

Présenté par :

Souyah Sara

Dirigé par :

Dr. Boudjellali Latifa

Année universitaire



2024-2025

شكر وتقدير

الحمد لله الذي بنعمته تتم الصالحات، وبتوفيقه يكتمل العمل.

أتقدم بجزيل الشكر والامتنان أولاً إلى والديّ العزيزين، لما قدّماه لي من دعم، دعاء، وتضحيات طوال مشواري الدراسي، فبفضلهما بعد الله، وصلت إلى هذه المرحلة من النجاح.

ثم أتوجه بخالص الشكر والتقدير إلى أستاذتي المشرفة الفاضلة، التي لم تبخل عليّ بتوجيهاتها السديدة وملاحظاتها القيمة، وكانت خير مُرافقة في هذا المسار الأكاديمي، فجزاها الله عني خير الجزاء.

كما لا يفوتني أن أقدم أسمى عبارات الشكر إلى جميع أساتذة الجامعة الأفاضل، الذين تعلمت على أيديهم واكتسبت من علمهم وتجربتهم، وإلى جامعتي الغالية، التي وقرت لنا بيئة علمية محفزة ومناسبة للتعلم والتطور.

ولا أنسى أن أخصّ بالشكر والتقدير مؤسسة النهضة بالقرآن الكريم، التي غرست فينا حب العلم والقرآن، وكانت منارة نور في بدايات طريقنا العلمي، فدمت لنا فخراً وذخراً.

كما أتوجه بجزيل الامتنان إلى الدكتورة أسماء من ولاية مستغانم، والدكتورة لعباني نور الهدى من جامعة سعيدة، اللّتين لم تبخلا عليّ بوقتتهما وجهدهما، ووقفنا إلى جانبي بكل كرم وتفانٍ.

وأتوجه بالشكر الخالص إلى عائلتي الكريمة، وإلى إخوتي وأخواتي الأعمام، الذين كانوا لي دعماً معنوياً لا ينضب، وملاًداً في كل لحظات الضعف والتعب، فلكم مني كل الحب والامتنان والدعاء من القلب.

كما لا يفوتني أن أذكر صديقتي العزيزات، اللواتي كنّ النور في أيام التعب، والسند في لحظات التحدي. وأشكر من القلب كل من درس معي وشارك هذه الرحلة، فلولا دفنكم وتشجيعكم، لما كان الطريق بهذه الروعة.

ولا يسعني إلا أن أخصّ بالشكر فلسطين الجريحة، الصامدة في وجه الألم والخذلان، رمز الصبر والكرامة والعزّة، لك يا فلسطين، كل الدعاء والنور، ولأجلك نرتقي ونُنجز ونكتب، علّ العلم يحمل بعض الوفاء.

جزيل الشكر لكل من ساهم في هذا الإنجاز، من قريب أو بعيد،

ودمتُم جميعاً في القلب.

Sommaire

Chapitre 1 : État de l'art sur les systèmes d'aide à la décision médicale	5
1.1 Introduction.....	5
1.2 Le système d'aide à la décision médicale	6
1.2.1 Définition.....	6
1.2.2 La modélisation d'un système d'aide à la décision médicale	7
1.2.3 Types des systèmes d'aide à la décision médicale.....	8
1.3 Travaux connexes.....	10
1.3.1 Des systèmes d'aide à la décision pour le diagnostic de diabète	10
1.3.2 Des systèmes d'aide à la décision pour le diagnostic des maladies cardiovasculaire	11
1.3.3 Des systèmes d'aide à la décision pour le diagnostic des maladies rénales.....	12
1.4 Les techniques intelligentes en aide à la décision	13
1.4.1 Les fouilles de données (Data mining).....	14
1.4.2 Le raisonnement à partir de cas	15
1.4.3 Les réseaux de neurones	16
1.4.4 L'apprentissage automatique (Machine Learning)	17
1.5 Conclusion	18
Chapitre 2 : Architecture proposée pour l'aide au diagnostic des maladies rénales	19
2.1 Introduction.....	19
2.2 Diagnostic d'une maladie rénale	19
2.2.1 Définition et classification	19
2.2.2 Symptômes fréquents	20
2.3 Architecture proposée : Adaptation par DFG (débit de filtration glomérulaire).....	20
2.3.1 L'architecture	20
2.3.2 Description de la base de cas	23
2.3.3 Description de la base de données des patients	23
2.3.4 Algorithme d'adaptation par débit de filtration glomérulaire (GFR).....	25
2.4 Intégration des plans de surveillance pour un patient.....	28
2.4.1 Stades 1 et 2 :	28
2.4.2 Stade 3 (a et b).....	29
2.4.3 Stades 4 et 5	29
2.5 Description de l'UML.....	30
2.5.1 Définition et historique de l'UML.....	30
2.5.2 Les diagrammes d'UML	30
2.6 Représentation graphique des diagrammes UML	32
2.7 Conclusion	39

Chapitre 3 : L'implémentation de notre logiciel ClinicPro	41
3.1. Introduction	41
3.2. Description de l'environnement de travail	41
3.2.1 Description du matériel utilisé	41
3.2.2 Les langages utilisés dans l'application	41
3.2.2.1 <i>Python</i>	41
• PyQt6	42
• SQLite	43
3.2.2.2 <i>HTML & CSS pour la génération des rapports médicaux</i>	43
3.3 Environnement de développement :	44
3.3.1 Visual Studio Code	44
3.3.2 BD Browser for SQLite	44
3.4 Intégration du Machine Learning dans un Système d'Aide à la Décision Médicale (SADM)	45
3.4.1 Contexte	45
3.4.2 Définition Technique du Machine Learning	45
3.4.3 Type d'Apprentissage Utilisé : Supervised Learning	46
3.4.4 Modèle Choisi : Arbre de Décision (Decision Tree Classifier)	46
3.5 Architecture Générale du SADM Intégré au ML	47
3.5.1 Acquisition et structuration des données :	47
3.5.2 Prétraitement des données :	47
3.5.3 Sélection des Caractéristiques (Feature Selection) :	48
3.5.4 Division des données en ensembles d'entraînement et de test	49
3.5.5 Entraînement du modèle DecisionTreeClassifier	49
3.5.6 Evaluation du modèle	50
3.5.7 Sauvegarde du modèle	53
3.5.8 Visualisation	53
3.5.9 Implémentation dans l'application ClinicPro	55
3.6 Fonctionnement de notre application ClinicPro	55
3.6.1 La page Login	56
3.6.2 <i>Page de sélection de rôle</i>	58
3.6.3 Interface infirmière	59
3.6.4 Interface médecin	61
3.7 Discussion	74
3.8 Conclusion	75
Références	76

Chapitre 1 : État de l'art sur les systèmes d'aide à la décision médicale

1.1 Introduction

Le but des ordinateurs était de décharger l'être humain des calculs longs et fastidieux qu'il était amené à réaliser. Dans les années 1980, les ordinateurs dépassent largement les premiers objectifs de simples machines à calculer. De nouveaux objectifs plus ambitieux sont envisagés : la simulation, les phénomènes sociologiques, des systèmes biologiques, des processus de décision...

Des la fin de la deuxième guerre mondiale dans les années 1945, l'intelligence artificielle (IA) se développe. Son but est d'imiter les comportements intelligents. L'IA cherche soit à rendre des systèmes autonomes, soit à aider l'homme dans ses activités quotidiennes. Pour cela, elle s'inspire de la nature humaine : raisonnements, sociologie, psychologie, structure cérébrale, le développement de projet de recherche. Dans les années 70, les premiers systèmes experts sont mis au point. Ceux-ci sont basés sur des règles et un moteur d'inférence. Ils permettent de prendre en considération des raisonnements humains fortement structurés, basé sur la logique propositionnelle. Les processus d'aide à la décision sont utilisés au quotidien par des millions de personnes, ils interviennent dans des domaines très divers : GPS automobile, finance, environnement et développement durable, domaine spatial, politiques publiques, gestion des ressources hospitalières, etc. L'aide à la décision peut permettre de trouver une solution où l'homme n'en trouvait pas ou encore une solution à des problèmes nouveaux où l'homme n'a aucune expérience. Elle peut encore aider à juger de la qualité d'une solution, mais aussi à confirmer/justifier des décisions. Elle favorise une culture scientifique dans la prise de décision [1].

Les **systèmes d'aide à la décision médicale** (SADM) permettent théoriquement une amélioration de la qualité des soins, mais leur intégration pratique est encore loin d'être une réalité pour la plupart des professionnels de santé. Les principaux écueils à cette intégration pratique sont l'évolution constante des connaissances en santé, les difficultés d'interopérabilité au sein des systèmes d'information médicaux et les résistances des professionnels [2].

Dans ce chapitre, nous recensons plusieurs travaux connexes récents qui permettent de pallier partiellement une ou plusieurs de ces problématiques et qui devraient faciliter l'intégration des SADM dans les établissements de santé. Néanmoins, l'intrication étroite entre les recommandations, les moyens à la disposition des médecins et les pratiques locales imposent une évaluation précise de l'impact des SADM de manière à en limiter les conséquences négatives pour les patients.

1.2 Le système d'aide à la décision médicale

Grâce aux progrès de l'informatique et à la collecte systématique d'informations sur les patients, de nouveaux outils médicaux ont été développés. Certains servent à guider le médecin dans sa décision, d'autres peuvent prendre automatiquement des décisions. Appelés « systèmes experts » ou « systèmes d'aide à la décision », ils peuvent participer au diagnostic, au pronostic ou au soin. Ils fonctionnent tous au moyen d'algorithmes traitant les données. Ces algorithmes sont plus ou moins complexes, plus ou moins supervisés, mais ils sont aussi très souvent opaques. Dès lors que ces outils prennent une place toujours plus grande dans le monde médical, il apparaît nécessaire de mieux informer les utilisateurs et les bénéficiaires – médecins et patients – de leurs logiques de fonctionnement, de leurs limites et de leurs éventuels biais [3].

1.2.1 Définition

L'intelligence artificielle désigne la capacité des machines à imiter le comportement intelligent de l'humain. Elle englobe diverses techniques permettant aux machines de percevoir leur environnement, de tirer des conclusions ainsi que de prendre des décisions. Les systèmes d'IA (intelligence artificielle) vont des simples algorithmes de prise de décision aux systèmes complexes capables d'apprentissage et d'adaptation autonomes [4].

Serroussi et Bouaud [5] définissent les systèmes d'aide à la décision médicale (SADM) comme étant « des applications informatiques dont le but est de fournir aux cliniciens, en temps et en lieu utiles, les informations décrivant la situation clinique d'un patient, ainsi que des

connaissances appropriées à cette situation qui soient correctement filtrées et présentées afin d'améliorer la qualité des soins et la santé des patients ».

Berner précise aussi que pour « améliorer la qualité des soins et la santé des patients les informations doivent également être : correctement filtrées et présentées ; fournies en temps et lieu utiles » [6].

1.2.2 La modélisation d'un système d'aide à la décision médicale

Le modèle global d'un SADM tel qu'il est présenté dans la littérature se compose essentiellement de trois couches, Benmimoune L.[7] a proposé un modèle global d'un SADM, la figure 1.1 illustre ce modèle.

- **Couche d'acquisition des données** : cette couche permet l'acquisition et l'introduction de données au système. Ces données sont relatives au patient, elles peuvent être de différentes natures : données médicales (signes, symptômes, antécédents médicaux, données physiologiques, etc.), données personnelles (sexe, âge, profession, etc.), données sur l'hygiène de vie (activité physique, tensions artérielles, taux de glycémies, etc.), etc. La manière dont les données sont introduites varie d'un système à l'autre. Certains systèmes imposent aux médecins de sélectionner les données depuis un vocabulaire contrôlé à l'aide de formulaires de saisie. D'autres systèmes extraient ces données depuis les dossiers médicaux électroniques. Des capteurs médicaux (tensiomètre, oxymètre, pèse-personne, etc.) et des questionnaires informatisés peuvent aussi être utilisés pour l'acquisition de données médicales.
- **Couche décisionnelle** : appelée aussi couche de raisonnement, c'est la partie intelligente du système qui permet, à partir des données introduites par la couche d'acquisition de données, de fournir des résultats à la couche de présentation. Nous pouvons distinguer un ou plusieurs modes de raisonnement. Ces modes de raisonnement permettant d'arriver à un résultat voire à plusieurs résultats pouvant être intriqués. Généralement les modes de raisonnement font appel à différents aspects de l'intelligence artificielle et des techniques de fouille de données dont le but est de simuler le raisonnement humain et, en l'occurrence, le raisonnement des experts médicaux.
- **Couche de présentation des données** : cette couche permet de présenter, à l'utilisateur, les résultats fournis par le système. Plusieurs types et formes de résultats peuvent être donnés en fonction du mode d'intervention du SADM.

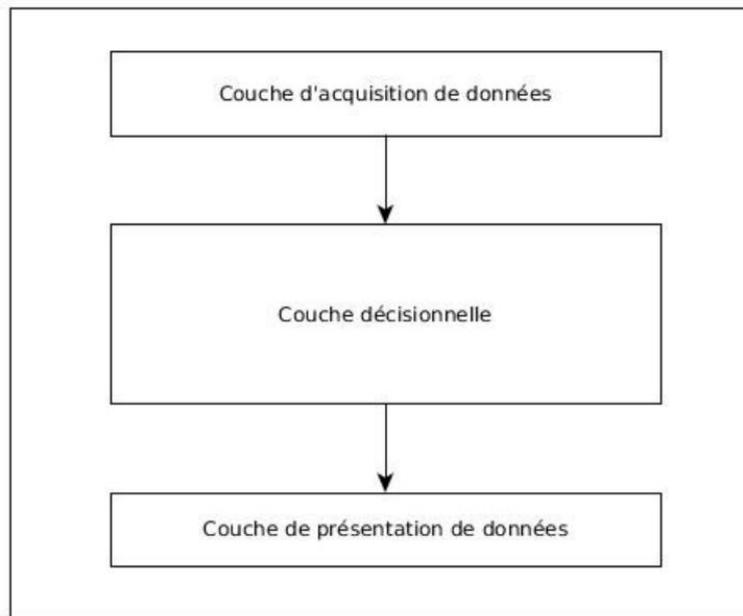


Figure 1.1 Modèle global d'un SADM [7]

1.2.3 Types des systèmes d'aide à la décision médicale

Afin d'accompagner les professionnels de santé dans leur démarche clinique, plusieurs systèmes de suivi et de prise en charge médicale ont été construits et déployés dans le milieu hospitalier. Ces systèmes permettent principalement de collecter des données médicales sur les patients, de les analyser et de présenter les résultats de différentes manières. Ils représentent un appui et une aide aux professionnels de santé dans leur prise de décision par rapport à l'évolution de l'état de santé des patients suivis. L'utilisation de tels systèmes nécessite systématiquement une adaptation à la fois au domaine médical concerné et au mode d'intervention. Il est nécessaire, dans un milieu hospitalier, que ces systèmes puissent s'adapter et évoluer d'une manière simple, en limitant toute maintenance corrective ou évolutive. Ils doivent être en mesure de prendre en compte dynamiquement des connaissances théoriques et empiriques du domaine issues des experts médicaux.

Un système d'aide à la décision clinique (CDSS – Clinical Decision Support System) est un aspect important des technologies de l'information en santé [8]. L'objectif principal des CDSS est d'aider les cliniciens dans les processus de prise de décision. Reed T. Sutton et al. [9] avaient défini deux types de CDSS :

- **Le CDSS traditionnel** comme une aide directe à la prise de décision clinique, dans lequel les facteurs individuels du patient sont adaptés à une base de connaissances cliniques informatisée afin de présenter des évaluations spécifiques au clinicien.

- **Les CDSS modernes** comme un outil au point de service que les cliniciens utilisent pour combiner leurs connaissances avec les suggestions fournies par le CDSS.

Selon A. T. M. Wasylewicz et al. [10], pour comprendre et se familiariser avec les CDSS, il est important de les catégoriser en se basant sur :

- La fonction du système
- Le modèle de conseil
- Le style de communication
- Le processus décisionnel sous-jacent
- L'Interaction Humain-Machine (IHM)

La fonction du système vise à déterminer « ce qui est vrai ? » sur la base d'un ensemble de données saisies par l'utilisateur ou facilement disponibles, et « que faire ? », par exemple pour aider au pronostic du patient. Le modèle de conseil est l'ensemble des approches actives ou passives utilisées pour donner des conseils. Le style de communication catégorise le CDSS en distinguant deux modèles :

1. **Modèle de consultation** : le système est un conseiller qui interroge et propose des solutions/ actions. Par exemple suggestion d'un traitement en fonction du diagnostic.

2. **Modèle de critique** : le système permet à l'utilisateur de décider par lui-même et l'alerte s'il y a un risque autour de la décision prise.

Le processus décisionnel sous-jacent représente les méthodes utilisées pour l'aide à la décision, que ce soit des modèles basés sur la pratique médicale, la littérature, des approches statistiques ou l'intelligence artificielle, etc. L'IHM désigne la communication et l'interaction entre l'utilisateur et la machine par le biais d'une interface utilisateur qui permet de transmettre des informations de la machine à l'humain de manière interactive en utilisant des transactions de type requête-réponse. Les interfaces humain-machine [11] ont le potentiel de faciliter l'accès et la compréhension des approches de la médecine de précision. Elles permettent l'interaction avec les données en temps réel, afin d'obtenir les informations nécessaires pour prendre la bonne décision au bon moment. Un CDSS est une combinaison de données et de connaissances conçue pour fournir aux professionnels de la santé et aux patients des informations portant sur des procédures médicales susceptibles d'aider à prendre une décision individualisée pour chaque patient. Les CDSS offrent un large éventail de fonctions. Ils peuvent intervenir comme support au diagnostic, au pronostic ou encore au choix thérapeutique pour accélérer l'identification de pathologies souvent complexes et aider à leur prise en charge.

1.3 Travaux connexes

Le développement des CDSS est de plus en plus marqué : analyse des données d'IRM (Imagerie par Résonance Magnétique) pour caractériser les tumeurs [12], aide à la décision médicale dans les maladies complexes [13], aide aux pharmaciens pour fournir à leurs clients des conseils factuels en toute simplicité [14], aide à la gestion des pandémies [15], etc. En tant qu'espace d'interaction et de visualisation des données, les CDSS ont le potentiel d'accroître l'adhésion aux procédures cliniques en enrichissant l'interaction médecin-patient. Dans ce chapitre, nous avons sélectionné trois familles des maladies chroniques pour l'état de l'art qui sont le diabète, les maladies cardiovasculaires et les maladies rénales.

1.3.1 Des systèmes d'aide à la décision pour le diagnostic de diabète

Le diabète est une maladie chronique fréquente, emblématique des enjeux liés à la transition épidémiologique et aux inégalités sociales de santé. Un système de surveillance du diabète de type 1 et de type 2, reposant sur un dispositif multi-sources et des développements méthodologiques innovants, a été mis en œuvre en France afin de produire un ensemble d'indicateurs pertinents, permettant de mesurer le fardeau du diabète, qu'il s'agisse de la charge de comorbidités, de mortalité ou des facteurs de risque ainsi que son évolution spatiotemporelle [16]. Cette surveillance doit s'intégrer dans un contexte territorial et socio-économique afin de guider l'élaboration des politiques de prévention, de prise en charge et d'en évaluer leur impact, à un niveau national, local, ou au sein de populations spécifiques.

Merzoug et al. [17] ont conçu un outil d'aide au diagnostic basé sur une approche d'apprentissage supervisée, qui consiste, à un détecteur automatique de la rétinopathie diabétique, basé sur la combinaison d'une segmentation d'images, d'une extraction de caractéristiques et d'une classification binaire de type SVM. Les auteurs ont comme objectif de créer un système de détection automatique de la rétinopathie diabétique, et l'évaluation de ces performances, en termes de son pouvoir discriminatif entre une population rétinopathique et une autre saine.

L'utilisation de systèmes de classification dans le domaine médical augmente progressivement, et ces systèmes peuvent réduire le temps nécessaire au diagnostic et fournir une aide précieuse pour réduire les erreurs. Le système immunitaire artificiel est un domaine de recherche qui relie les domaines de l'immunologie, de l'informatique et de l'ingénierie. Hamdi et al. [18] ont développé une méthode AIS telle que la sélection négative. Dans leur

travail ils ont appliqué cette méthode pour identifier le diabète, puis essayer de donner des perspectives à leur étude, appliquée sur la base du diabète. Une modélisation synthétique de sélection négative avec machine Learning algorithmes proposée pour implémenter cette application.

1.3.2 Des systèmes d'aide à la décision pour le diagnostic des maladies cardiovasculaire

Les maladies cardiovasculaires deviennent l'un des problèmes de santé publique les plus importants dans la plupart des régions du monde, une approche préventive éprouvée aux problèmes de santé cardiovasculaire est la pratique de l'activité physique régulière. Les exercices d'activité physique sont associés à une diminution de du risque de développer une maladie cardiovasculaire et de sa létalité. Pour aider à ancrer et à suivre les habitudes d'activité physique au sein de la population, dans le même ordre d'idée, Zinnour et al. [19] ont développé un système d'aide à la décision clinique PODiSanté (Plateforme Opérationnelle d'analyse de Données sur la pratique de l'activité physique appliqué à la santé) pour prescrire des exercices et proposer des recommandations aux patients. Leur système d'aide à la décision clinique utilisant l'apprentissage profond AP-SADC, un domaine de l'apprentissage automatique pour aider les cliniciens à prendre de meilleures décisions concernant l'activité physique des patients visant à améliorer leur santé. L'élément clé de leur AP-SADC proposé est l'estimation de l'incertitude qui l'accompagne dans la phase de traitement en raison de différents facteurs. Ils se sont concentrés sur le AP-SADC et la façon bayésienne de penser pour la prise de décision, et de capture de l'incertitude. Différents modèles avec différentes architectures ont été proposés dans ce travail pour trouver le meilleur modèle possible. Ils ont réalisé de multiples évaluations sur les différents modèles testés afin de mieux établir la performance et la faisabilité du POD-iSanté.

Une autre contribution est proposée par Mokadem S. [20] concerne la fouille de traces qui se base sur l'aide à la décision médicale pour proposer un système d'aide au diagnostic clinique basé traces, fondé sur l'apprentissage automatique et le raisonnement flou. L'auteur a choisi les maladies cardiovasculaires, qui sont la première cause de mortalité, pour la conception et la validation de son système. Il s'agit de mettre en place une nouvelle approche de classification SIPINA et la modélisation floue. Les performances de cette contribution ont été validé à travers des expérimentations et les résultats ont été encourageants.

Dans un autre travail récent, Zemouli [21] a développé un système de prédiction des maladies cardiaques capable de détecter l'apparition précoce d'une maladie cardiaque, qui peut

s'avérer fatale. La démarche de son projet de recherche consiste à détecter les maladies cardiaques à l'aide de techniques d'apprentissage automatique supervisé. Pour ce faire, elle a exploré plusieurs algorithmes d'apprentissage automatique supervisé afin d'identifier le modèle le plus efficace pour obtenir des résultats optimaux. Elle a concentré son étude sur l'application de ces modèles à la détection précoce des maladies cardiaques, ce qui pourrait contribuer à améliorer les taux de pré-diagnostic tout en prévenant les complications associées à ces maladies. Son travail s'est concentré sur l'amélioration de l'algorithme de l'arbre de décision (DT) pour la détection des maladies cardiaques, qui était déjà considéré comme le meilleur algorithme dans ce domaine. En outre, elle a proposé des améliorations spécifiques à cet algorithme afin d'accroître son efficacité. En utilisant le même ensemble de données que d'autres algorithmes d'apprentissage automatique supervisé, L'auteur a pu démontrer que son algorithme amélioré produisait de meilleurs résultats. Ces résultats renforcent l'idée que son approche offre une solution plus efficace pour la détection des maladies cardiaques que les autres algorithmes existants.

1.3.3 Des systèmes d'aide à la décision pour le diagnostic des maladies rénales

L'insuffisance rénale chronique (IRC) est de plus en plus reconnue comme un problème de santé majeur en raison de sa prévalence croissante. La survie moyenne sans reins fonctionnels est généralement limitée à environ 18 jours, ce qui crée un besoin important de transplantations rénales et de dialyses. La détection précoce de l'IRC est cruciale, et les méthodes d'apprentissage automatique se sont avérées efficaces pour diagnostiquer cette maladie, malgré des processus décisionnels souvent opaques. L'auteur dans [22] a utilisé l'apprentissage automatique explicatif pour prédire l'IRC, surmontant ainsi le caractère « boîte noire » des prédictions traditionnelles de l'apprentissage automatique. Parmi les six algorithmes d'apprentissage automatique évalués, l'algorithme Extreme Gradient Boost (XGB) a démontré la plus grande précision. Pour l'interprétabilité, l'étude a utilisé les explications additives de Shapley (SHAP) et les graphiques de dépendance partielle (PDP), qui éclairent la logique des prédictions et soutiennent le processus décisionnel. De plus, pour la première fois, une interface utilisateur graphique explicative a été développée pour diagnostiquer la probabilité d'une IRC.

Compte tenu de la nature critique et des enjeux élevés de l'IRC, selon Gangani et al., l'utilisation de l'apprentissage automatique explicable peut aider les professionnels de la santé à établir des diagnostics précis et à identifier les causes profondes.

Dans les scénarios de prise de décision incertains, les modèles et méthodes de prédiction scientifique quantitative peuvent apporter une aide précieuse à la prise de décisions scientifiques. Cependant, les caractéristiques des informations d'attributs hybrides peuvent engendrer une série de problèmes, notamment des difficultés de comparaison et d'évaluation exhaustive de différents types d'attributs, des relations non linéaires entre eux et un manque de méthodes d'aide à la décision efficaces. Pour surmonter ces difficultés, l'étude proposée par Jiqian et al. [23] introduit une fonction noyau permettant d'abstraire la similarité des différents types d'attributs et propose un modèle appelé « Kernel Multi Granularity Fuzzy Rough Sets » (KMGFRS). Le modèle proposé dans cette étude facilite l'exploration et l'analyse approfondies des incertitudes associées aux objets de décision. De plus, Les auteurs ont proposé une méthode de réduction d'attributs basée sur KMGFRS pour traiter les attributs redondants dans les systèmes d'information hybrides. Leur méthode élimine les attributs ayant une influence minimale sur les résultats de la décision, simplifie le processus décisionnel et améliore son efficacité. Cette étude intègre les concepts de KMGFRS et d'apprentissage profond hybride pour proposer une nouvelle méthode de prédiction visant à améliorer la précision et la robustesse. Du point de vue des informations sur les attributs hybrides, cette méthode permet de prédire avec plus de précision les attributs inconnus des objets de décision, offrant ainsi un support solide à la prédiction des maladies dans le diagnostic médical et la prise de décision thérapeutique. Les résultats expérimentaux de cette étude ont montré que le modèle construit gérait efficacement les scénarios de prise de décision incertains impliquant des attributs hybrides et des objets de décision flous donc les auteurs ont fourni une aide à la décision précise et fiable pour l'insuffisance rénale chronique (IRC), améliorant considérablement la précision prédictive des types d'IRC.

1.4 Les techniques intelligentes en aide à la décision

Les techniques intelligentes sont utiles pour analyser les données et fournir des prévisions, quantifier l'incertitude, fournir facilement des informations et suggérer la marche à suivre. La figure 1.2 souligne comment la prise de décision est influencée par une gamme variée de techniques intelligentes.

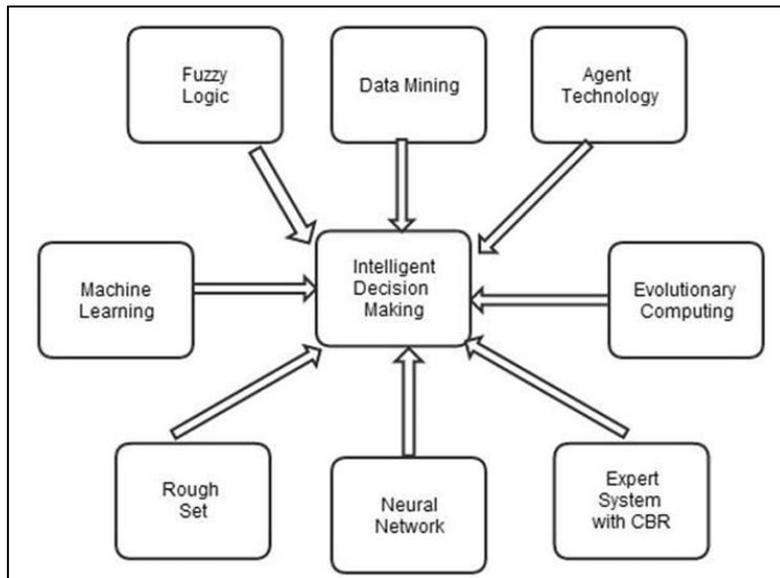


Figure 1.2 Les techniques intelligentes en aide à la décision [24][25].

1.4.1 Les fouilles de données (Data mining)

Les avancées technologiques tant au niveau des volumes de stockage qu’au niveau des capacités de calcul et de transfert de données ont profondément modifié l’environnement de travail dans le domaine de la santé. L’informatisation des dossiers patients et le nombre exponentiel de données produites permettent d’envisager la réutilisation des données patient dans un autre cadre que celui du soin. Cette seconde vie des données est rendue possible par leur désenclavement au sein d’entrepôts de données conçus pour intégrer des données biomédicales provenant de sources multiples et aux formats très hétérogènes : texte libre (compte rendu hospitalier), donnée codée (code diagnostic), donnée numérique (résultat biologique). Dans ce contexte, des équipes de recherche en bio-informatique et en informatique médicale développent des outils de fouille de données (extraction de connaissance) pour optimiser la recherche biomédicale mais aussi pour améliorer la prise en charge du patient. La fouille de données met en œuvre des méthodes statistiques, de traitement automatique du langage, de visualisation de données voire d’intelligence artificielle [26].

Le fouille de données est le cœur du processus d’extraction de connaissances, ce dernier est défini par Zighed et al. [27] : « L’Extraction de Connaissances à partir des Données (ECD) est un processus itératif et interactif d’analyse d’un grand ensemble de données brutes afin d’en extraire des connaissances exploitables par un utilisateur analyste qui y joue un rôle central ».

Le “data mining” que l’on peut traduire par “fouille de données” apparaît au milieu des années 1990 aux États-Unis comme une nouvelle discipline à l’interface de la statistique et des

technologies de l'information : bases de données, intelligence artificielle, apprentissage automatique (« machine learning »). Hand [28] en donne la définition suivante : « Data Mining consists in the discovery of interesting, unexpected, or valuable structures in large data sets ».

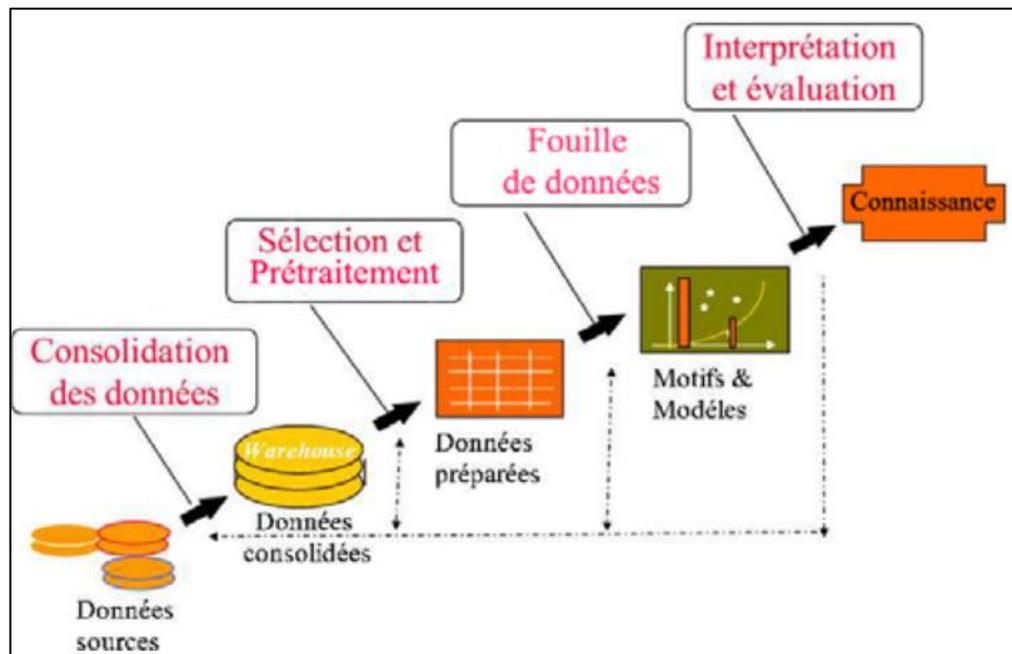


Figure 1.3. Le Processus d'extraction de connaissances à partir des données [29][30].

1.4.2 Le raisonnement à partir de cas

Le raisonnement à partir de cas est basé sur l'exploitation d'expériences antérieures pour résoudre, par analogie, de nouveaux problèmes ou cas (Voir la figure 1.4). Le RàPC est donc considéré comme un outil de calcul dont la seule source de connaissance est une mémoire des cas importants passés. Quand un problème est résolu, sa solution est retenue pour résoudre des problèmes similaires. Cela permet alors d'éviter de répéter les erreurs commises dans le passé.

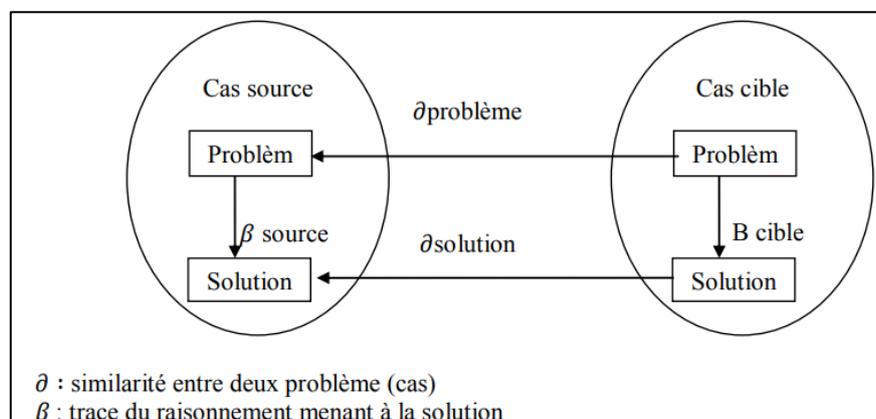


Figure 1.4. Le carré d'analogie [31].

1.4.3 Les réseaux de neurones

La naissance du domaine des réseaux de neurones artificiels remonte aux années 1940 avec les travaux de Warren McCulloch et Walter Pitts qui ont montré qu'avec de tels réseaux, on pouvait, en principe, calculer n'importe quelle fonction arithmétique ou logique. Vers la fin des années 1940, Donald Hebb a ensuite proposé une théorie fondamentale pour l'apprentissage [32]. Les réseaux de neurones artificiels sont définis par : "Des réseaux fortement connectés de processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit. Toute structure hiérarchique de réseaux est évidemment un réseau." [33]

Actuellement, l'utilisation des réseaux de neurones artificiels (RNA) s'est développée dans de nombreuses disciplines (sciences économiques, écologie et environnement, biologie et médecine...). Ils sont notamment appliqués pour résoudre des problèmes de classification, de prédiction, de catégorisation, d'optimisation, de reconnaissance ainsi dans les systèmes d'aide au diagnostic médical. Les Réseaux de Neurones Artificiels (RNA) sont inspirés de la biologie et représentent un modèle mathématique du fonctionnement du neurone biologique.

Un réseau de neurones est un maillage de plusieurs neurones, généralement organisé en couches. Pour construire une couche de S neurones, il s'agit simplement de les assembler comme à la figure 1.5. Les S neurones d'une même couche sont tous branchés aux R entrées. On dit alors que la couche est totalement connectée [34].

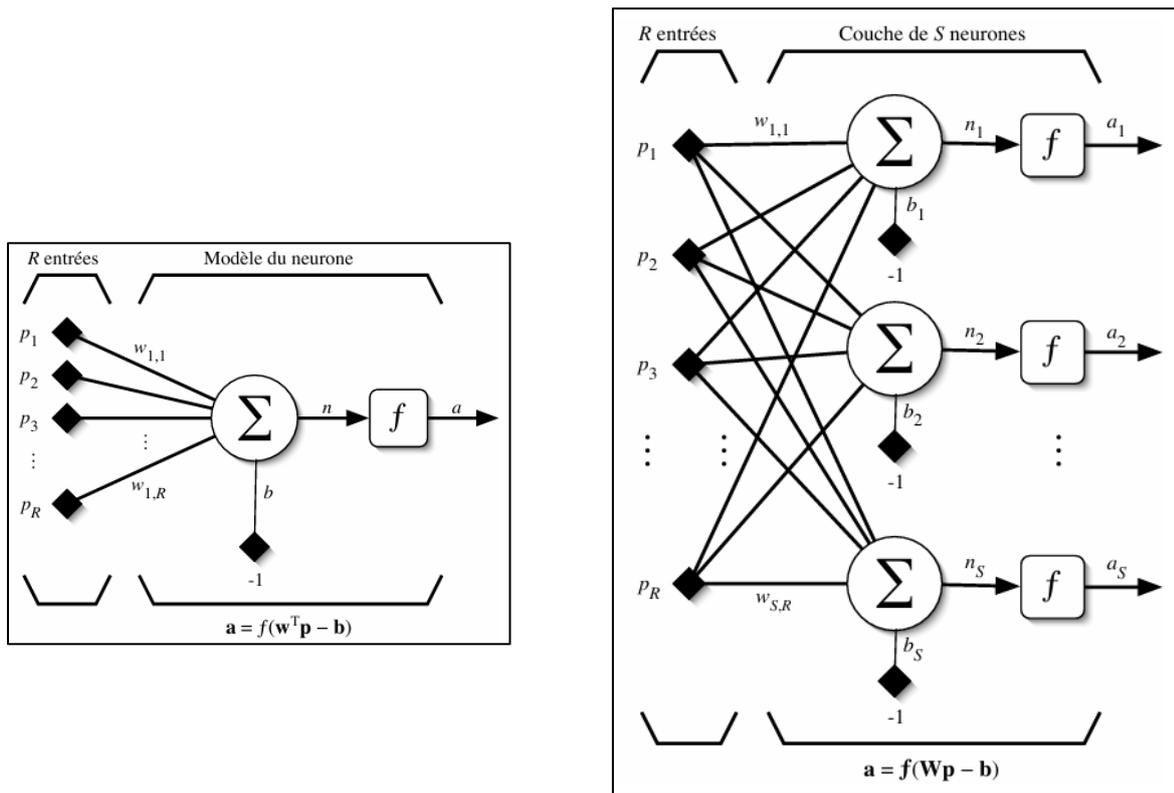


Figure 1.5. Modèle d'un neurone artificiel et la composition d'une couche de S neurones.

1.4.4 L'apprentissage automatique (Machine Learning)

L'apprentissage automatique est un sous-domaine de l'IA qui fait référence aux programmes qui calculent des règles de décision à partir de données empiriques annotées par des experts et qui optimisent la performance en termes de prédiction de résultats confrontés aux annotations des experts, par des essais itératifs prenant en compte les essais déjà réalisés. Ces programmes requièrent donc des données annotées ou classifiées et ont révélé leur potentiel avec des données complexes et structurées telles que des images, des vidéos, ou du texte [35].

Afin de développer des modèles d'apprentissage efficaces à partir d'ensembles volumineux de données (les datasets), il est recommandé de respecter un processus précis, nous avons résumé les étapes de ce processus comme suivant :

- Identifier les besoins et les objectifs du domaine d'application (Formuler le problème)
- Collecter les informations nécessaires
- Nettoyer et à normaliser (rendre comparables) les données brutes

- **Choisir le bon algorithme** pour traiter le problème (les arbres de décision par exemple)
- Recourir à des données d'entraînement en sélectionnant une partie du dataset (**échantillonnage**) afin d'entraîner plus efficacement le modèle et de perfectionner ses prédictions.
- Tester et déployer le modèle proposé.

1.5 Conclusion

L'objectif de ce chapitre est de présenter les différentes approches d'aide à la décision médicale et de dresser un panorama des techniques intelligentes qui ont vocation à entourer les pratiques dans un futur proche. Dans un premier temps, nous avons commencé par une recherche sur les définitions formelles existées dans la littérature d'un système d'aide à la décision médicale ainsi que sa modélisation composée de trois couches. Puis, nous avons présenté les types des SADM, et nous avons trouvé qu'il y a deux types (Traditionnel et moderne) et deux modèles : un système considéré comme conseiller qui interroge et propose des solutions appelé modèle de consultation, et un autre modèle critique qui permet à l'utilisateur de décider par lui-même et l'alerte s'il y a un risque autour de la décision prise.

Dans un autre temps, nous avons proposé quelques travaux connexes, notre recherche se limitera aux systèmes d'aide à la décision appliqué aux trois maladies majeures : Le diabète, les maladies cardiovasculaires et le diagnostic des maladies rénales.

À la fin de ce chapitre, nous avons sélectionné les techniques décisionnelles les plus utilisées par les chercheurs, qui sont des outils d'aide à la décision puissants offerts pour les praticiens. Nous avons synthétisé le raisonnement à partir de cas, les réseaux de neurones et l'apprentissage automatique (machine learning).

Dans le prochain chapitre, nous allons présenter le domaine du diagnostic des maladies rénales puis nous allons proposer et détailler notre approche d'aide à la décision pour les équipes praticiens spécialistes à cette maladie.

Chapitre 2 : Architecture proposée pour l'aide au diagnostic des maladies rénales

2.1 Introduction

Ce chapitre est consacré à l'approche que nous avons proposée et qui s'inscrit dans le cadre de l'aide à la décision pour faire un bon diagnostic des maladies rénales dans un temps réduit.

Avant de réaliser notre approche, nous avons commencer par une recherche sur la description et la définition d'une maladie rénale, les types, les symptômes, les analyses demandées par le médecin, le diagnostic et les prescriptions.

Afin de consolider notre recherche, nous avons réalisé une interview avec le « Dr Labani Nourelhouda », docteur à l'Université Moulay Tahar de Saïda. Cette rencontre nous a permis de mieux comprendre le processus médical lié au diagnostic des maladies rénales, les pratiques courantes en consultation, ainsi que les éléments cliniques et biologiques pris en compte.

2.2 Diagnostic d'une maladie rénale

L'insuffisance rénale chronique (IRC) est devenue un problème de santé publique dans beaucoup de pays à travers le monde et risque de le devenir en Algérie, si des mesures adéquates pour sa prévention et sa prise en charge ne sont pas prises dans un avenir immédiat [36].

2.2.1 Définition et classification

Une maladie rénale chronique est définie par la persistance pendant plus de 3 mois d'une atteinte rénale ou d'une baisse du débit de filtration glomérulaire (DFG) $< 60 \text{ mL/min pour } 1,73 \text{ m}^2$ de surface corporelle, indépendamment du type de néphropathie causale. L'atteinte rénale (kidney damage) est caractérisée, qu'il y ait ou non baisse du DFG, par la présence d'anomalies biologiques. (Protéinurie, microalbuminurie, leucocyturie, hématurie) et/ou histologiques et/ou morphologiques. Les méthodes de mesure du DFG étant complexes (clairance de l'inuline, méthode isotopique, etc.), il est recommandé en pratique d'apprécier la fonction rénale à partir d'équations d'estimation du DFG tenant compte du dosage de créatinine sérique, et de certains des facteurs suivants : l'âge, le sexe, l'ethnie et le poids, qui permettent de corriger l'effet de la masse musculaire sur le niveau de la créatininémie. Une classification en stades de sévérité, telle qu'il en existe pour l'insuffisance cardiaque et hépatique, a été proposée par deux équipes de recherches l'Anaes et K/DOQI. Cette classification est fondée sur le niveau de DFG, dont la valeur normale chez l'adulte jeune est d'environ $120 \text{ mL/min/1,73 m}^2$. Au-dessous de $60 \text{ mL/min/1,73 m}^2$, tous les individus sont considérés comme malades, qu'ils aient ou non des

marqueurs d'atteinte rénale. Au-dessus de ce seuil, la présence persistante de marqueurs d'atteinte rénale est nécessaire au diagnostic [37]. Les classifications de l'Anaes et des K/DOQI figurent dans le tableau 2.1. Si ces 2 classifications sont très comparables, quelques différences sont intéressantes à souligner pour la pratique clinique. La première concerne le nombre de stades, 4 pour l'Anaes et 5 pour les K/DOQI, les experts de l'Anaes ayant choisi de ne pas distinguer de stades de sévérité pour les DFG > 60 mL/min/1,73 m². La seconde a trait au choix de l'équation d'estimation du DFG [38].

Classification selon l'Anaes			Classification selon les K/DOQI		
Stade	Description	DFG (mL/min/1,73 m ²)	Stade	Description	GFR (mL/min/1,73 m ²)
1	Maladie rénale chronique* avec DFG ≥ 60 mL/min/1,73 m ²	≥ 60	1	Kidney damage† with normal or ↑ GFR	≥ 90
2	IRC modérée	30-59	2	Kidney damage with mild or ↓ GFR	60-89
3	IRC sévère	15-29	3	Moderate ↓ GFR	30-59
4	IRC terminale	< 15 ou traitement de suppléance	4	Severe ↓ GFR	15-29
			5	Kidney failure	< 15 or dialyse

DFG : débit de filtration glomérulaire (en anglais, GFR : *Glomerular Filtration Rate*).
 * Maladie rénale chronique définie par la persistance pendant plus de 3 mois d'anomalies rénales biologiques et/ou histologiques et/ou morphologiques.
 † *Kidney damage* : atteinte rénale définie par la présence pendant plus de 3 mois d'anomalies histologiques ou biologiques sanguines ou urinaires ou morphologiques.

Tableau 2.1 Classification en stade de la maladie rénale chronique selon les recommandations 2002 de l'agence nationale d'accréditation et d'évaluation en santé (Anaes) et la National Kidney foundation(K/DOQI)

2.2.2 Symptômes fréquents

La maladie rénale chronique est une détérioration ancienne et progressive de la fonction rénale. Les symptômes se développent lentement et aux stades avancés comprennent une anorexie, des nausées, des vomissements, une stomatite, une dysgueusie, une nycturie, une asthénie, une fatigue, un prurit, une diminution de l'acuité mentale, des contractions et des crampes musculaires, une rétention d'eau, une dénutrition, des neuropathies périphériques et des convulsions.

2.3 Architecture proposée : Adaptation par DFG (débit de filtration glomérulaire)

2.3.1 L'architecture

La démarche que nous avons adoptée est illustrée dans la figure 2.1

1. A partir de la base de données Pre-processed Chronic Kidney Disease Dataset qui contient 401 cas. Chaque cas est décrit par 25 attributs dont 24 représentent les descripteurs et le dernier la classe.

2. Après une élaboration du nouveau cas, une deuxième étape du raisonnement à partir de cas se lance en appliquant l'algorithme d'intelligence artificielle, afin de sélectionner les cas similaires au nouveau cas.
3. La troisième étape du raisonnement à partir de cas consiste à une adaptation par le débit de filtration glomérulaire (DFG, GFR en anglais), nous avons rédigé cet algorithme pour sélectionner le stade de la maladie rénale, l'algorithme est présenté dans la section 2.3.4.
4. Passer à l'étape importante qui donne la main à l'expert (le néphrologue) pour réviser le cas adapté en ajoutant les prescriptions et les conseils nécessaires.
5. La dernière étape consiste à ajouter le cas révisé à la base des patients.

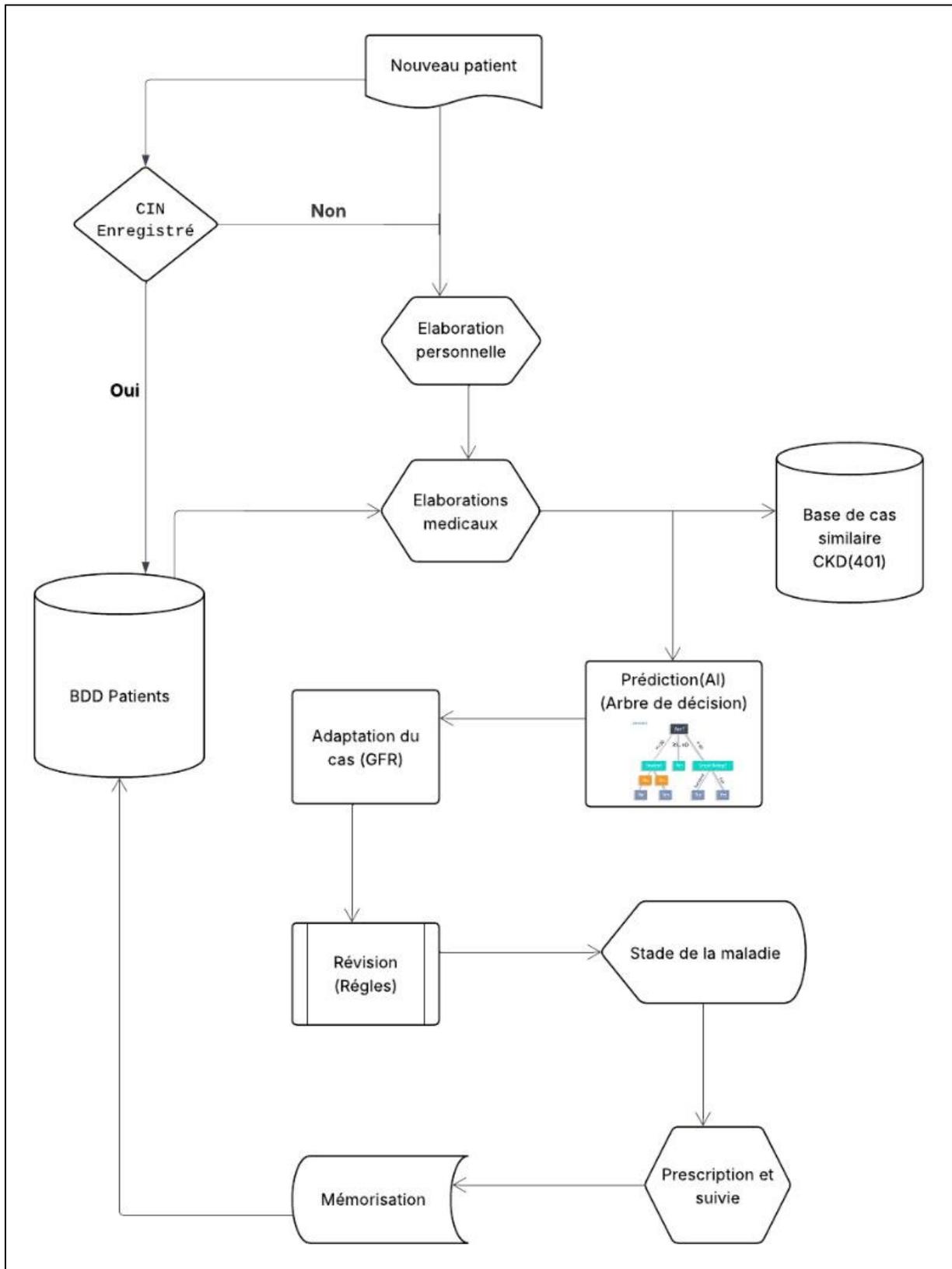


Figure 2.1. Schéma fonctionnel de l'approche

2.3.2 Description de la base de cas

La base de données Pre-processed Chronic Kidney Disease Dataset du KAGGLE intègre des données médicales concernant 401 patients, destinées à accélérer la détection précoce de la maladie rénale chronique. Chaque exemple de registre contient 24 caractéristiques cliniques et biologiques (par exemple, pression artérielle, densité urinaire, taux de créatinine, etc.) et une étiquette indiquant si le patient est ou non atteint d'une maladie rénale chronique. Cette base est souvent employée à des fins de projet de classification et d'analyse prédictive en santé.

2.3.3 Description de la base de données des patients

La base de données est composée de trois tables :

- Table Patients : contient les informations personnelles
- Table Prédiction : contient les informations médicales
- Table Medical_History : contient les informations nécessaires pour le diagnostic médical

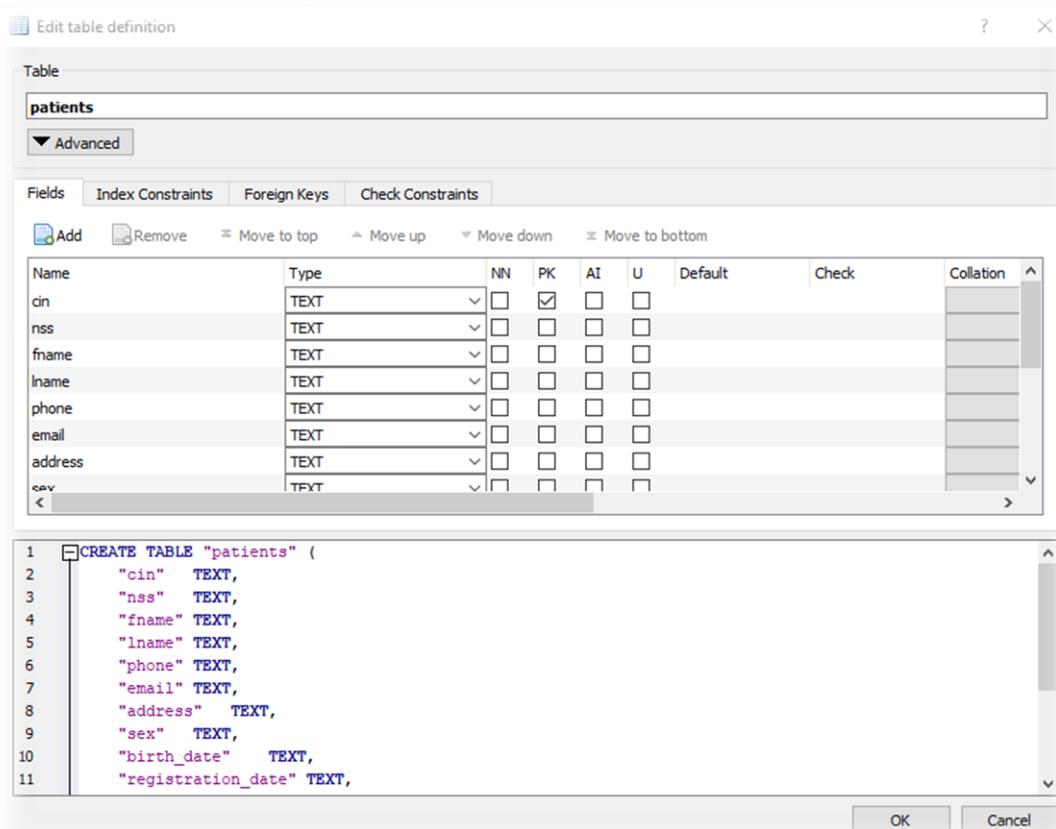


Figure2.2. Table patients

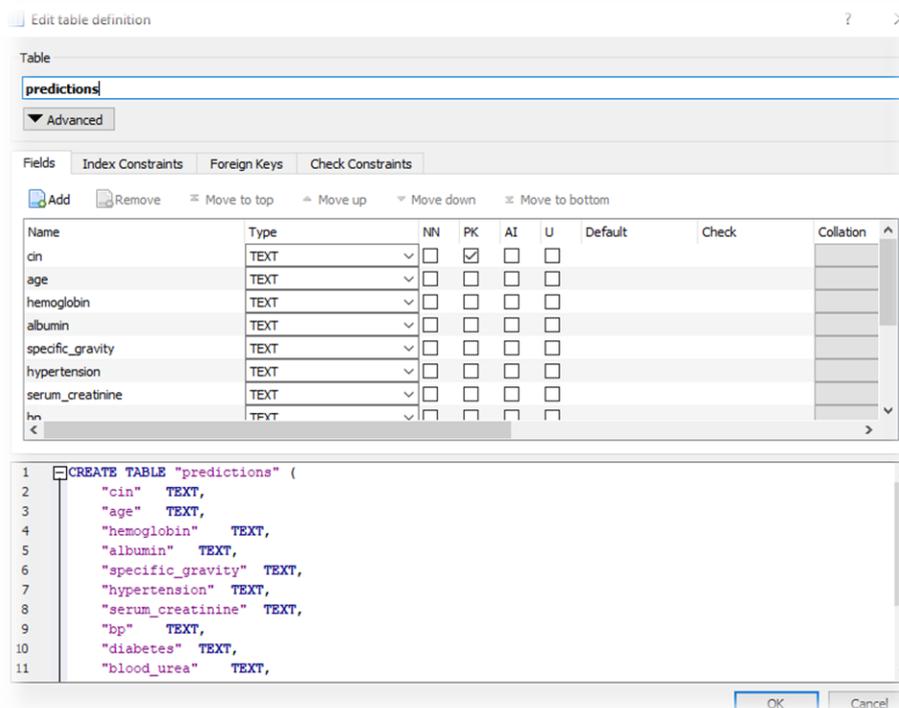


Figure 2.3. Table prediction

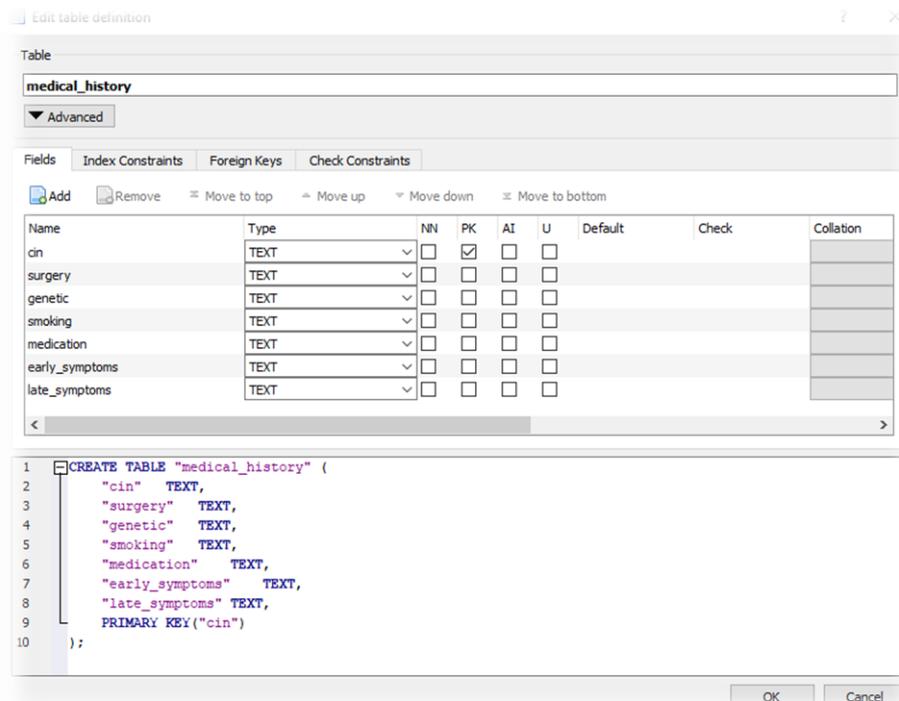


Figure 2.4. Table medical_history

La figure suivante illustre un exemple d'un enregistrement dans la table Patients :

cin	nss	fname	lname	phone	email	address	sex	birth_date	registration_date	height	weight	blood_pressure	illnesses	description
5555 5555 5	8888 8888 8888	sara	sh	0755 5555 55	sara souy ah@gmail.com	said a	Fema le	2002 -05- 25	2025 -05- 25	150	55	120/ 80	Diab etes , Hype rten sion , Hear t dise ase, Othe r	memo ryyy

Figure 2.5. Exemple d'un enregistrement dans la table patients

2.3.4 Algorithme d'adaptation par débit de filtration glomérulaire (GFR)

Après avoir faire des recherches et de plusieurs interviews avec des médecins, nous avons implémenter un algorithme pour classifier les 5 stades d'une maladie rénale.

Algorithme : Calcul d'eGFR et classification du stade CKD

Entrées :

- creatinine (float) : Valeur de créatinine sérique (mg/dL)
- age (int) : Âge du patient (années)
- sex (string) : Sexe du patient ("male" ou "female")

Sorties :

- egfr (float) : Valeur arrondie à 2 décimales (mL/min/1.73m²)
- stage (string) : Stade CKD (S1 à S5)
- icon_path (string) : Chemin de l'icône illustrative
- title/desc (string) : Description clinique du stade

```
// Fonction pour calculer l'eGFR
```

```
calculate_egfr(creatinine : réel, age: entier , is_female: booleen) : réel
```

Début

```
// Constantes CKD-EPI 2021
```

```
k, alpha, sex_factor : réel;
```

```
Si (is_female=Vrai) alors
```

```
  k ← 0.7f;
```

```
  alpha ← -0.241f;
```

```

    sex_factor ← 1.012f;
Sinon // Homme
    k ← 0.9f;
    alpha ← -0.302f;
    sex_factor ← 1.0f;
finsi

// Calcul eGFR
float term1 = pow(fmin(creatinine / k, 1.0f), alpha);
float term2 = pow(fmax(creatinine / k, 1.0f), -1.200f);
float term3 = pow(0.9938f, age);
float egfr = 142.0f * term1 * term2 * term3 * sex_factor;

Retourner roundf(egfr * 100) / 100; // Arrondi à 2 décimales
Fin

// Fonction pour afficher le stade CKD
Procedure print_ckd_stage(egfr : réel) {
    const char *stage, *title, *desc;

Si (egfr >= 90) alors
    stage ← "S1";
    title ← "Normal Kidney Function";
    desc ← "eGFR ≥ 90 mL/min/1.73m2. No kidney damage signs.";
Sinon
Si (egfr >= 60) alors
    stage ← "S2";
    title ← "Mildly Decreased Function";
    desc ← "eGFR 60–89 mL/min/1.73m2. Monitor for kidney damage.";
Sinon
Si (egfr >= 45) alors
    stage ← "S3a";
    title ← "Mild to Moderate Loss";
    desc ← "eGFR 45–59. Moderate risk of complications.";

```

Sinon**Si** (egfr >= 30) **alors**

stage ← "S3b";

title ← "Moderate to Severe Loss";

desc ← "eGFR 30–44. Monitor and manage progression.";

Sinon**Si** (egfr >= 15) **alors**

stage ← "S4"; title = "Severely Decreased Function";

desc ← "eGFR 15–29. Prepare for possible renal replacement.";

Sinon

stage ← "S5"; title = "Kidney Failure";

desc ← "eGFR < 15. Needs dialysis or transplant evaluation.";

finsi**finsi****finsi****finsi****finsi**

Ecrire("--- CKD Stage --- ");

Ecrire("Stage : ", stage, "Title: ", title, " Description : ", desc);

Fin**Début //Programme principal**

Creatinine, egfr : réel;

age : entier;

sex : chaine de caractère ;

is_female : booléen

Ecrire("Enter serum creatinine (mg/dL): ");

Lire(creatinine);

Ecrire ("Enter age: ");

Lire (age);

Ecrire ("Enter sex (male/female): ");

Lire (sex);

// Convertir la saisie en booléen

```
Si (sex= "female") alors
```

```
    is_female ← 0 ;
```

```
    Sinon
```

```
        is_female ← 1 ;
```

```
    fin
```

```
// Calcul et affichage
```

```
egfr ← calculate_egfr(creatinine, age, is_female);
```

```
Ecrire("eGFR: ", egfr, "mL/min/1.73m²");
```

```
print_ckd_stage(egfr);
```

```
FIN
```

2.4 Intégration des plans de surveillance pour un patient

L'intégration des plans de surveillance dans un système d'aide à la décision médicale constitue une étape essentielle pour garantir un **suivi personnalisé, standardisé et évolutif** des patients atteints de **maladie rénale chronique (MRC)**. Ce processus vise à assurer une continuité des soins, à anticiper la progression de la maladie et à améliorer la qualité de vie du patient grâce à des interventions précoces et ciblées.

Conformément aux recommandations des lignes directrices internationales (KDIGO 2022), le système développé distingue clairement les modalités de surveillance en fonction des stades de la MRC, allant du stade 1 au stade 5. Cette segmentation permet d'adapter les fréquences d'examen, les mesures cliniques à collecter, les interventions à prévoir, et les ressources médicales à mobiliser.

2.4.1 Stades 1 et 2 :

Les patients à ces stades précoces bénéficient d'un suivi axé sur la prévention des facteurs aggravants. L'interface intègre un plan de contrôle de la tension artérielle et de la glycémie sous forme de **tableau de suivi structuré**, avec saisie des valeurs mesurées, des dates et des remarques médicales. Un volet **conseils hygiéno-diététiques** permet au médecin ou au personnel infirmier d'accompagner le patient dans ses habitudes de vie (alimentation, activité

physique, arrêt du tabac, etc.). En complément, l'interface intègre un **plan de surveillance médicale** (monitoring plan), conforme aux recommandations des sociétés néphrologiques. Enfin, une **fonctionnalité d'impression** est proposée, permettant de générer une **version PDF du plan de suivi** personnalisé. Cela offre une **traçabilité médicale**, facilite la communication avec d'autres professionnels de santé et permet au patient d'avoir un **support écrit** de son plan de prise en charge [39].

2.4.2 Stade 3 (a et b)

Le **stade 3 de la maladie rénale chronique (MRC)** marque une phase charnière dans l'évolution de la pathologie, nécessitant un suivi plus rigoureux et une évaluation multidimensionnelle du patient. **Le stade 3 est subdivisé en 3a et 3b, avec des fréquences de suivi ajustées : tous les 6 mois pour le stade 3a, et tous les 3 à 4 mois pour le stade 3b (respectivement)**. Cette distinction permet une meilleure adaptation des interventions cliniques en fonction du niveau de dégradation de la fonction rénale. Un volet **plan de surveillance structuré** est proposé pour assurer un suivi régulier de la fonction rénale et des marqueurs biologiques associés. Ce plan est présenté à l'écran et peut être imprimé automatiquement sous forme de **liste d'analyses cliniques**, facilitant la prescription et la transmission au patient ou au laboratoire. L'interface permet au médecin d'effectuer une **revue systématique des traitements en cours**, dans une logique de **sécurité thérapeutique**. Enfin, un volet spécifique est dédié aux **conseils hygiéno-diététiques** et à **l'éducation thérapeutique du patient** [40].

2.4.3 Stades 4 et 5

Les **stades 4 et 5** de la maladie rénale chronique correspondent aux phases les plus avancées de l'insuffisance rénale, marquées par une altération sévère du **débit de filtration glomérulaire**. À ce stade, une **prise en charge spécialisée, pluridisciplinaire et anticipative** devient impérative pour ralentir la progression vers l'insuffisance rénale terminale et préparer la suppléance rénale.

Un **plan de surveillance biologique** comprenant les examens rénaux, les électrolytes, le statut calcique, conformément aux recommandations KDIGO.

Une **revue médicamenteuse** approfondie, permettant de détecter les traitements inadaptés, **d'éviter les médicaments néphrotoxiques**, et **d'ajuster les posologies** en fonction du DFG.

Les patients à ces stades nécessitent une **prise en charge nutritionnelle personnalisée**, gérée en collaboration avec un diététicien. Des conseils de **style de vie** (activité physique douce, sommeil, arrêt du tabac) sont également intégrés pour optimiser l'état général du patient.

La planification de la dialyse constitue une section spécifique du système, conçue pour accompagner le professionnel de santé dans la préparation du traitement de suppléance. Elle permet de définir le **type de dialyse approprié**, d'organiser les **séances selon un planning hebdomadaire**, de sélectionner le **centre de dialyse adapté**, et d'assurer un **suivi structuré des symptômes post-dialyse** afin d'ajuster la prise en charge de manière individualisée et proactive.

Enfin, l'interface intègre également un rappel pour référer le patient à un psychologue, dans une démarche d'accompagnement psychologique essentielle à ce stade. Le suivi du statut de transplantation rénale, avec les options : candidat, en liste d'attente, opération planifiée, transplanté, ou non éligible, permet de mieux suivre l'évolution vers une greffe éventuelle. Toutes les informations saisies sont automatiquement centralisées dans un tableau récapitulatif, assurant une gestion dynamique et sécurisée des informations médicales [41].

2.5 Description de l'UML

2.5.1 Définition et historique de l'UML

UML (unified modeling language) est un langage unifié pour la modélisation dans le cadre de la conception orientée objet. Il s'agit d'un langage graphique de modélisation objet permettant de spécifier, de construire, de visualiser et de décrire les détails d'un système logiciel. L'UML est né de la fusion des trois méthodes qui ont le plus influencé la modélisation objet au milieu des années 90 : OMT, Booch et OOSE. Principalement issu des travaux de Grady Booch, James Rumbaugh et d'Ivar Jacobson.

UML est devenu une norme officielle de l'**OMG (Object Management Group)** en **1997**. L'OMG, fondé en **1989** à l'initiative de grandes entreprises technologiques telles que **HP, IBM**, etc., est chargé de la standardisation des langages de modélisation. La version actuelle de la norme UML est la **2.5.1**, publiée en **décembre 2017**. [42].

2.5.2 Les diagrammes d'UML

Les diagrammes sont des éléments graphiques. Ceux-ci décrivent le contenu des vues qui sont des notions abstraites. Ils sont répartis selon trois points de vue (fonctionnelle, statique, dynamique), comme le montre la figure suivante :

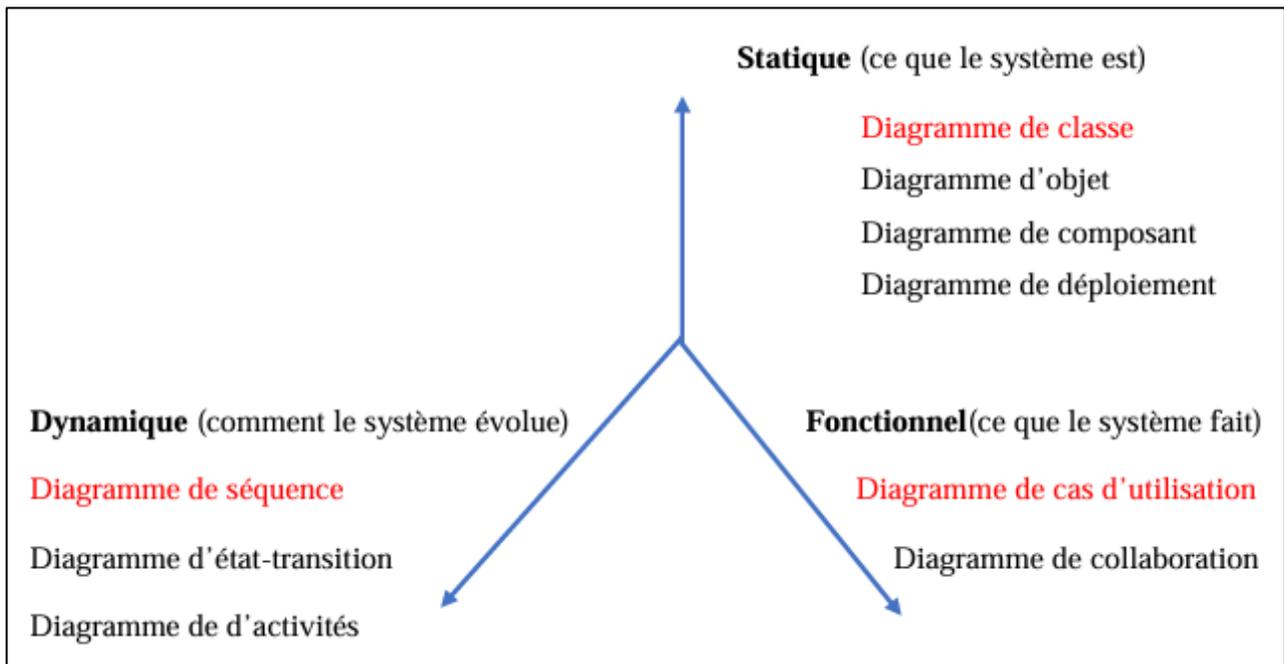


Figure 2.6 Les axes de modélisation

Comme le montre la figure (04) chaque vue de système est modélisée par des diagrammes.

- **Les diagrammes fonctionnels**

Un diagramme fonctionnel permet de rendre compte des interactions entre des utilisateurs et le système. Ce diagramme doit comporter tous les traitements métier (fonctionnels). On retrouve : Le diagramme de cas d'utilisation : représente un ensemble de séquence d'action qui sont réalisées par le système et qui produise un résultat observable intéressant pour un acteur particulier.

- **Les diagrammes dynamiques**

Un diagramme dynamique permet de montrer le comportement du système. Les interactions des objets et leur évolution dans le temps. L'accent est mis sur la chronologie des interactions. On trouve :

Le diagramme de séquence : représente la vue dynamique du fonctionnement de l'application. il fait apparaître les interactions entre les acteurs et l'application ; Les descriptions de l'interaction (par l'envoi de messages) ; Les interactions entre les intervenants ; La description de l'ordre des interactions entre les objets qui compose le système.

- **Les diagrammes statiques**

Un diagramme statique permet de décrire la structure du système en termes de Composant du système : objet, classe, paquetage, composants et de relation entre ces composants : Spécialisation, Association, dépendance, etc. On trouve :

Les diagrammes de classes : est un diagramme UML statique qui représente la structure d'un système logiciel. Il montre les **classes**, leurs **attributs**, leurs **méthodes**, ainsi que les **relations** entre elles (comme l'héritage ou l'association). Il est utilisé pour modéliser les entités d'un système orienté objet, faciliter la conception, et préparer la génération de code.

2.6 Représentation graphique des diagrammes UML

Nous présentons ci-dessus les diagrammes de cas d'utilisation pour chaque acteur.

Un administrateur est représenté par :

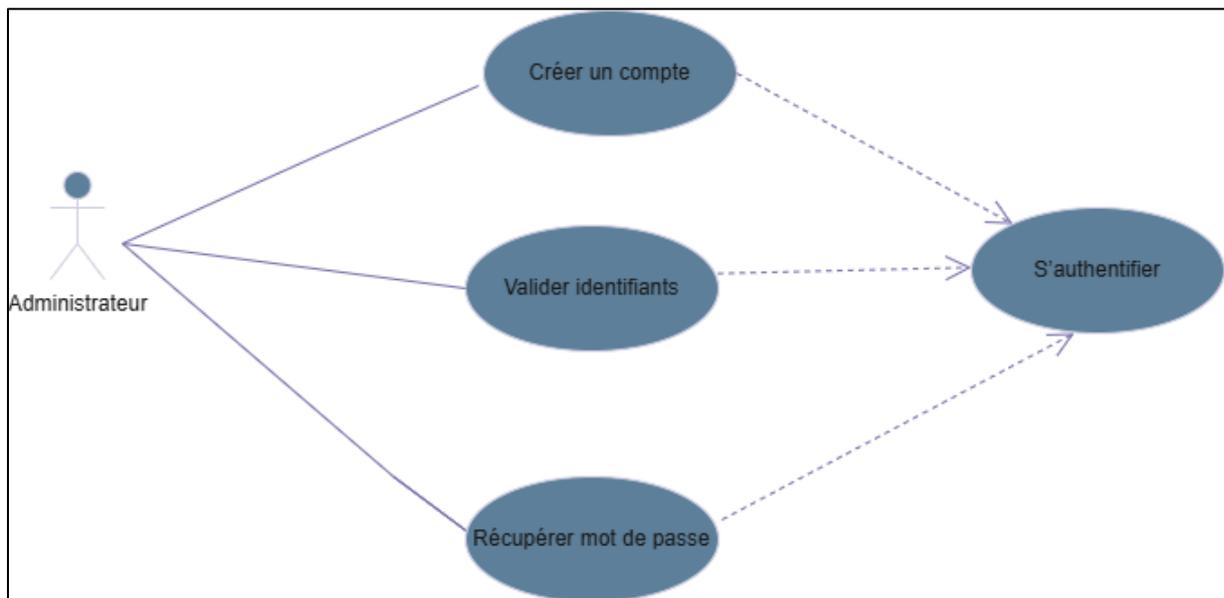


Figure 2.7 Diagramme de cas d'utilisation pour l'administrateur.

L'infirmière est représentée par :

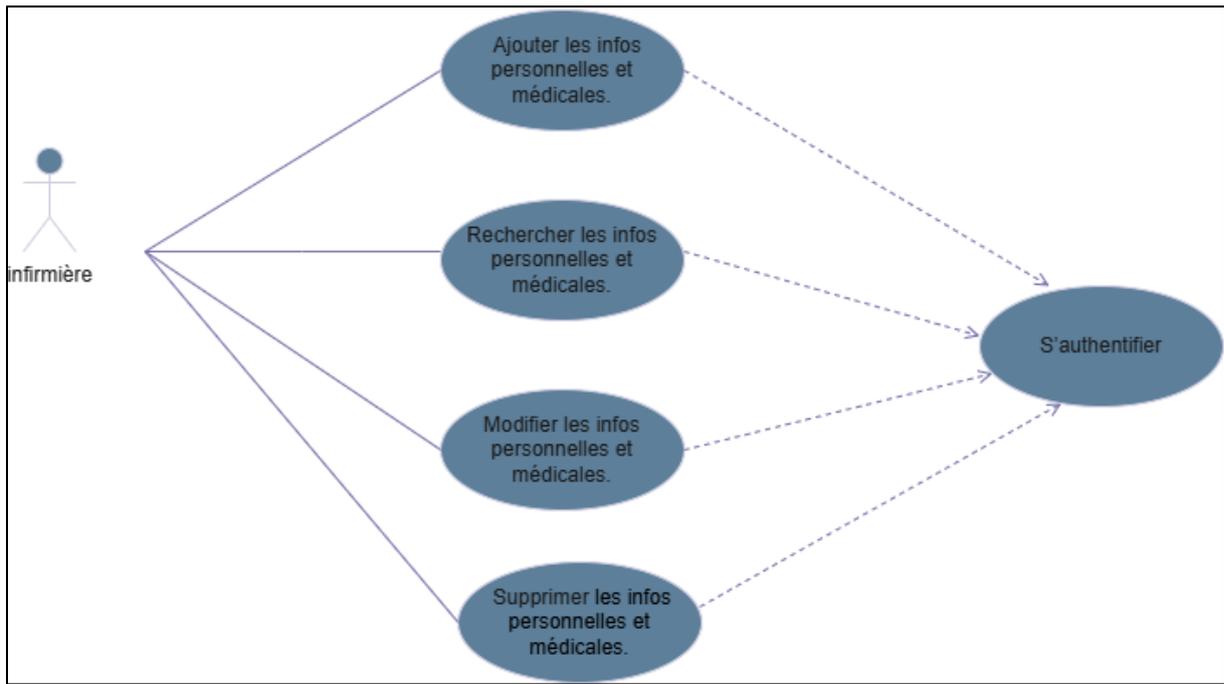


Figure 2.8 Diagramme de cas d'utilisation pour l'infirmière.

Le médecin est représenté par :

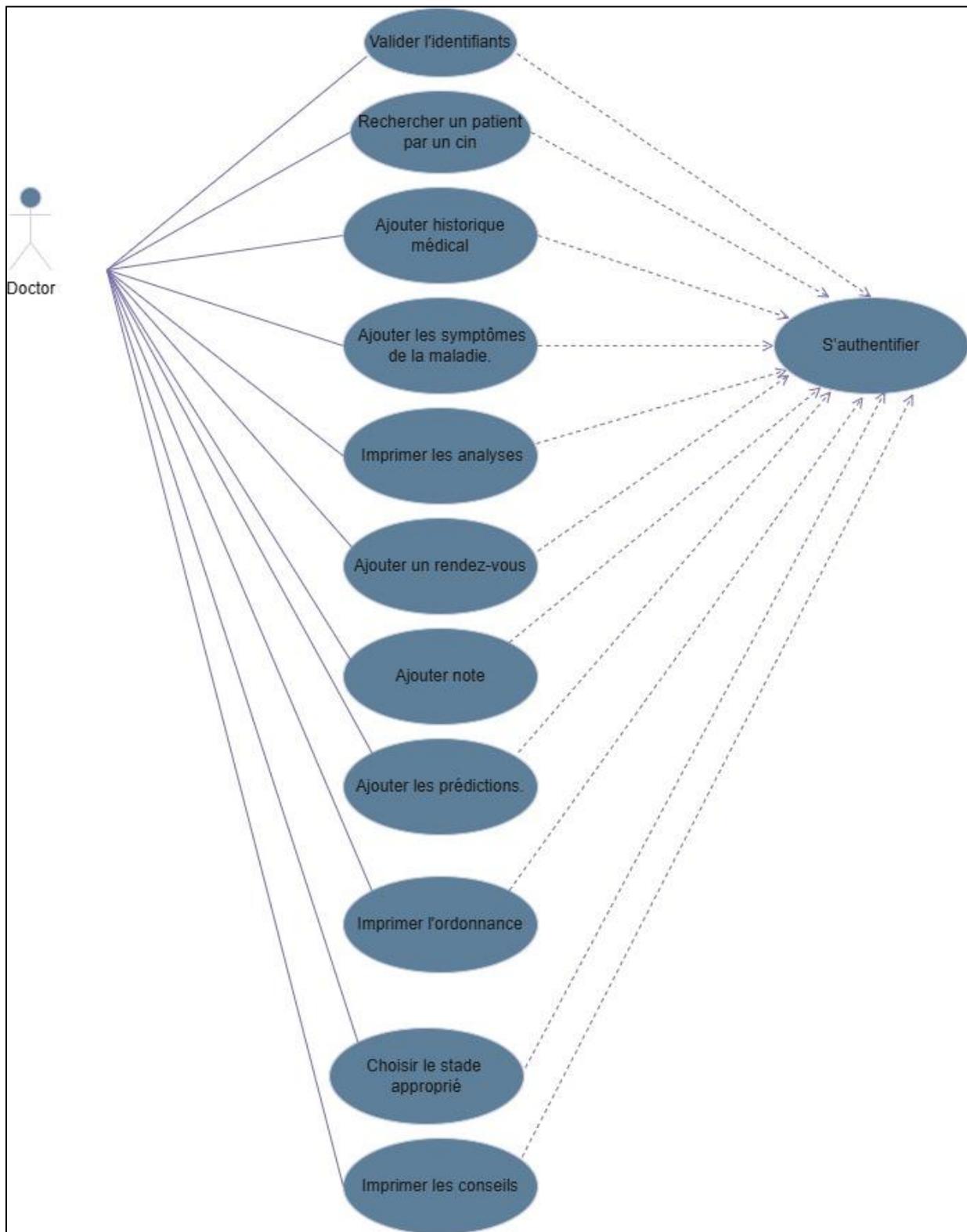


Figure 2.9 Diagramme de cas d'utilisation pour le médecin.

Nous allons présenter ci-après, les diagrammes de séquences suivants :

Un administrateur est représenté par :

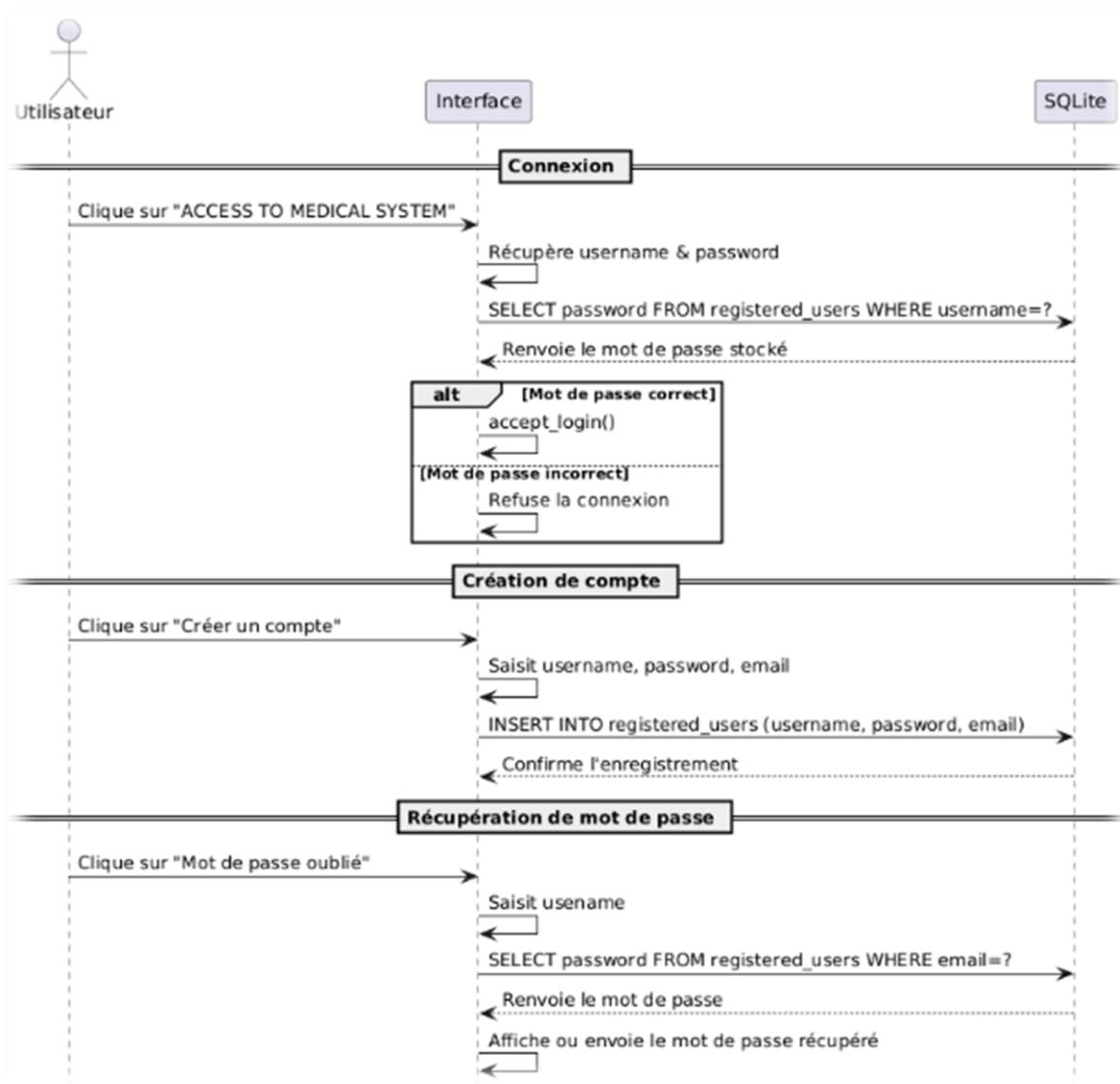


Figure 2.10 : diagramme de séquence pour l'administrateur.

L'infirmière est représentée par :

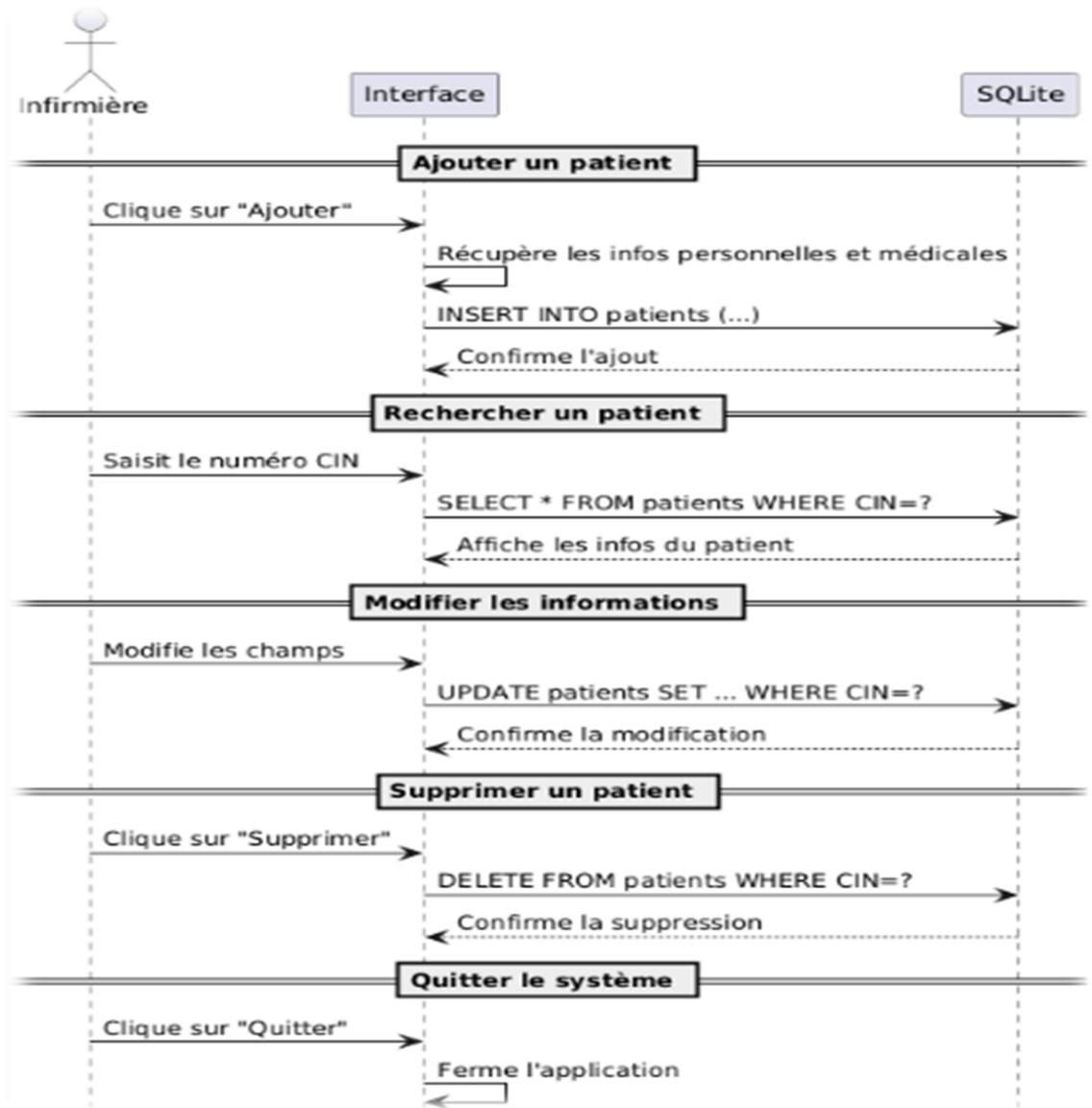


Figure 2.11 Diagramme de séquence pour l’infirmière.

Le médecin est représenté par :

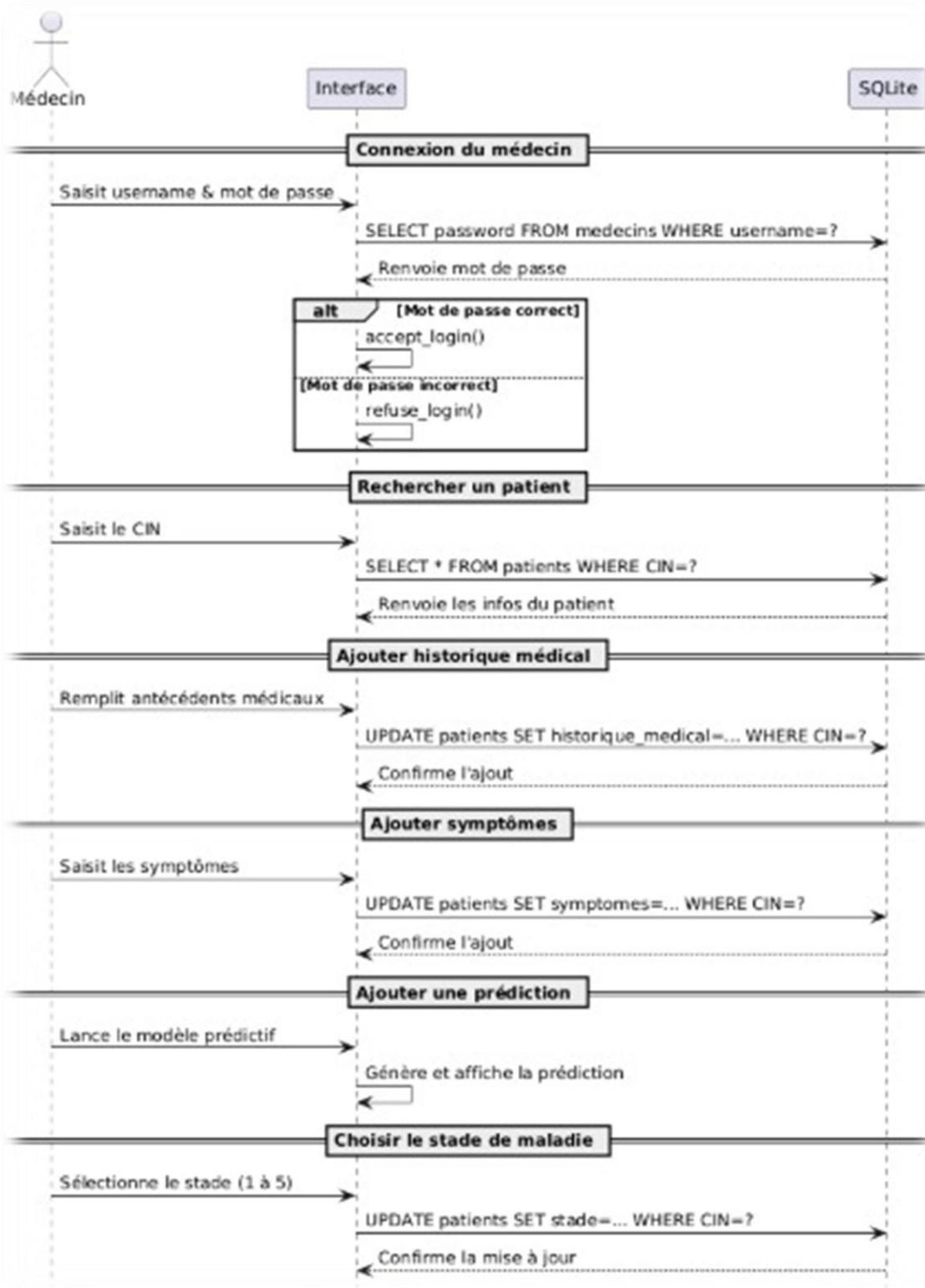


Figure 2.12 Diagramme de séquence pour le médecin.

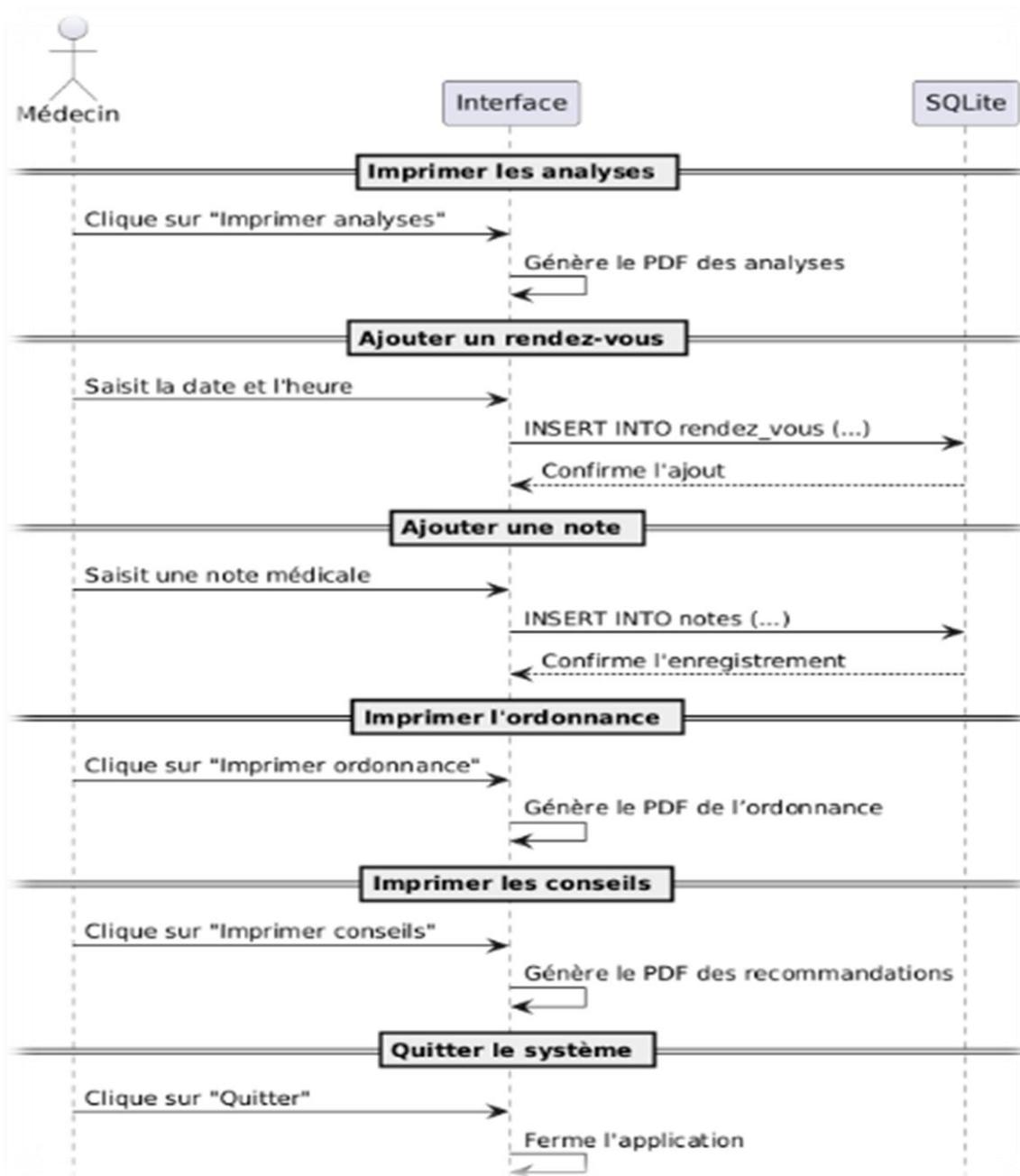


Figure 2.13 Diagramme de séquence pour le médecin (la suite).

Nous présentons ci-dessus le diagramme de classe.

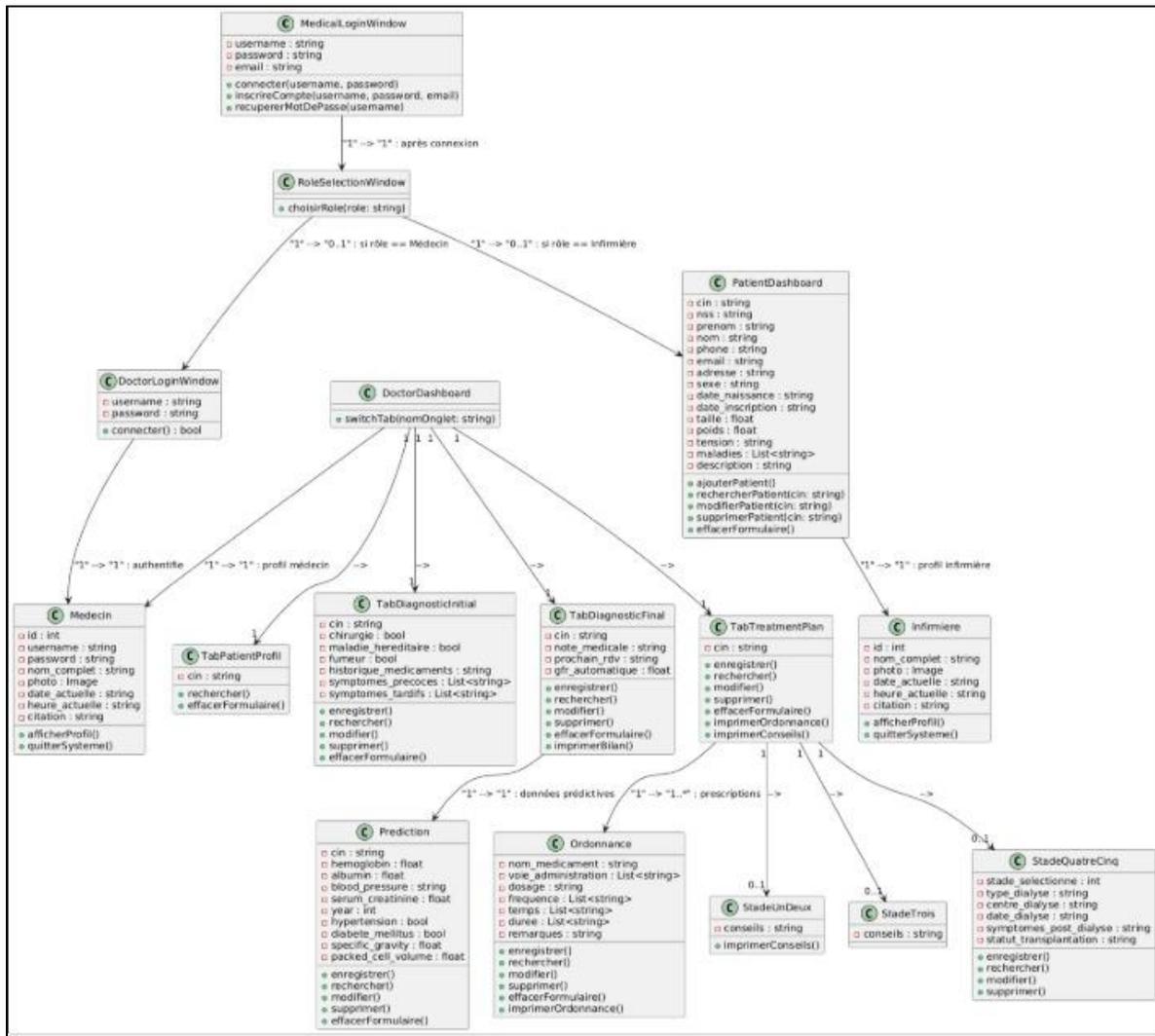


Figure 2.14 Diagramme de classe.

2.7 Conclusion

Ce chapitre a présenté une étude approfondie sur le diagnostic des maladies rénales, en mettant l'accent sur une architecture intelligente adaptée au contexte clinique. Après avoir défini les maladies rénales et leurs classifications, ainsi que les symptômes fréquents, nous avons introduit une architecture décisionnelle basée sur le Débit de Filtration Glomérulaire (DFG), un indicateur essentiel de la fonction rénale.

L'architecture proposée s'appuie sur deux composantes principales : une base de cas regroupant des profils types de patients et leurs diagnostics associés, et une base de données réelle de patients alimentée par des données cliniques. L'algorithme d'adaptation par DFG permet une personnalisation du diagnostic, en ajustant les décisions à partir du niveau de

fonction rénale mesuré. Ce processus repose sur des méthodes de raisonnement à base de cas (Case-Based Reasoning), assurant ainsi une prise de décision progressive, transparente et réutilisable.

Enfin, l'intégration de plans de surveillance adaptés au stade de la maladie garantit un suivi médical structuré, notamment pour les stades avancés (4 et 5) nécessitant une dialyse ou une transplantation. L'ensemble est complété par une série de diagrammes UML, représentant les composants logiciels, les interactions et les flux d'information du système proposé.

Ainsi, cette architecture représente une avancée dans le développement d'un système d'aide à la décision médical (SADM), combinant intelligence artificielle, médecine personnalisée et outils de modélisation pour améliorer le diagnostic et la gestion des maladies rénales chroniques.

Chapitre 3 : L'implémentation de notre logiciel ClinicPro

3.1. Introduction

Ce chapitre présente les outils, techniques et technologies utilisés pour la conception et le développement de notre application. Il couvre la création et l'implémentation des différents programmes, interfaces graphiques et bases de données nécessaires au bon fonctionnement du système. Nous commencerons par décrire l'environnement de développement choisi ainsi que les outils logiciels utilisés pour la mise en place du système et de la base de données. Ensuite, nous détaillerons quelques interfaces clés de l'application, illustrant leur rôle et leur interaction avec les différentes fonctionnalités implémentées.

3.2. Description de l'environnement de travail

3.2.1 Description du matériel utilisé

Afin de réaliser cette application dans les conditions les plus favorables, nous avons mis en disposition un micro-ordinateur portable ayant les configurations suivantes :

- **Processeur** : Intel® Core™ i3-7020U @ 2.30 GHz;
- Fréquence d'horloge : 2,40 GHz;
- RAM : 4Go ;
- Système d'exploitation : Windows 10 Pro ;

3.2.2 Les langages utilisés dans l'application

Le développement de notre système d'aide à la décision pour le diagnostic des maladies rénales repose sur un ensemble de langages et technologies modernes, choisis pour leur adéquation avec les exigences fonctionnelles, l'implémentation d'algorithmes d'intelligence artificielle, ainsi que la conception d'interfaces graphiques intuitives :

3.2.2.1 Python

Python est un **langage de programmation interprété, orienté objet, de haut niveau** et à typage dynamique. Il a été conçu avec une **syntaxe claire et lisible**, ce qui facilite la maintenance des programmes et améliore la productivité des développeurs. Python prend en charge plusieurs paradigmes de programmation, notamment :

- **La programmation impérative ;**
- **La programmation orientée objet (POO) ;**
- **La programmation fonctionnelle ;**

Python est doté d'une **gestion dynamique de la mémoire**, et d'un **large écosystème de bibliothèques** pour la science des données, l'intelligence artificielle, les interfaces graphiques, etc. [43] [44]

Le langage est notamment utilisé dans les domaines de **l'intelligence artificielle**, du **traitement de données médicales**, et du **développement rapide d'applications** grâce à des bibliothèques puissantes telles que :

- **NumPy** (calcul numérique) ;
- **Pandas** (analyse de données) ;
- **Scikit-learn** (machine learning) ;
- **Tkinter / PyQt / Kivy** (interfaces graphiques) ;
- **Matplotlib / Seaborn** (visualisation) ;

Dans notre application, nous avons utilisé les bibliothèques suivantes :

- **PyQt6** :

PyQt6 est un **binding officiel en Python** de la bibliothèque **Qt 6** développée en C++. Elle permet de créer des **interfaces graphiques (GUI)** riches et modernes à l'aide du langage Python. PyQt6 est maintenu par *Riverbank Computing*. [45]

Ce Framework intègre :

- **Widgets standards (QPushButton, QLabel, etc.) ;**
- **Système de signal/slot** pour la gestion des événements ;
- **Gestion avancée des layouts** (QGridLayout, QVBoxLayout...) ;
- **Support multi-plateforme** (Windows, Linux, macOS) ;

Il est particulièrement adapté aux applications professionnelles nécessitant une interface graphique robuste, comme les logiciels médicaux ou d'analyse de données cliniques.

- **SQLite :**

SQLite est un **moteur de base de données relationnelle embarqué et sans serveur**. Contrairement aux SGBDR classiques (comme MySQL), il ne nécessite aucune configuration ni processus serveur. Les bases sont stockées dans un **fichier unique** sur disque. [46]

Caractéristiques :

- Langage **SQL-92 compatible** ;
- Très **léger**, rapide et fiable ;
- Idéal pour des **applications locales ou embarquées** ;
- Parfaitement intégré à Python via le module `sqlite3` ;

SQLite est largement utilisé dans les applications mobiles, les navigateurs web, et les systèmes médicaux locaux pour gérer des **bases de données patient légères mais sécurisées**.

3.2.2.2 HTML & CSS pour la génération des rapports médicaux

Bien que l'application soit une solution desktop développée avec **Python** et **PyQt6**, l'utilisation de **HTML** et **CSS** a été intégrée pour la **génération et la stylisation de rapports médicaux**. Cette approche permet de produire des **rapports professionnels**, imprimables, et **visuellement structurés**, notamment lors de l'exportation en PDF ou pour l'affichage dans des composants tels que `QTextDocument`. [47]

- **HTML** a servi à structurer les données médicales (informations patient, résultats d'analyse, diagnostic initial) sous forme de tableaux, sections, et titres.
- **CSS** a permis d'appliquer une mise en forme claire et professionnelle : couleurs médicales (bleu/vert), bordures, alignement typographique, espacement entre les éléments, et styles conditionnels (ex. couleur rouge pour un diagnostic critique).

Cette séparation du fond (HTML) et de la forme (CSS) suit les **bonnes pratiques de présentation documentaire** en informatique médicale.

3.3 Environnement de développement :

3.3.1 Visual Studio Code

Visual Studio Code (VS Code) est un **éditeur de code source moderne, léger et extensible**, développé par **Microsoft**. Il est largement utilisé dans le développement logiciel, notamment dans les domaines de l'intelligence artificielle, des applications desktop, et du traitement de données médicales, grâce à sa compatibilité avec un grand nombre de langages et frameworks. [48]

Caractéristiques techniques clés :

- **Support natif de Python** via l'extension officielle *Python* (IntelliSense, Pylint, Jupyter support, etc.) ;
- **Terminal intégré** pour exécuter des scripts directement depuis l'IDE ;
- **Debugging interactif** avec points d'arrêt, visualisation de variables, et step-by-step execution ;
- **Contrôle de version Git intégré** ;
- **Extensions dédiées** : scikit-learn, SQLite, Jupyter Notebooks, PyQt Designer Viewer ;
- **Cross-platform** : compatible avec Windows, macOS, et Linux ;

VS Code permet également d'intégrer des **environnements virtuels Python** facilitant la gestion des dépendances, ainsi que des **plug-ins de visualisation** pour les modèles de machine learning ou bases de données.

3.3.2 BD Browser for SQLite

DB Browser for SQLite est un outil graphique libre (open-source) permettant de **créer, concevoir, visualiser, modifier et interroger** des bases de données SQLite de manière intuitive, sans écrire de requêtes SQL manuellement. [49]

Il facilite :

- La **création de tables** et la définition de schémas relationnels ;
- L'**insertion, édition et suppression** de données via des formulaires ;
- L'exécution et le **débugage de requêtes SQL** personnalisées ;
- L'**exportation/importation** des données (CSV, SQL, etc.) ;

- Le **suivi du contenu** d'une base patient en contexte médical.

Dans le cadre de ce projet, **DB Browser** a été utilisé pour **concevoir le schéma de base de données, inspecter les tables patient, et vérifier l'intégrité des données** lors des tests fonctionnels.

3.4 Intégration du Machine Learning dans un Système d'Aide à la Décision Médicale (SADM)

3.4.1 Contexte

L'intégration du **Machine Learning (ML)** dans un **Système d'Aide à la Décision Médicale (SADM)** a pour but de fournir aux professionnels de santé un outil intelligent capable d'assister dans l'identification rapide et fiable de maladies, à partir de données cliniques et biologiques du patient. Dans notre cas, l'objectif est de détecter **la présence ou l'absence d'une maladie rénale chronique (CKD)**.

L'intégration du **Machine Learning** (apprentissage automatique) dans un SADM vise à :

- **assister le médecin** dans le diagnostic,
- **améliorer la précision** des décisions,
- **réduire les erreurs humaines**,
- et **accélérer la prise en charge** du patient.

3.4.2 Définition Technique du Machine Learning

Le Machine Learning (ML) ou apprentissage automatique est un sous-domaine de l'intelligence artificielle (IA) qui permet à une machine d'apprendre automatiquement à partir des données, sans être explicitement programmée. L'objectif principal est de développer des modèles prédictifs ou classificateurs capables de détecter des schémas ou relations cachées dans des jeux de données. Selon Mitchell (1997), *"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."* [50].

Dans un contexte médical, le ML peut être utilisé pour classer un patient comme atteint ou non atteint d'une maladie en se basant sur des données cliniques et biologiques.

3.4.3 Type d'Apprentissage Utilisé : Supervised Learning

Dans ce projet, nous utilisons l'**apprentissage supervisé (Supervised Learning)**. Ce type d'apprentissage repose sur un **jeu de données étiquetées**, c'est-à-dire que chaque exemple contient des caractéristiques (features) ainsi qu'une **étiquette de sortie (label)** connue — dans notre cas, la présence ou non d'une maladie rénale chronique (CKD: Chronic Kidney Disease).[51]

L'objectif est d'entraîner un modèle à **prédire cette étiquette** à partir de nouvelles observations.

3.4.4 Modèle Choisi : Arbre de Décision (Decision Tree Classifier)

Un arbre de décision est un algorithme d'apprentissage supervisé non paramétrique utilisé pour résoudre des problèmes de classification et de régression. Le modèle construit une structure arborescente où chaque nœud interne représente un test (décision) sur une ou plusieurs caractéristiques, chaque branche représente le résultat du test, et chaque feuille correspond à une prédiction de classe. [52]

Ce modèle est très apprécié dans le domaine médical pour sa capacité à fournir des règles décisionnelles claires, facilement interprétables par des praticiens.

Voici la structure générale d'un arbre de décision :

- Nœud racine : Le point de départ de l'arbre.
- Nœuds internes : Les points de décision intermédiaires entre le nœud racine et les nœuds feuilles.
- Nœuds feuilles : Les points terminaux qui représentent la classification finale ou la prédiction.
- Branches : Les connexions entre les nœuds, représentant les résultats des tests effectués à chaque nœud.

Pour chaque nœud dans un arbre de d'excision, les critères de division sont utilisés pour d'déterminer comment séparer les données en fonction des différents attributs. Les principaux critères sont :

- Gain d'information
- Indice de Gini (Gini Index)
- Ratio de gain (Gain Ratio)

3.5 Architecture Générale du SADM Intégré au ML

L'implémentation d'un SADM fondé sur le ML repose sur une architecture modulaire et hiérarchisée, structurée autour des étapes suivantes :

3.5.1 Acquisition et structuration des données :

Dans cette étude, nous avons utilisé le jeu de données relatif à la **maladie rénale chronique (Chronic Kidney Disease - CKD)**, disponible publiquement sur le dépôt de données de kaggle [53]. Il contient **401 observations** décrivant l'état de santé de patients, réparties sur **25 attributs médicaux** incluant des variables biologiques, cliniques, et démographiques. La variable cible classification indique si un patient est atteint de CKD (ckd) ou non (notckd) dans la figure 3.1. [54]

```

=== FIRST 10 ROWS OF CLEANED DATASET ===
  age_yrs  blood_pressure_mm/hg  specific_gravity  albumin  ...  appetite:poor  pedal_edema:yes  anemia:yes  ckd:yes
0    48.0             80           1.020    1.0  ...           0              0              0           1
1     7.0             50           1.020    4.0  ...           0              0              0           1
2    62.0             80           1.010    2.0  ...           1              0              1           1
3    48.0             70           1.005    4.0  ...           1              1              1           1
4    51.0             80           1.010    2.0  ...           0              0              0           1
5    60.0             90           1.015    3.0  ...           0              1              0           1
6    68.0             70           1.010    0.0  ...           0              0              0           1
7    24.0             60           1.015    2.0  ...           0              1              0           1
8    52.0            100           1.015    3.0  ...           0              0              1           1
9    53.0             90           1.020    2.0  ...           1              0              1           1

```

Figure 3.1 Aperçu de l'ensemble de données CKD_Preprocessed.

3.5.2 Prétraitement des données :

Les données brutes doivent être préparées pour être exploitables par les algorithmes de ML. [55]. Cela comprend :

1. **Nettoyage** : suppression des doublons, gestion des valeurs manquantes (imputation ou suppression) avec `dropna ()`.
2. **Encodage** : conversion des variables catégorielles (sexe, maladie chronique oui/non) en variables numériques avec `pd.to_numeric ()`.

3. **Normalisation/Standardisation** : pour homogénéiser les échelles des variables (notamment utile pour les modèles sensibles à l'échelle comme KNN ou SVM).

3.5.3 Sélection des Caractéristiques (Feature Selection) :

La sélection des caractéristiques vise à **réduire la dimensionnalité** du jeu de données en gardant les variables les plus pertinentes pour le modèle prédictif. [56]. Cela :

1. Améliore la précision du modèle,
2. Réduit le sur-apprentissage (overfitting),
3. Facilite l'interprétation par les médecins.

Dans notre approche, nous avons utilisé la méthode de l'importance de Gini, c'est une méthode intégrée dans les **arbres de décision** (ou forêts aléatoires). Elle mesure à quel point chaque variable réduit l'impureté (indice de Gini) lorsqu'elle est utilisée pour un split.

Parmi les 25 variables disponibles, dix ont été identifiées comme particulièrement discriminantes pour le diagnostic des maladies rénales chroniques. Cette sélection repose sur leur **pertinence clinique** et leur **poids statistique** dans les études antérieures. [57]

- **Age (yrs)** : facteur de risque naturel lié au vieillissement rénal.
- **Hemoglobin (g/dL)** : Évalue la capacité du sang à transporter l'oxygène.
- **Specific Gravity** : mesure la capacité des reins à concentrer l'urine.
- **Hypertension** : pression artérielle élevée pouvant endommager les glomérules rénaux.
- **Diabetes Mellitus**: cause majeure de néphropathie chronique.
- **Albumin (g/dL)** : présence d'albumine dans les urines, marqueur d'atteinte glomérulaire.
- **Blood Pressure (mmHg)** : reflet direct du stress vasculaire sur les reins.
- **Blood Urea (mg/dL)**: indicateur de l'accumulation de déchets azotés.
- **Serum Creatinine (mg/dL)** : principal biomarqueur de la fonction rénale.
- **Packed Cell Volume(%)** : Indique la concentration des globules rouges dans le sang.

La **variable cible** est ckd:_yes, indiquant la présence ou non de la maladie

3.5.4 Division des données en ensembles d'entraînement et de test

Afin d'évaluer objectivement la performance du modèle d'apprentissage automatique, l'ensemble des données a été divisé en deux sous-ensembles distincts : un ensemble d'entraînement (**train**) et un ensemble de test (**test**). Cette procédure permet de simuler des données jamais vues par le modèle, garantissant ainsi une évaluation réaliste de sa capacité de généralisation. [58]

Dans notre cas, une division aléatoire de **60 %** pour l'entraînement et **40 %** pour le test a été réalisée à l'aide de la fonction **train_test_split** de la bibliothèque *Scikit-learn*.

Cette répartition a été accompagnée de l'option **stratify=y**, ce qui signifie que la proportion des classes (par exemple : patients atteints ou non atteints de maladie rénale) est conservée dans les deux sous-ensembles. Cela est crucial en contexte médical, où les jeux de données sont souvent déséquilibrés. [59]

De plus, le paramètre **random_state=42** a été défini afin d'assurer la reproductibilité des résultats. En fixant une graine aléatoire, les mêmes échantillons seront sélectionnés à chaque exécution du code, ce qui est essentiel pour valider scientifiquement les performances du modèle.

Ce processus de séparation garantit que l'évaluation du modèle repose sur des observations indépendantes de celles utilisées lors de l'apprentissage, conformément aux bonnes pratiques de validation en science des données. [60]

3.5.5 Entraînement du modèle DecisionTreeClassifier

L'entraînement du modèle consiste à ajuster les paramètres internes d'un algorithme d'apprentissage supervisé à partir d'un ensemble de données étiquetées. Dans ce travail, un arbre de décision a été utilisé comme classificateur, à l'aide de la classe *DecisionTreeClassifier* de la bibliothèque Python *Scikit-learn* [61].

Les arbres de décision sont des modèles d'apprentissage non linéaire qui apprennent une série de règles de décision hiérarchiques à partir des attributs d'entrée, divisant l'espace des données de manière récursive pour maximiser un critère de pureté [62]. Dans ce projet, le critère utilisé est l'**entropie**, une mesure de l'impureté informationnelle issue de la théorie de

l'information de Shannon, qui guide le modèle à choisir les attributs les plus discriminants à chaque nœud [63].

Le modèle a été entraîné sur l'ensemble d'apprentissage à l'aide des **hyperparamètres** suivants :

- **criterion='entropy'** : Utilisation de l'entropie comme fonction de gain d'information pour la division des nœuds.

$$Entropie = - \sum p_i \log_2 (p_i)$$

- **max_depth=5** : Limitation de la profondeur de l'arbre à 5 niveaux pour éviter le surapprentissage (*overfitting*).
- **min_samples_split=10** : Un nœud n'est divisé que s'il contient au moins 10 exemples, afin d'assurer la robustesse statistique des règles.
- **min_samples_leaf=5** : Chaque feuille de l'arbre doit contenir au minimum 5 échantillons, ce qui réduit la variance.
- **max_features='sqrt'** : Le nombre de caractéristiques testées à chaque division est limité à la racine carrée du total, une approche inspirée des forêts aléatoires pour réduire la corrélation entre nœuds.
- **class_weight='balanced'** : Cette option permet de compenser automatiquement les déséquilibres entre classes en ajustant les poids attribués aux échantillons.
- **random_state=42** : Pour garantir la reproductibilité des résultats, une graine fixe est définie.

Ce paramétrage permet d'obtenir un compromis acceptable entre biais et variance, et est adapté à des contextes sensibles comme le diagnostic médical assisté par intelligence artificielle, où l'interprétabilité et la stabilité sont cruciales [64].

3.5.6 Evaluation du modèle

Après l'entraînement du modèle `DecisionTreeClassifier`, une phase d'évaluation rigoureuse est nécessaire afin d'estimer sa capacité à généraliser sur des données inédites. Cette étape repose sur l'application du modèle à un ensemble de test, distinct des données d'entraînement, et sur l'analyse de plusieurs métriques de performance.

3.5.6.1 Objectif de l'évaluation

L'évaluation permet de vérifier si le modèle est capable de fournir des prédictions fiables sur des données qu'il n'a jamais vues. Elle sert également à détecter des problèmes courants comme le **surapprentissage** (*overfitting*) ou le **sous-apprentissage** (*underfitting*), et à comparer les performances avec d'autres algorithmes ou configurations.

3.5.6.2 Méthodologie

Le jeu de données a été divisé en deux sous-ensembles : **60 % pour l'entraînement** et **40% pour le test**, en utilisant une division stratifiée afin de préserver la proportion des classes (CKD / Non-CKD). L'évaluation a été réalisée à l'aide des fonctions de la bibliothèque scikit-learn [65], notamment `accuracy_score`, `classification_report` et `confusion_matrix`.

3.5.6.3 Métriques utilisées

Les principales métriques utilisées sont les suivantes :

- **Accuracy:**

Mesure la proportion de prédictions correctes. Bien que simple, cette métrique peut être trompeuse en cas de classes déséquilibrées. (**précision globale**).

- **Precision:**

Indique la proportion de prédictions positives qui sont réellement positives. Elle est essentielle pour éviter les **faux positifs**, surtout dans le domaine médical. (**Précision**)

- **Recall:**

Mesure la capacité du modèle à détecter les cas positifs réels. Un bon rappel est crucial dans les contextes médicaux pour minimiser les **faux négatifs**, qui peuvent avoir des conséquences graves. (**Rappel**)

- **F1-score :**

Moyenne harmonique entre la précision et le rappel. Elle offre une évaluation équilibrée, notamment utile lorsque les classes sont déséquilibrées.

- **Matrice_de_confusion:**

Outil visuel qui synthétise les vrais positifs, faux positifs, vrais négatifs et faux négatifs. Elle permet de comprendre les types d'erreurs commises par le modèle.

3.5.6.4 Interprétation des résultats

Les résultats obtenus montrent que le modèle atteint un taux de précision de **96,25 %** a été enregistrée, suggérant que le modèle distingue efficacement les cas de maladie rénale chronique de ceux sans pathologie.

La matrice de confusion illustre une bonne détection des cas de **maladie rénale chronique (CKD)**, avec un nombre limité de faux positifs et de faux négatifs dans la figure 2 :

- **VN (Vrais Négatifs) = 60** : le modèle a correctement prédit 60 patients comme *non atteints* de maladie rénale chronique.
- **VP (Vrais Positifs) = 94** : 94 cas de CKD ont été correctement identifiés.
- **FN (Faux Négatifs) = 6** : 6 patients atteints de CKD ont été classés à tort comme non malades.
- **FP (Faux Positifs) = 0** : aucun patient sain n'a été faussement classé comme malade

L'utilisation conjointe de ces métriques permet une évaluation plus robuste que l'accuracy seule. Cela est particulièrement important dans le domaine biomédical, où certaines erreurs de classification peuvent entraîner des retards de diagnostic ou des traitements inappropriés [66].

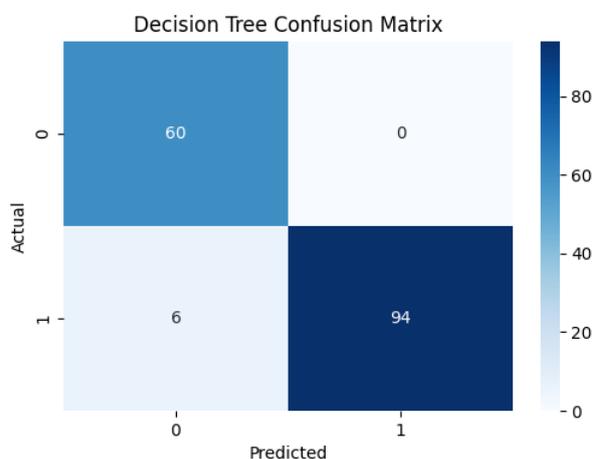


Figure 3.2 Matrice de confusion du modèle Decision Tree pour la détection de la maladie rénale chronique.

3.5.7 Sauvegarde du modèle

Une fois l'entraînement du modèle terminé, il est crucial de **sauvegarder** celui-ci afin d'en permettre la réutilisation sans devoir le réentraîner à chaque exécution. Cela optimise le temps de traitement et assure la reproductibilité des résultats.

En Python, des bibliothèques comme **joblib** ou **pickle** permettent de sérialiser et désérialiser efficacement des objets complexes comme les modèles entraînés.

Selon [67], la sauvegarde des modèles est une étape indispensable dans tout pipeline de machine Learning pour garantir la pérennité et la portabilité des solutions développées.

Dans notre projet, le modèle est sauvegardé au format **.pkl** à l'aide de la méthode **joblib.dump()**, incluant un horodatage afin de suivre les différentes versions entraînées.

3.5.8 Visualisation

La visualisation joue un rôle fondamental dans l'**interprétabilité** des modèles, en particulier pour les arbres de décision, qui permettent une lecture intuitive des règles de classification. À l'aide de la fonction **plot_tree()** de Scikit-learn, l'arbre entraîné peut être représenté graphiquement avec ses nœuds, attributs, seuils de décision, et prédictions associées dans la figure 3.

D'après [68], la transparence des arbres de décision les rend particulièrement adaptés aux domaines sensibles comme la médecine, où les décisions doivent être justifiables auprès des experts et des patients.

Dans notre étude, la visualisation de l'arbre facilite l'analyse des critères médicaux les plus déterminants dans le diagnostic des maladies rénales chroniques.

La figure illustre la structure d'un arbre de décision construit à l'aide de l'algorithme DecisionTreeClassifier de la bibliothèque Scikit-learn, entraîné pour prédire la présence ou l'absence de la maladie rénale chronique (Chronic Kidney Disease - CKD).

Chaque nœud interne de l'arbre représente un test sur une variable clinique (ex. : `hemoglobin_gms <= 12.945`), tandis que chaque nœud terminal (feuille) correspond à une classe prédite : CKD ou Not CKD. Les critères de séparation sont basés sur l'**entropie**, une mesure de l'impureté des classes dans un ensemble de données. [69]

L'arbre met en évidence les caractéristiques cliniques les plus déterminantes pour réduire fortement l'entropie :

- hemoglobin_gms,
- blood_urea_mgs/dl,
- serum_creatinine_mgs/dl,
- pus_cell_clumps:_present.
- Chaque boîte affiche :
 - Le critère de séparation (feature <= valeur),
 - Le niveau d'entropie (plus proche de 0 = pureté),
 - Le nombre d'échantillons (patients),
 - La répartition des classes (value = [Non-CKD, CKD]),
 - La classe majoritaire prédite.
- L'utilisation de la **coloration** (orange pour Not CKD, bleu pour CKD) facilite l'interprétation des décisions prises par le modèle.

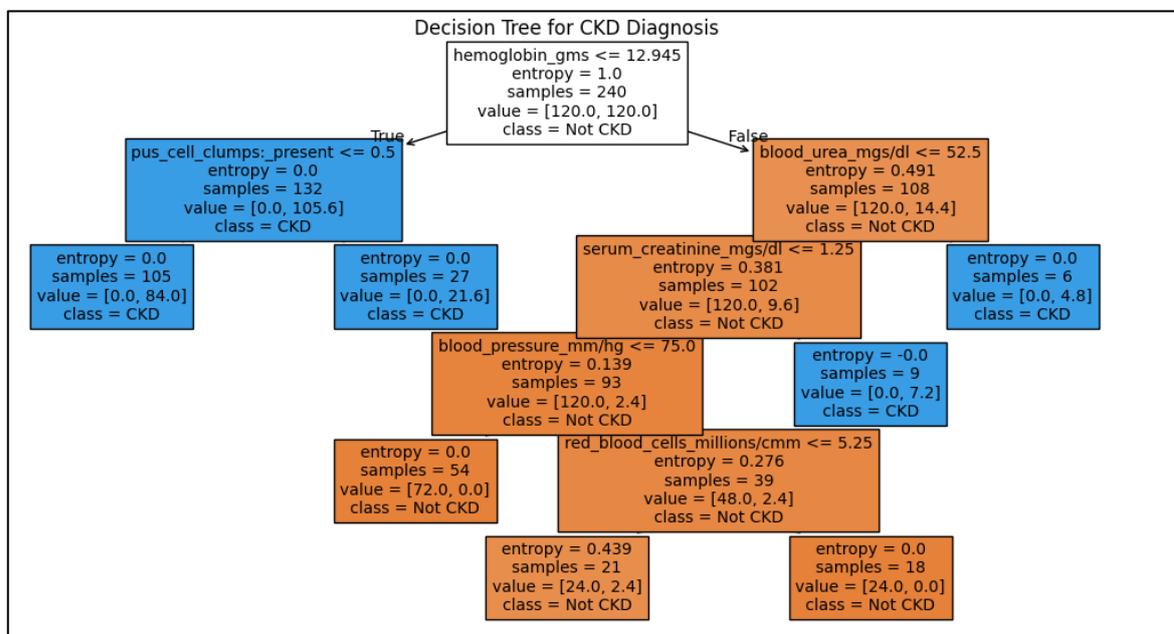


Figure 3.3 Arbre de décision généré pour le diagnostic de la maladie rénale chronique

3.5.9 Implémentation dans l'application ClinicPro

L'intégration d'un modèle de machine learning dans une application clinique représente l'étape finale du processus de développement, visant à transformer un prototype algorithmique en un outil décisionnel accessible aux professionnels de santé. Dans ce projet, le modèle `DecisionTreeClassifier`, une fois entraîné, évalué et sauvegardé, est intégré à une interface graphique développée avec PyQt6. Cette interface permet aux médecins et infirmiers de saisir les paramètres médicaux d'un patient, d'obtenir une prédiction instantanée sur la présence ou non d'une maladie rénale chronique (CKD), et d'en visualiser les facteurs déterminants.

Selon [70], l'intégration de l'intelligence artificielle dans les systèmes cliniques peut renforcer la qualité du diagnostic, améliorer la rapidité des décisions médicales, et contribuer à la personnalisation des soins.

Cette approche s'inscrit dans une logique de **système d'aide à la décision médicale (SADM)**, où l'algorithme ne remplace pas le médecin, mais **le soutient par des analyses prédictives objectives**, exploitant les données historiques et les connaissances extraites des bases cliniques.

3.6 Fonctionnement de notre application ClinicPro

L'application présente une interface claire et intuitive, adaptée aux besoins du personnel médical. Elle facilite la saisie des données, l'affichage des résultats du modèle d'IA, et soutient efficacement la prise de décision clinique. Nous présentons nos interfaces utilisateurs de notre application par les figures ci-dessous.

3.6.1 La page Login

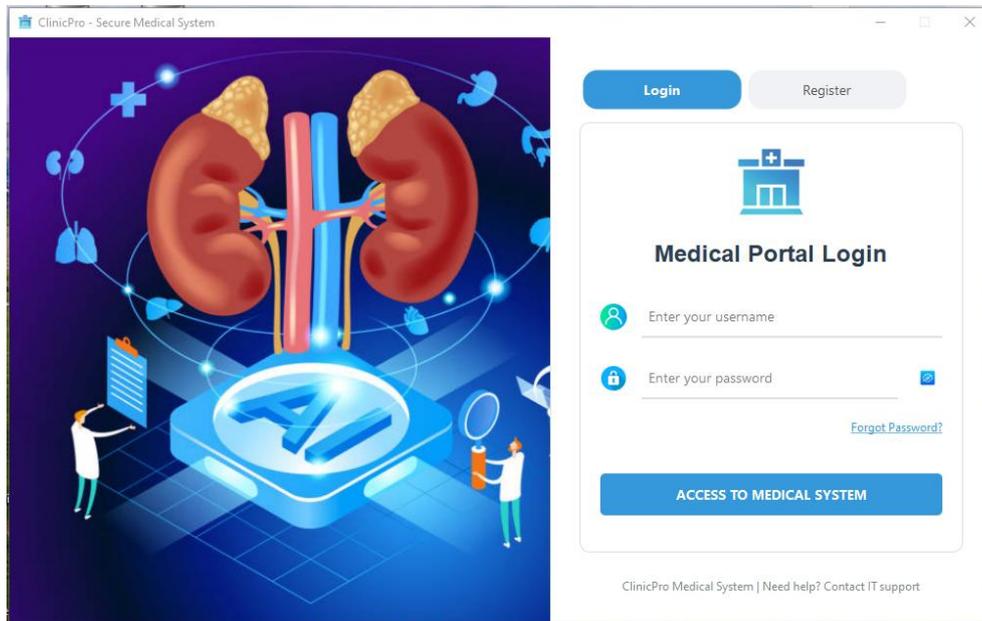


Figure 3.4 Interface « Inscription »

Cette interface offre aux utilisateurs la possibilité de s'enregistrer et de se connecter de manière sécurisée au système.

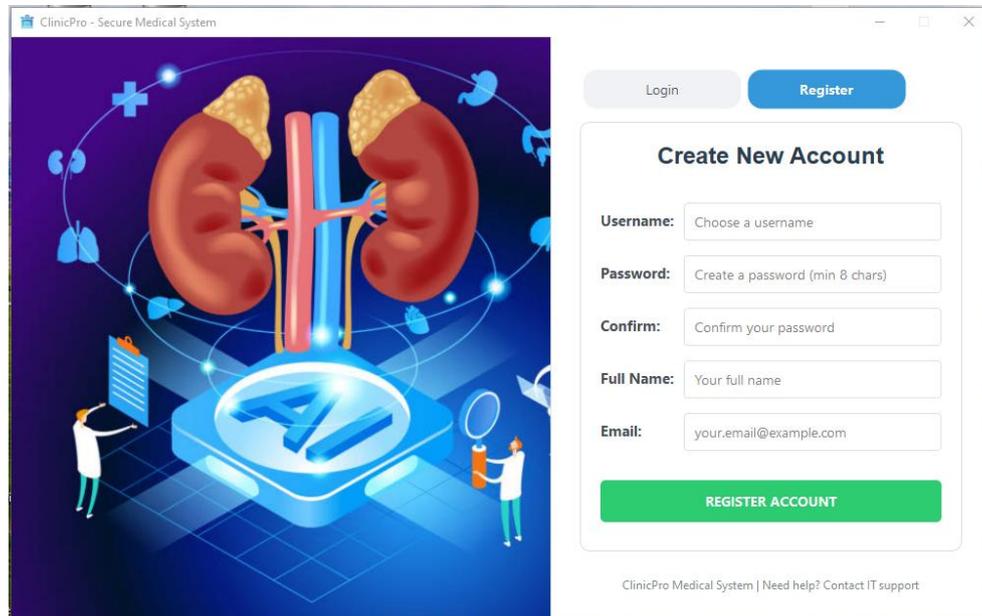


Figure 3.5 Interface « Authentification »

La page d'inscription permet à l'utilisateur de remplir un formulaire avec ses informations personnelles afin de créer un compte sécurisé dans le système.

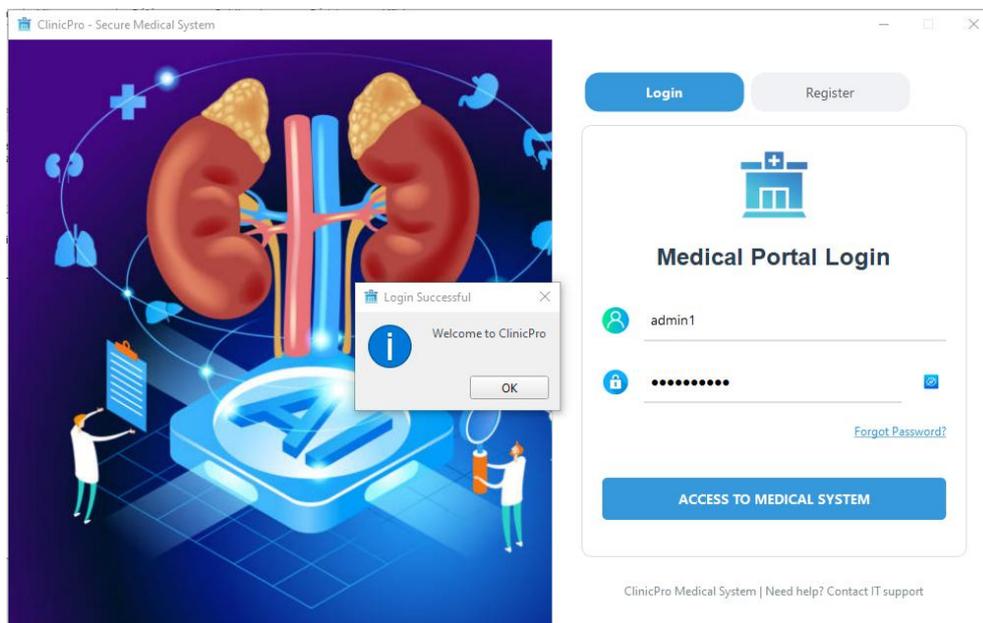


Figure 3.6 Interface «Création d'un compte sécurisé»

L'interface demande la saisie du nom d'utilisateur et du mot de passe pour l'authentification

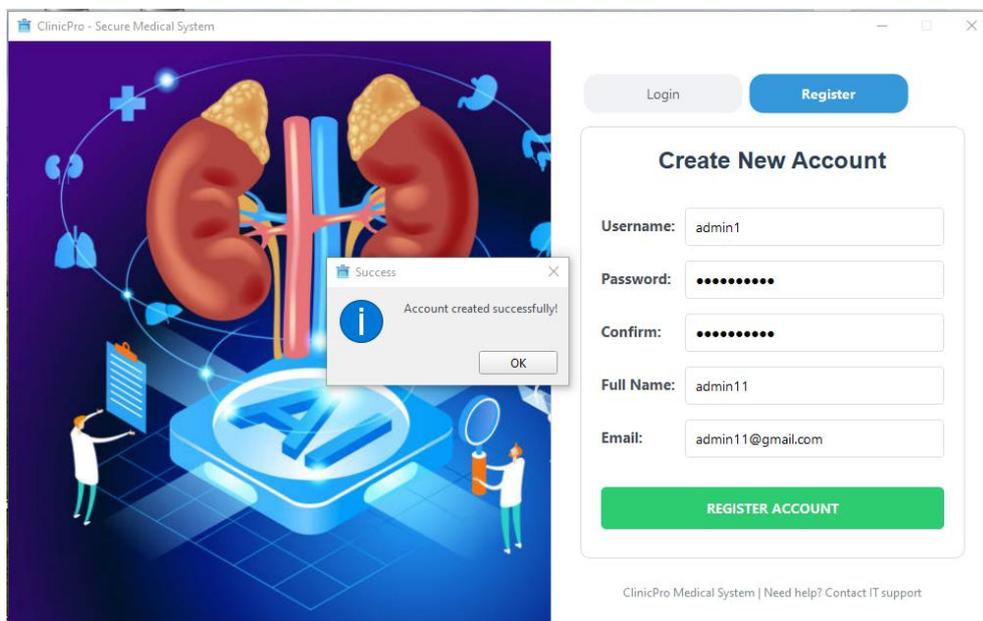


Figure 3.7 Interface «Authentification à un compte réussie»

3.6.2 Page de sélection de rôle

Après l'authentification, l'utilisateur est redirigé vers une page de sélection de rôle lui permettant de choisir son profil (ex. : médecin ou infirmier) pour accéder aux fonctionnalités correspondantes.

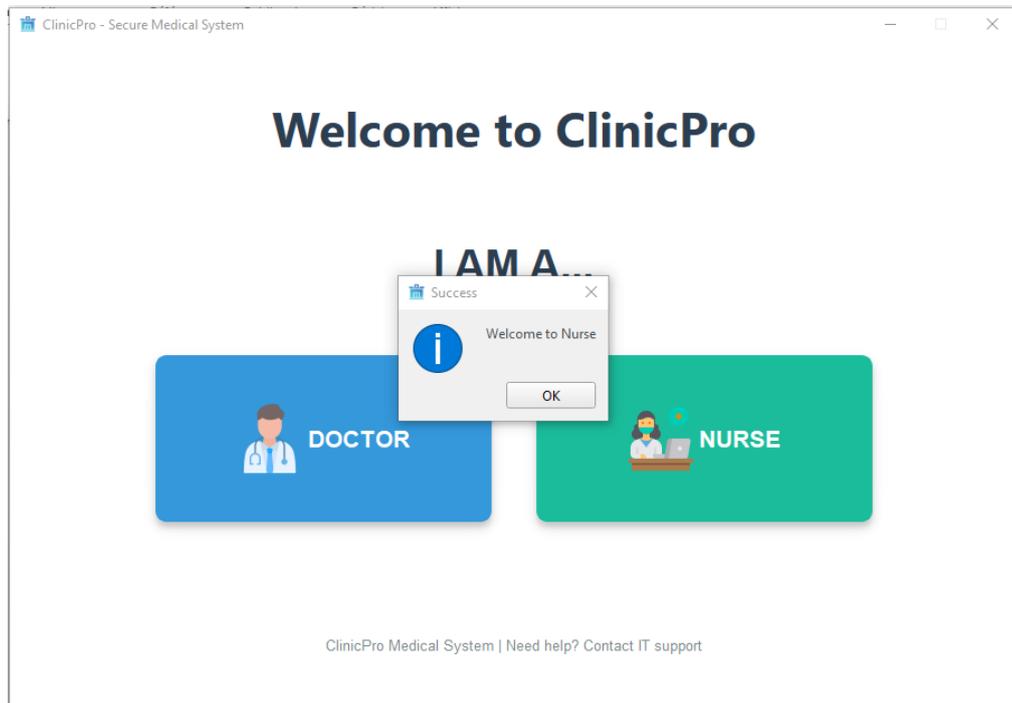


Figure 3.8 Interface «sélection de rôle».

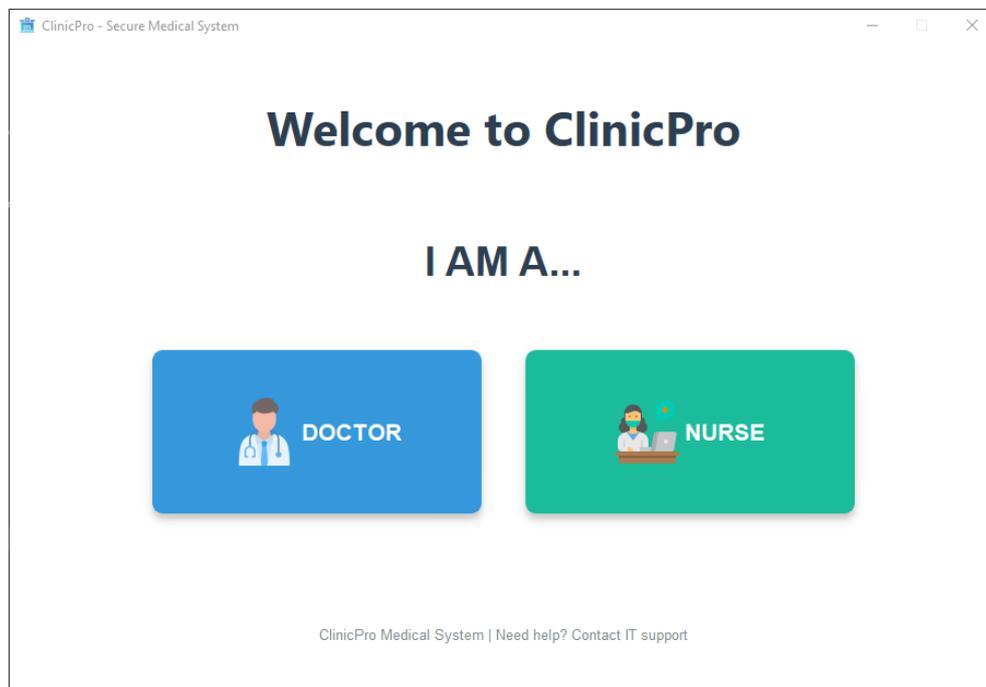


Figure 3.9 Interface «sélection de rôle de l'infirmière».

Espace dédié à l'infirmière pour la gestion et le suivi des dossiers médicaux.

3.6.3 Interface infirmière

Elle dédiée à la saisie des informations personnelles des patients.

F_name	L_name	CIN	NSS	Phone	Email	Address
1 sara	sh	55555555	888888888888	0755555555	sarasouyah@g...	saidajdfhsdfsh...

Figure 3.10 Interface « L'infirmière accède aux données personnelles ».

Elle dédiée à la saisie des informations médicales des patients.

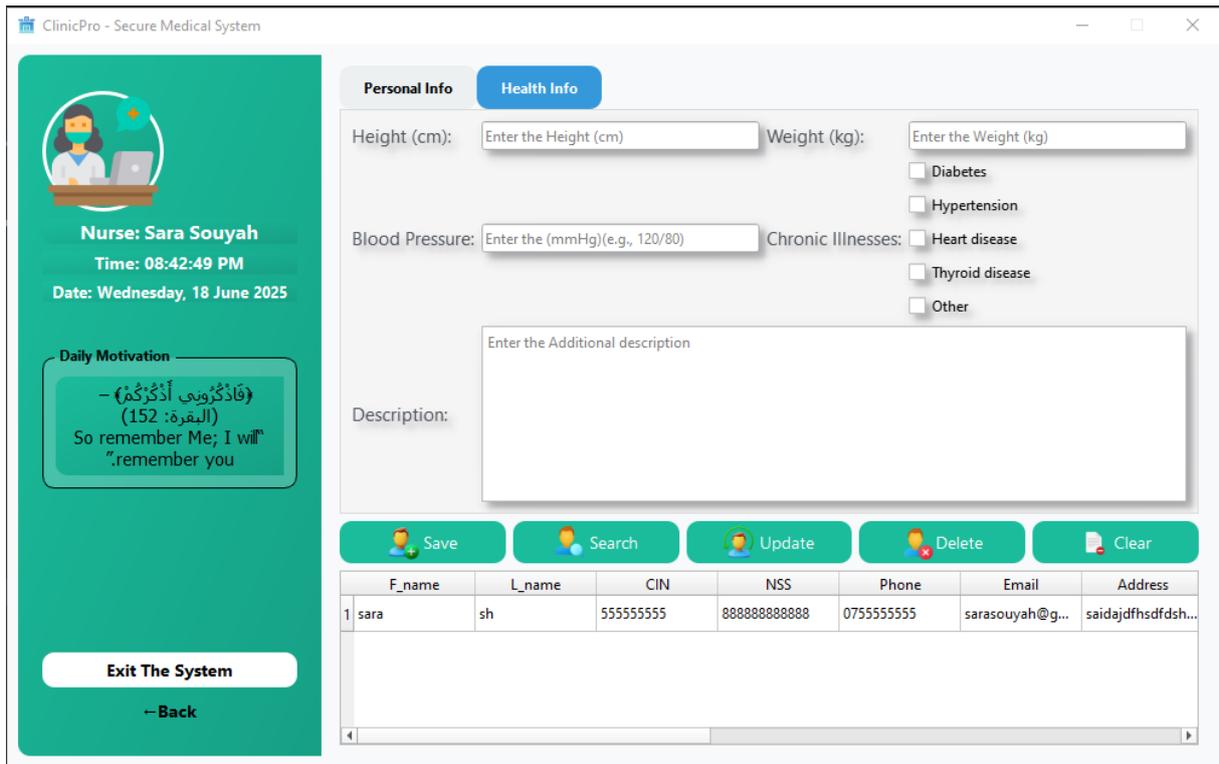


Figure 3.11 Interface «L’infirmière accède aux données médicales.».

Enregistrement réussi des informations personnelles par l’infirmière.

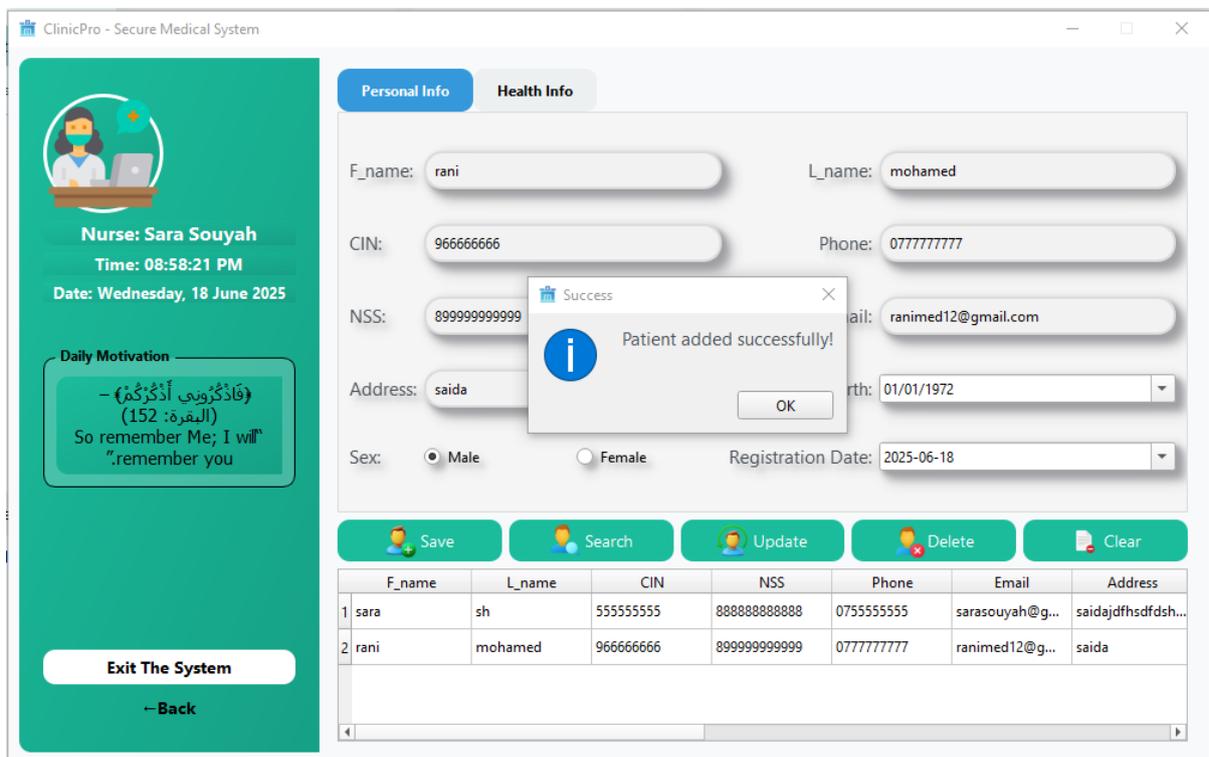


Figure 3.12 Interface « l’infirmière ajoute les informations personnelles du patient ».

Enregistrement réussi des informations médicales par l’infirmière.

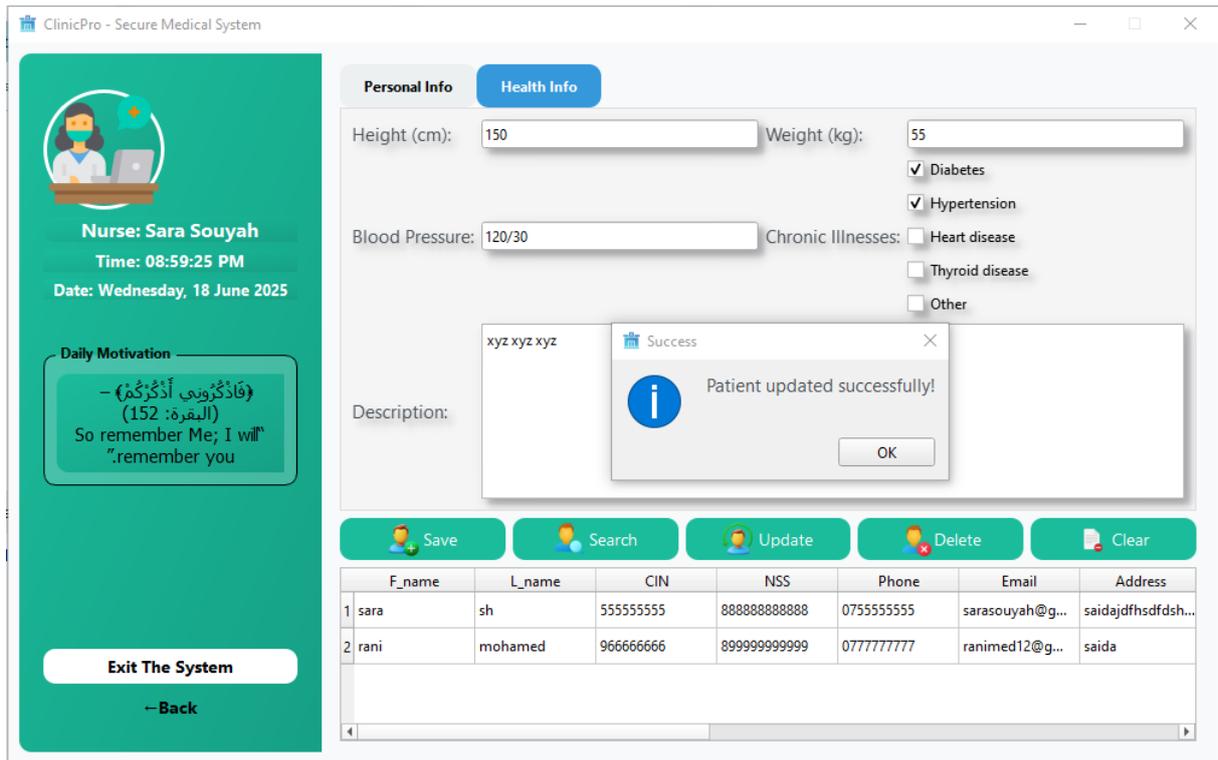


Figure 3.13 Interface «l’infirmière modifier les informations médicales du patient».

3.6.4 Interface médecin

Seul le médecin, après authentification par nom d'utilisateur et mot de passe, peut accéder à son espace personnel. L’infirmier ne dispose pas de cet accès.

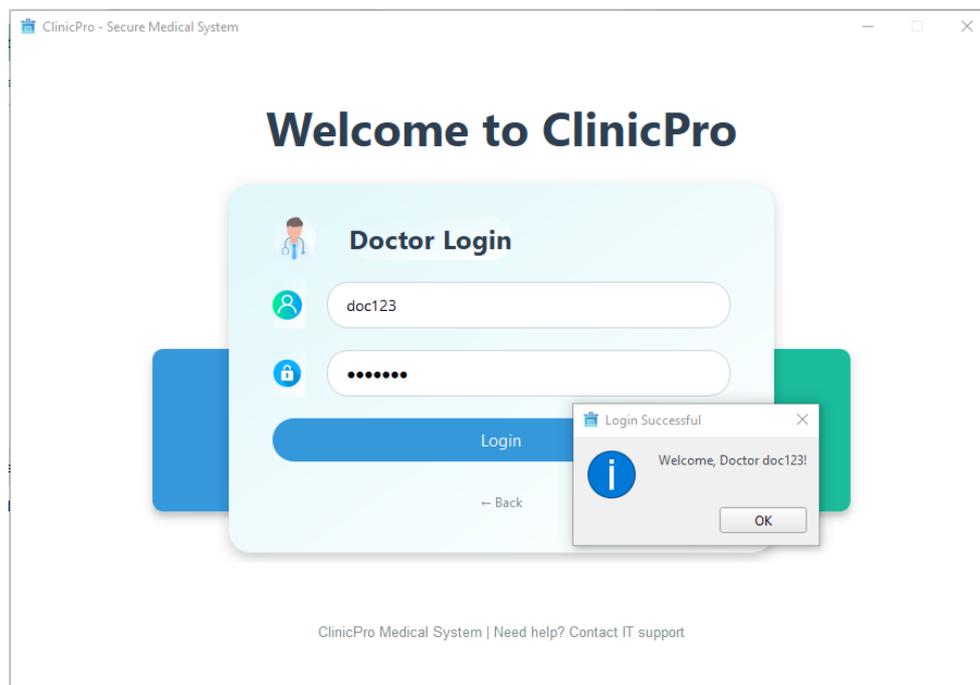


Figure 3.14 Interface «Login Médecin».

3.6.4.1 Recherche de patient

Le médecin peut rechercher un patient à l'aide d'un identifiant unique (ex. : numéro CIN), afin de consulter son profil complet.

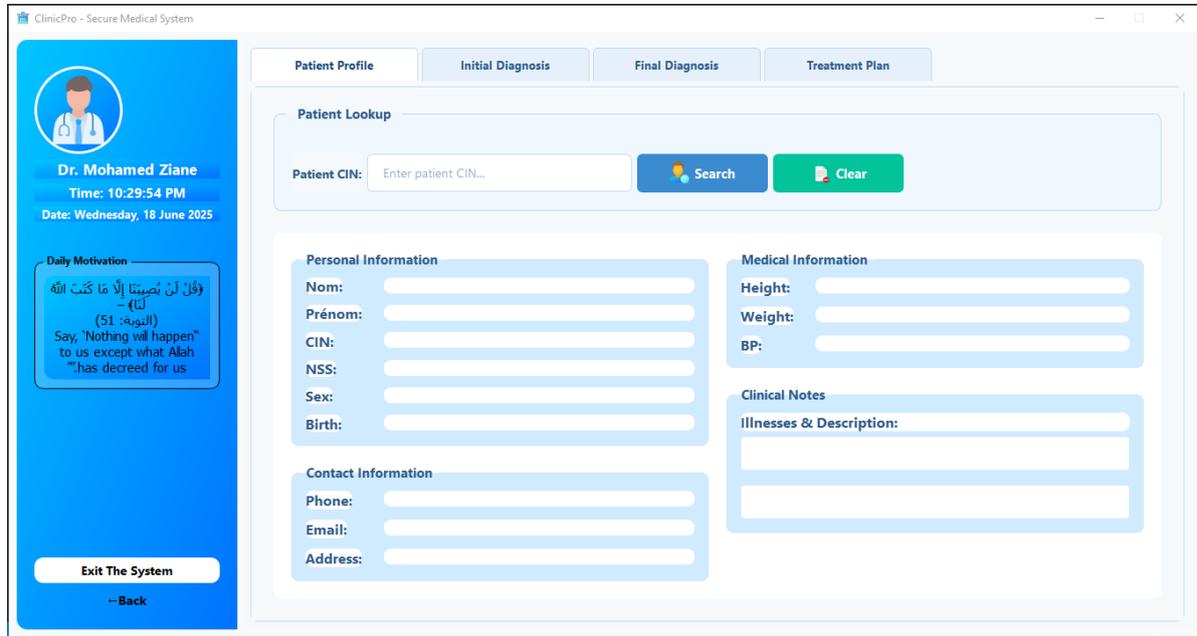


Figure 3.15 Interface «recherche de patient».

Le médecin peut accéder avec succès à l'ensemble de son profil médical et personnel.

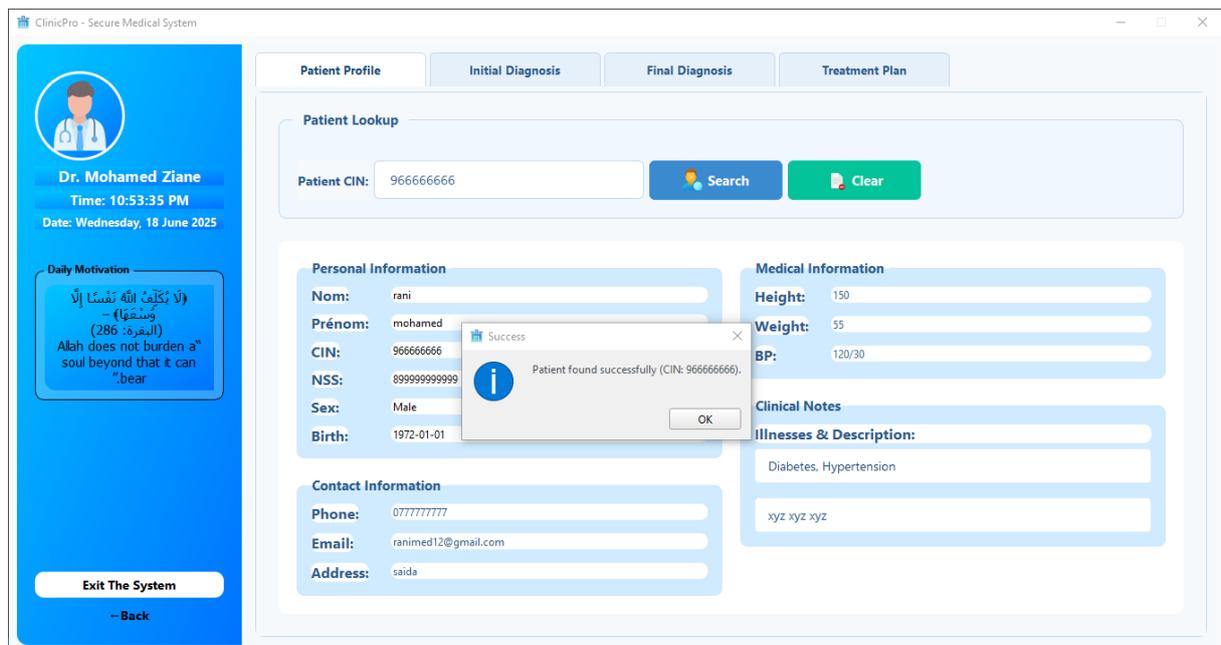


Figure 3.16 Interface «patient trouvé avec succès».

3.6.4.2 Diagnostic initial

Une fois les données cliniques analysées, le médecin peut initier un **premier diagnostic assisté**, en s'appuyant sur les **antécédents médicaux du patient** ainsi que sur les **symptômes précoces et tardifs observés**.

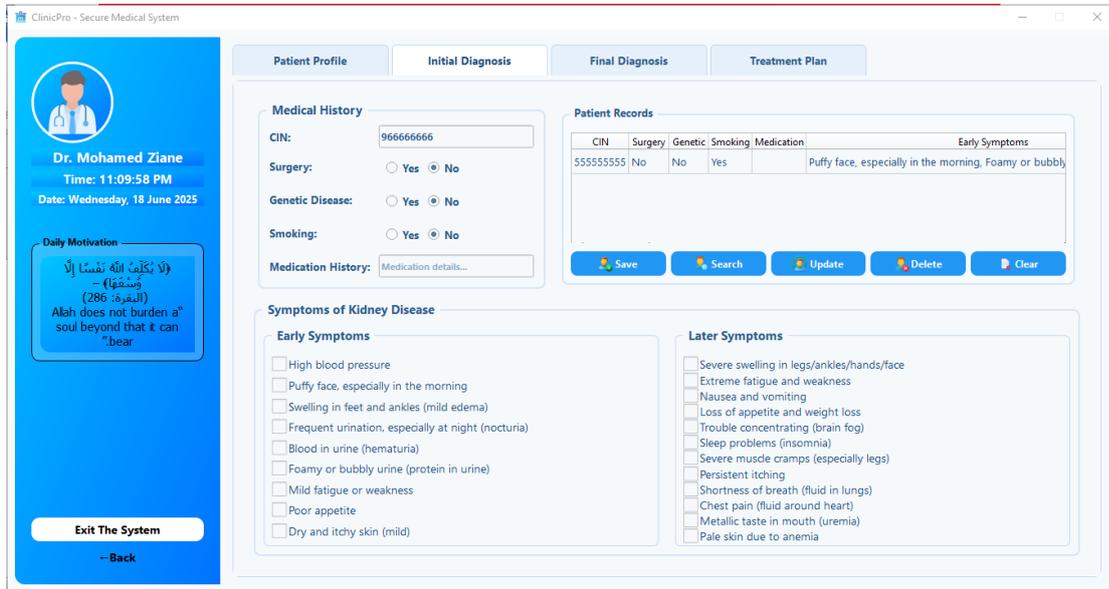


Figure 3.17 Interface «diagnostic initial».

Le médecin initie le diagnostic assisté après validation réussie des antécédents et symptômes cliniques.

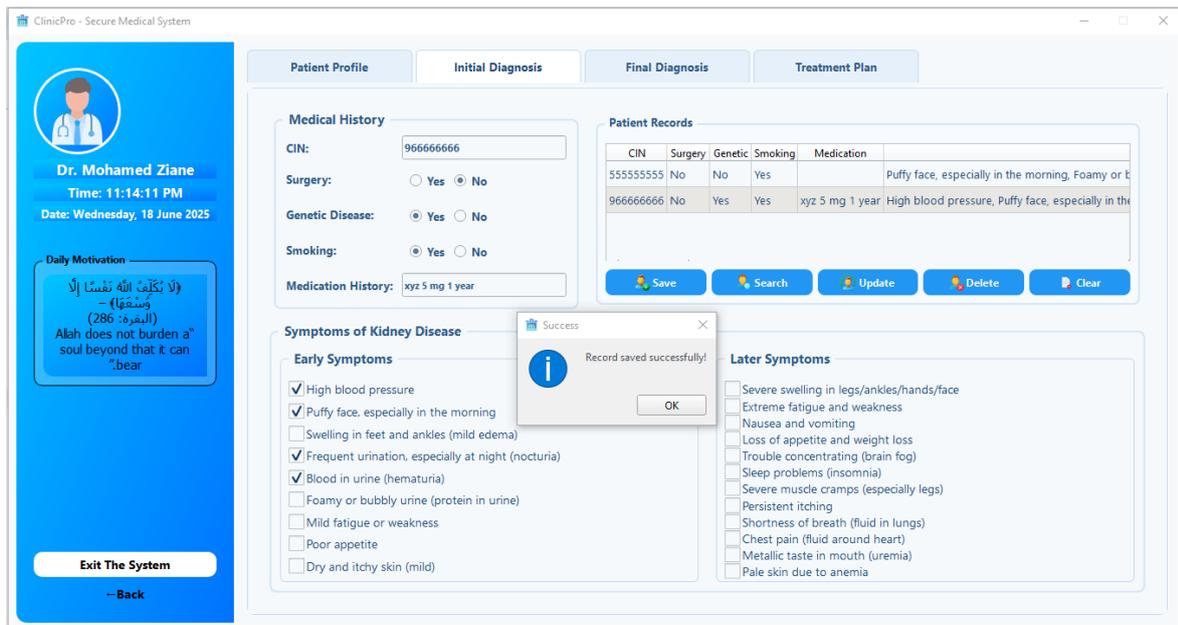


Figure 3.18 Interface « diagnostic initial avec patient réussi».

3.6.4.3 Diagnostic finale

À partir du diagnostic initial, le médecin produit une analyse personnalisable et imprimable, destinée à appuyer la décision clinique.

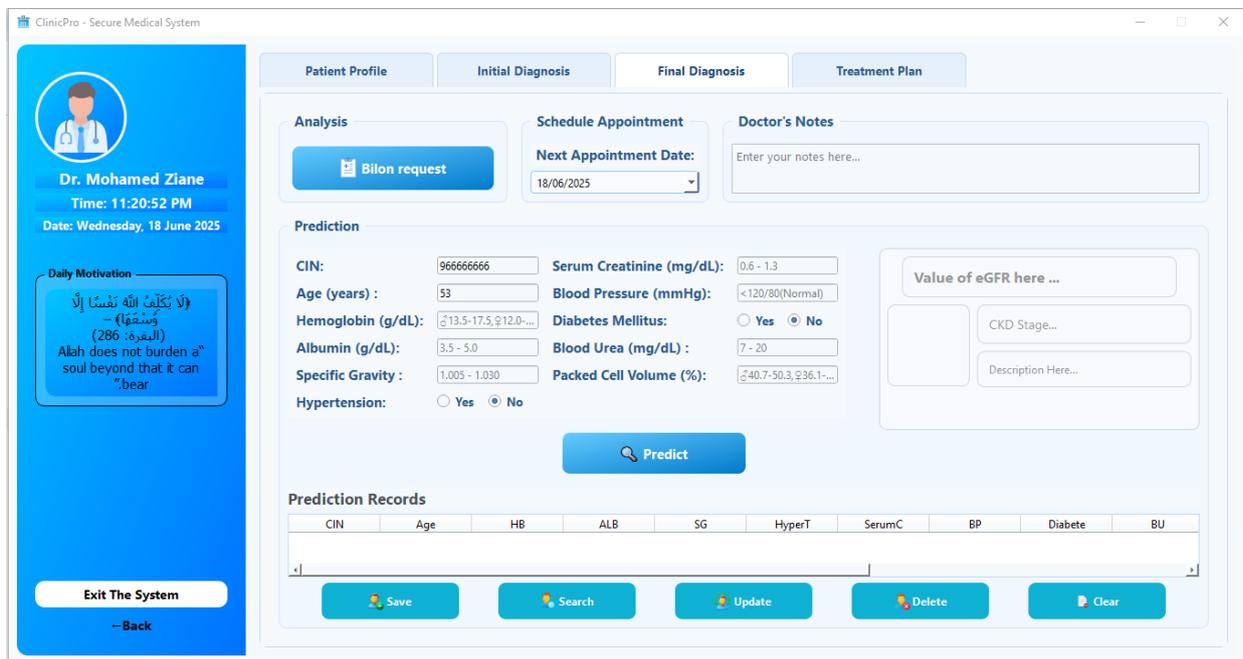


Figure 3.19 Interface « diagnostic finale ».

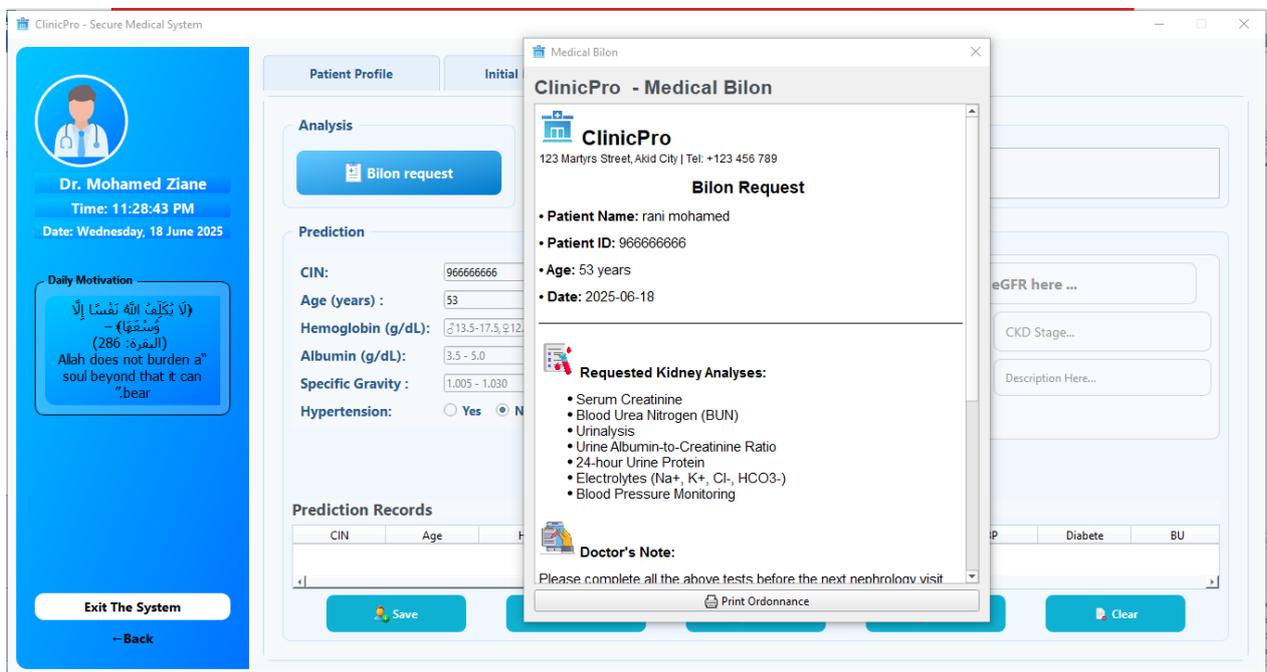


Figure 3.20 Interface «demande de bilan lancée par le médecin.».

Génération automatisée de l'ordonnance médicale en format PDF pour impression et archivage.



ClinicPro
123 Martyrs Street, Akid City | Tel: +123 456 789

Bilon Request

- **Patient Name:** rani mohamed
- **Patient ID:** 966666666
- **Age:** 53 years
- **Date:** 2025-06-18



Requested Kidney Analyses:

- Serum Creatinine
- Blood Urea Nitrogen (BUN)
- Urinalysis
- Urine Albumin-to-Creatinine Ratio
- 24-hour Urine Protein
- Electrolytes (Na+, K+, Cl-, HCO3-)
- Blood Pressure Monitoring



Doctor's Note:

Please complete all the above tests before the next nephrology visit. Contact us if any urgent symptoms arise.



Dr. Mohamed Ziane
Nephrologist

This ordonnance was generated by ClinicPro - Secure Medical System.

Figure 3.21 Interface «les analyses demandées par le médecin».

À partir des analyses demandées, une prédiction médicale est automatiquement établie afin d'affiner le diagnostic du patient. Le médecin ajoute une note explicative ainsi que la date du prochain rendez-vous pour assurer le suivi médical. En complément, le système calcule le **taux de filtration glomérulaire (GFR)** à l'aide de la formule CKD-EPI, ce qui permet une évaluation plus précise de la fonction rénale (Voir la figure 3.22)

Le médecin saisit les résultats des analyses demandées, qui sont ensuite utilisés par le système d'intelligence artificielle pour générer une prédiction médicale. En parallèle, le taux de filtration glomérulaire (GFR) est automatiquement calculé, permettant de déterminer le stade d'évolution de la maladie rénale conformément aux critères cliniques.

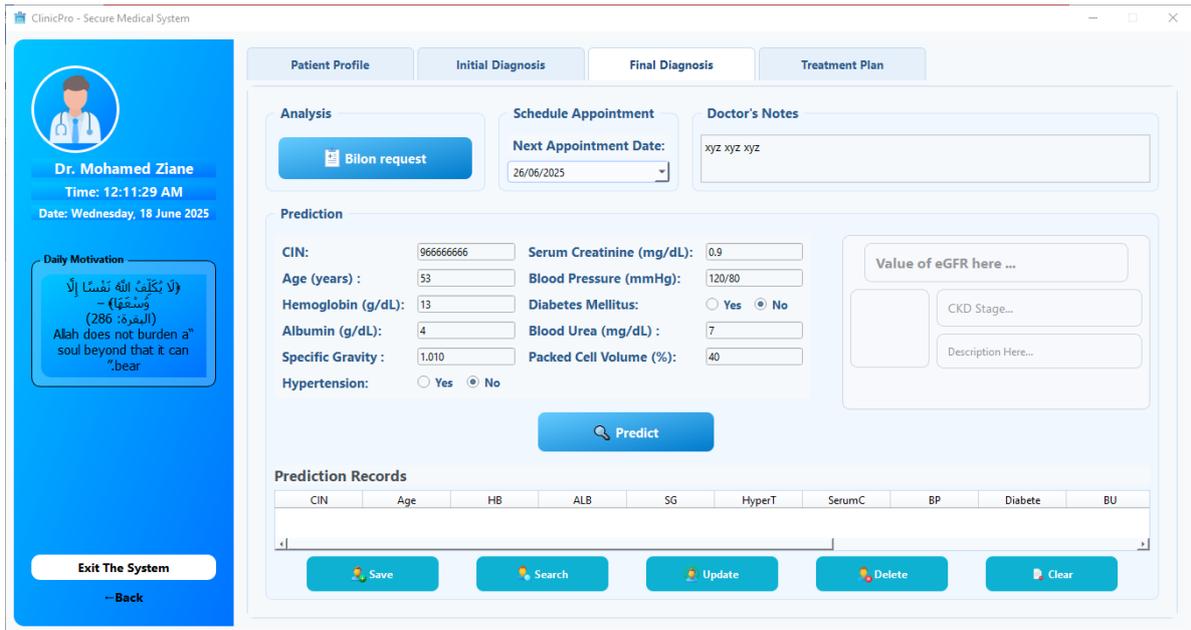


Figure 3.22 Interface «médecin remplir les champs pour faire une prédiction».

La prédiction indique que le patient est en bonne santé.

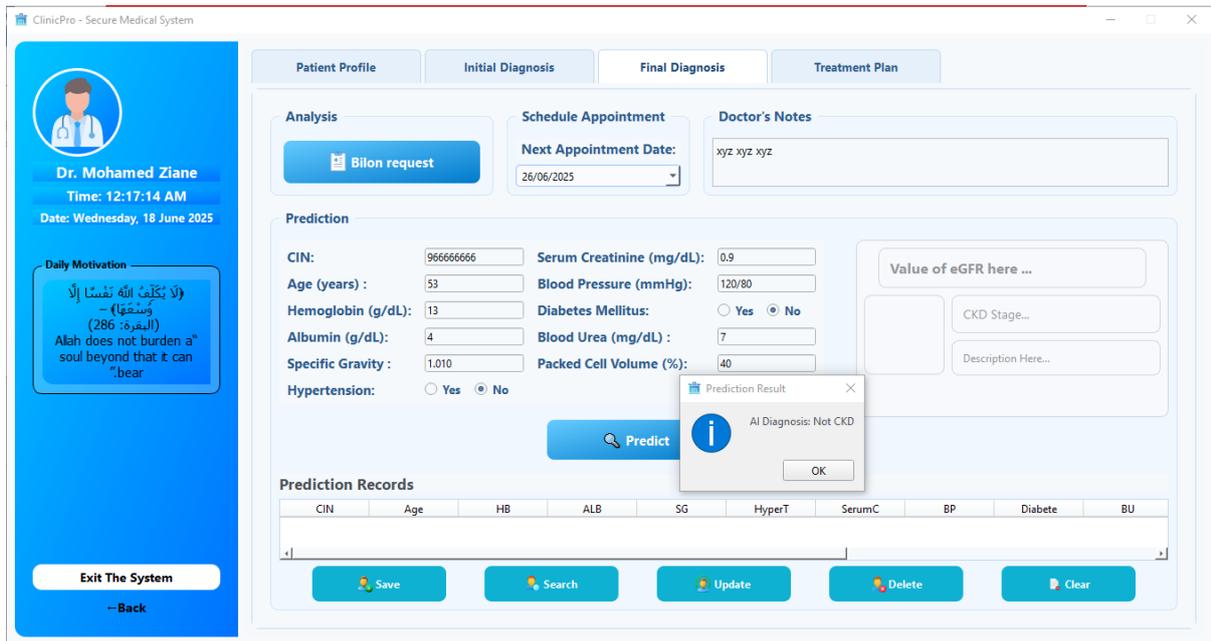


Figure 3.23 Interface «le système fait la prédiction et donne la décision Not CKD».

Le système calcule automatiquement le taux de filtration glomérulaire (GFR), puis identifie et mentionne le stade de la maladie rénale conformément aux classifications cliniques.

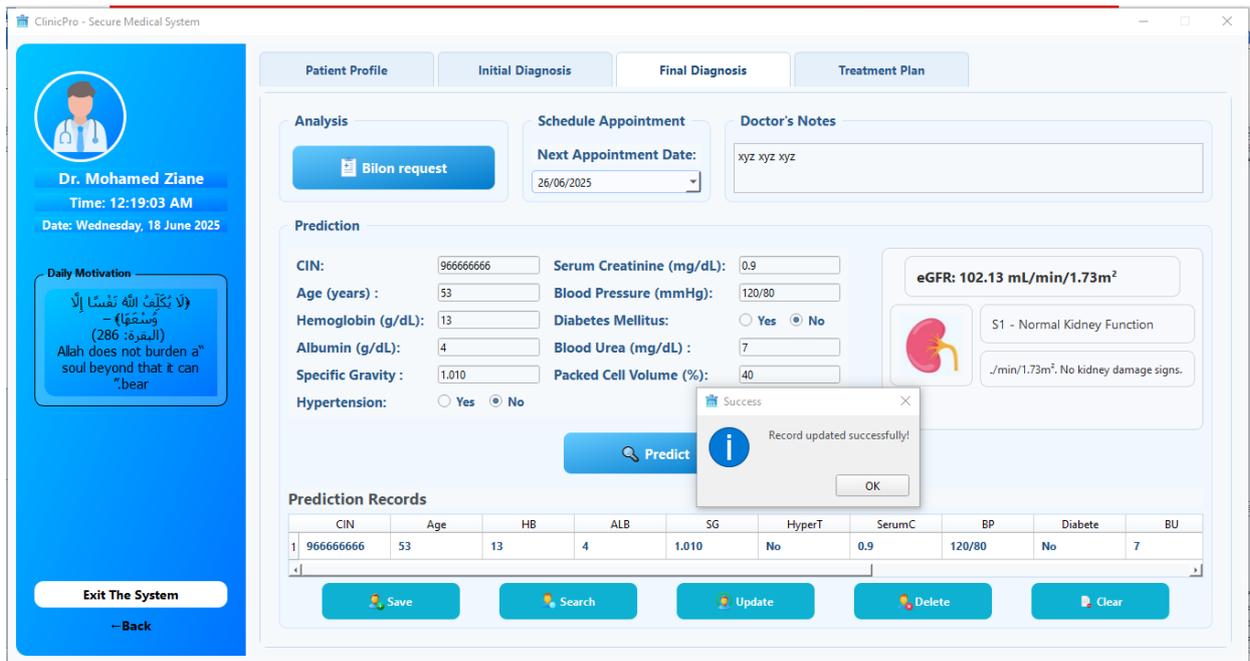


Figure 3.24 Interface « Le médecin ajoute un patient dans le diagnostic final et le GFR est calculé automatiquement ».

La prédiction indique que le patient est malade.

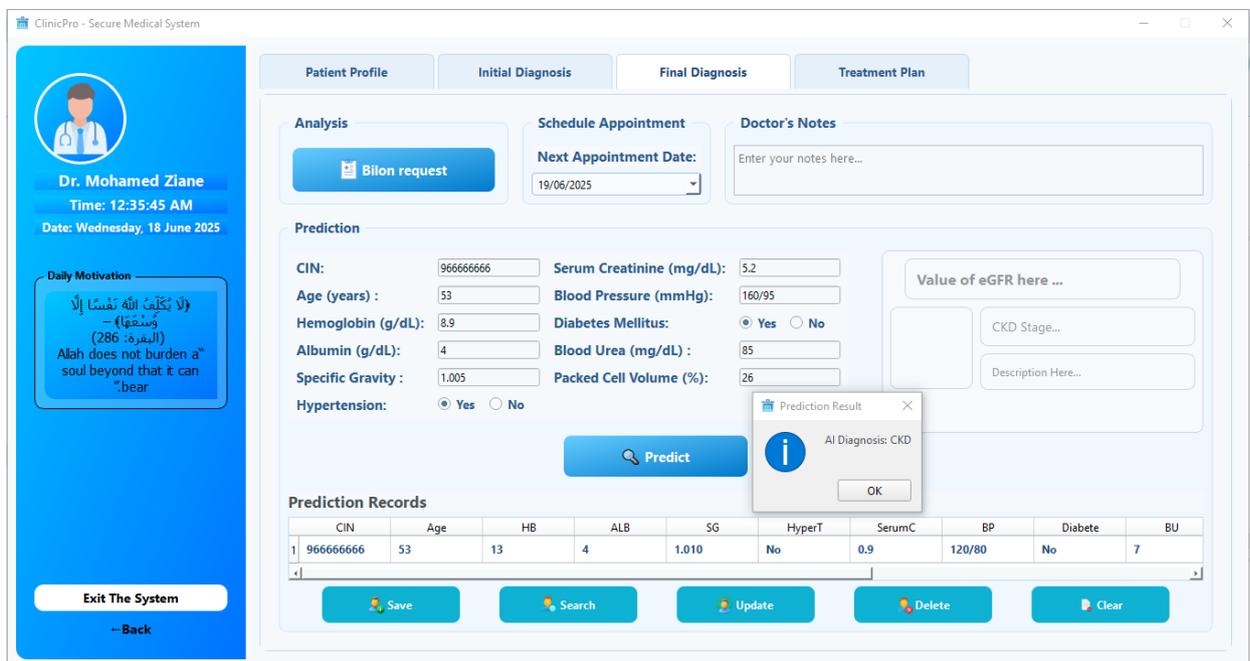


Figure 3.25 Interface «le système fait la prédiction et donne la décision CKD».

Le système calcule automatiquement le taux de filtration glomérulaire (GFR), puis identifie et mentionne le stade de la maladie rénale conformément aux classifications cliniques.

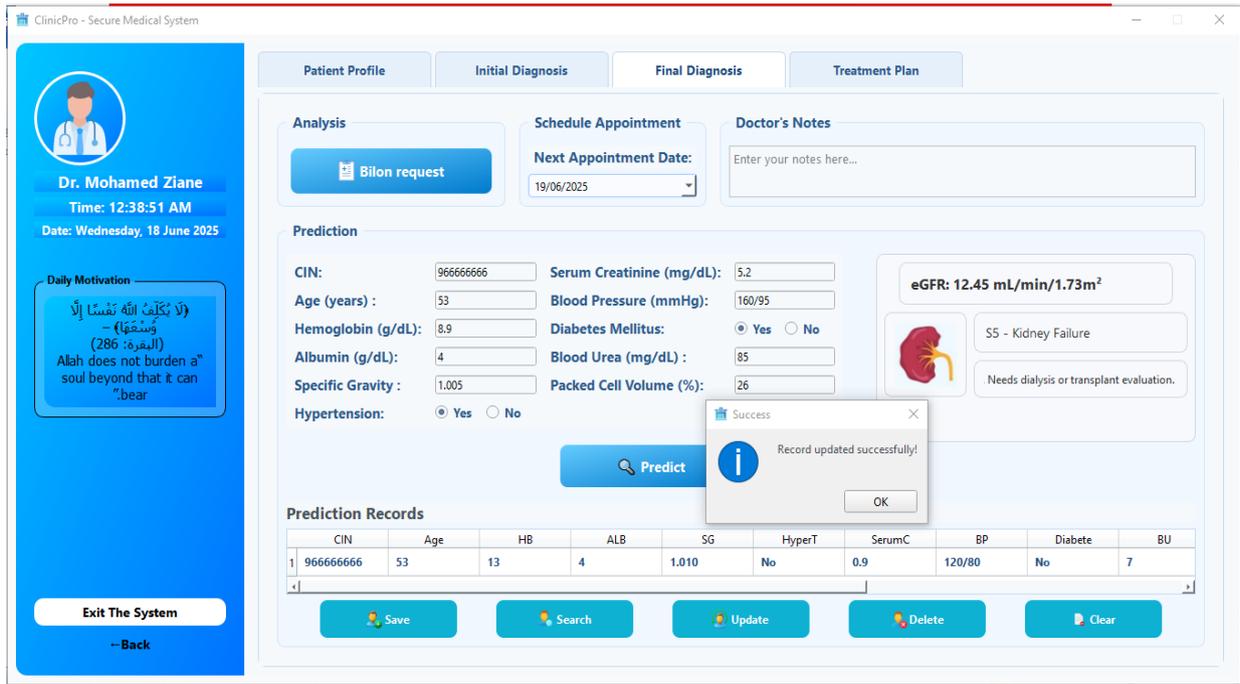


Figure 3.26 Interface «le médecin ajoute un patient dans le diagnostic final et le GFR est calculé automatiquement».

3.6.4.4 Plan de traitement

Il comporte une **ordonnance médicale adaptée au stade de la maladie rénale**, que le médecin peut remplir manuellement ou générer à l'aide du système. En parallèle, un **suivi médical personnalisé** est proposé, avec des recommandations spécifiques à chaque stade, afin d'assurer une **prise en charge continue** et conforme aux protocoles cliniques.

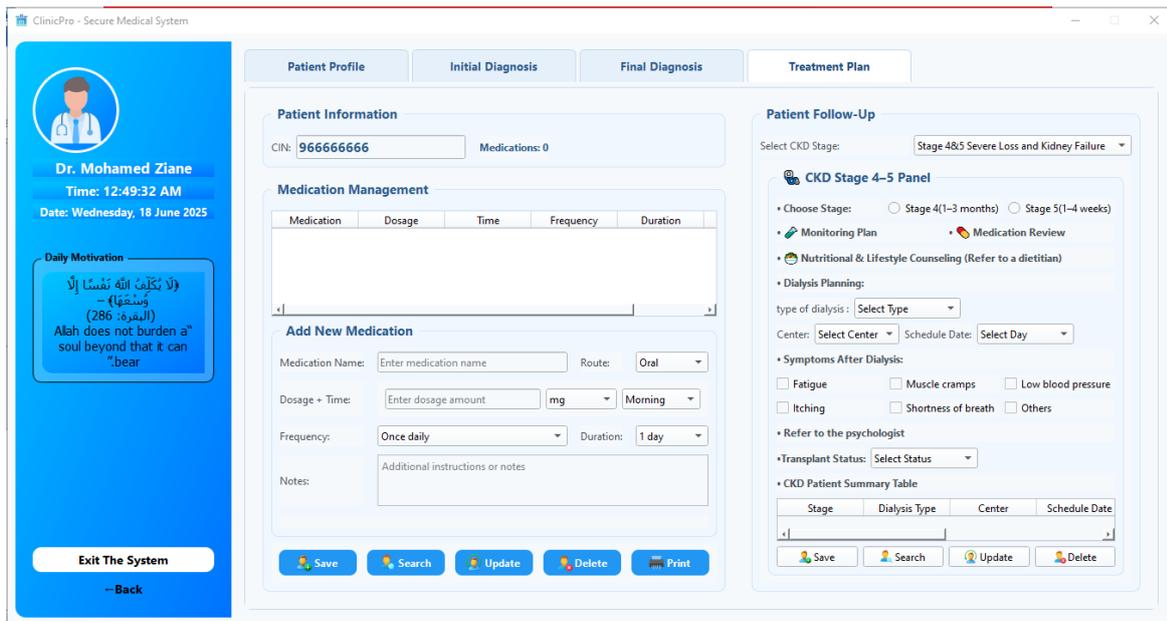


Figure 3.27 Interface «plan de traitement».

Le système permet au médecin de **proposer un traitement adapté**, en saisissant les **médicaments prescrits** selon le stade de la maladie.

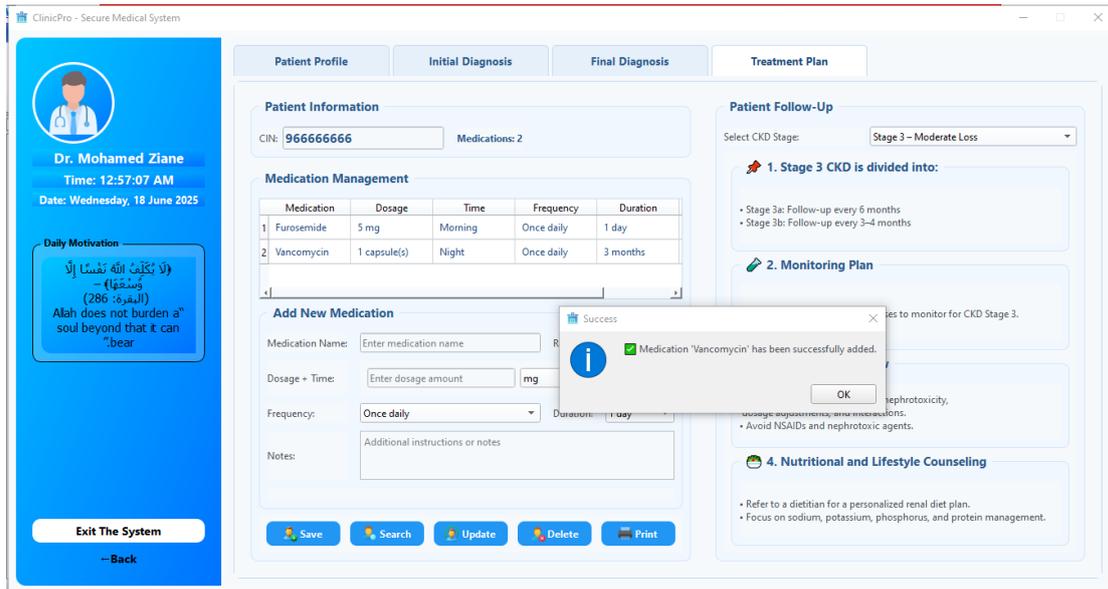


Figure 3.28 Interface « Le médicament a été ajouté avec succès.».



Figure 3.29 Interface «le médecin imprime l'ordonnance ».

Une fois l'ordonnance complétée, elle peut être **générée automatiquement au format PDF** afin de faciliter l'impression, l'archivage ou la remise au patient.

4.6.4.5 Suivi du patient

Un suivi médical adapté est proposé en fonction du stade de la maladie rénale identifié.

Pour les patients au stade **1 et 2** de la maladie rénale chronique (MRC), le système propose un **document de recommandations personnalisées**, incluant des **conseils médicaux, diététiques et comportementaux** visant à ralentir la progression de la maladie.

Patient Follow-Up

Select CKD Stage: Stage 1&2 Normal Function and Mild Loss

BP_Diabetes Monitoring Plan

Date	BP(mmHg)	Glucose(mg/dL)	Notes

Lifestyle Advice

- Reduce sodium and limit phosphorus.
- Limit protein intake and stay hydrated.
- Exercise regularly and maintain a healthy weight.
- Quit smoking and limit alcohol.
- Manage stress & sleep (7–9 hours, relaxation techniques).
- Avoid NSAIDs and kidney-harming drugs.

Monitoring Plan

- Check serum creatinine (every 3–6 months)
- Urine albumin-to-creatinine ratio (UACR) test (every 6 months)
- Hemoglobin(every 6 months)

Print Plan

Figure 3.30 Interface «le suivi des patients aux stades 1 et 2 (MRC)».

ClinicPro
123 Martyrs Street, Akid City | Tel: +123 456 789

CKD Stage 1–2 Follow-Up Plan

- Patient Name: rani mohamed
- Patient ID: 966666666
- Age: 53 years
- Date: 2025-06-19

BP & Diabetes Monitoring

Date of control	BloodP(mmHg)	Glucose(mg/dL)	Notes/observation

Lifestyle Advice

- Reduce sodium and limit phosphorus
- Limit protein intake and stay hydrated
- Exercise regularly and maintain healthy weight
- Quit smoking and limit alcohol
- Manage stress and sleep (7–9 hours)
- Avoid NSAIDs and kidney-harming drugs

Monitoring Plan

- Check serum creatinine (every 3–6 months)
- Urine albumin-to-creatinine ratio (UACR) test (every 6 months)
- Hemoglobin (every 6 months)



Dr. Mohamed Ziane
Nephrologist

This ordonnance was generated by ClinicPro - Secure Medical System.

Figure 3.31 Interface «plan de suivi imprimable pour les stades 1 et 2 (MRC)».

Le médecin peut, en un clic, **générer et imprimer ce document au format PDF.**

Le suivi des patients aux stades 3a et 3b de la maladie rénale chronique repose sur une approche intégrée combinant :

- une **surveillance biologique régulière** (DFG, créatinine, électrolytes, hémoglobine),
- une **prise en charge médicamenteuse adaptée** (IEC/ARA II, traitement de l’anémie, ajustement des antidiabétiques),
- des **recommandations nutritionnelles ciblées,**
- et des **mesures hygiéno-diététiques**

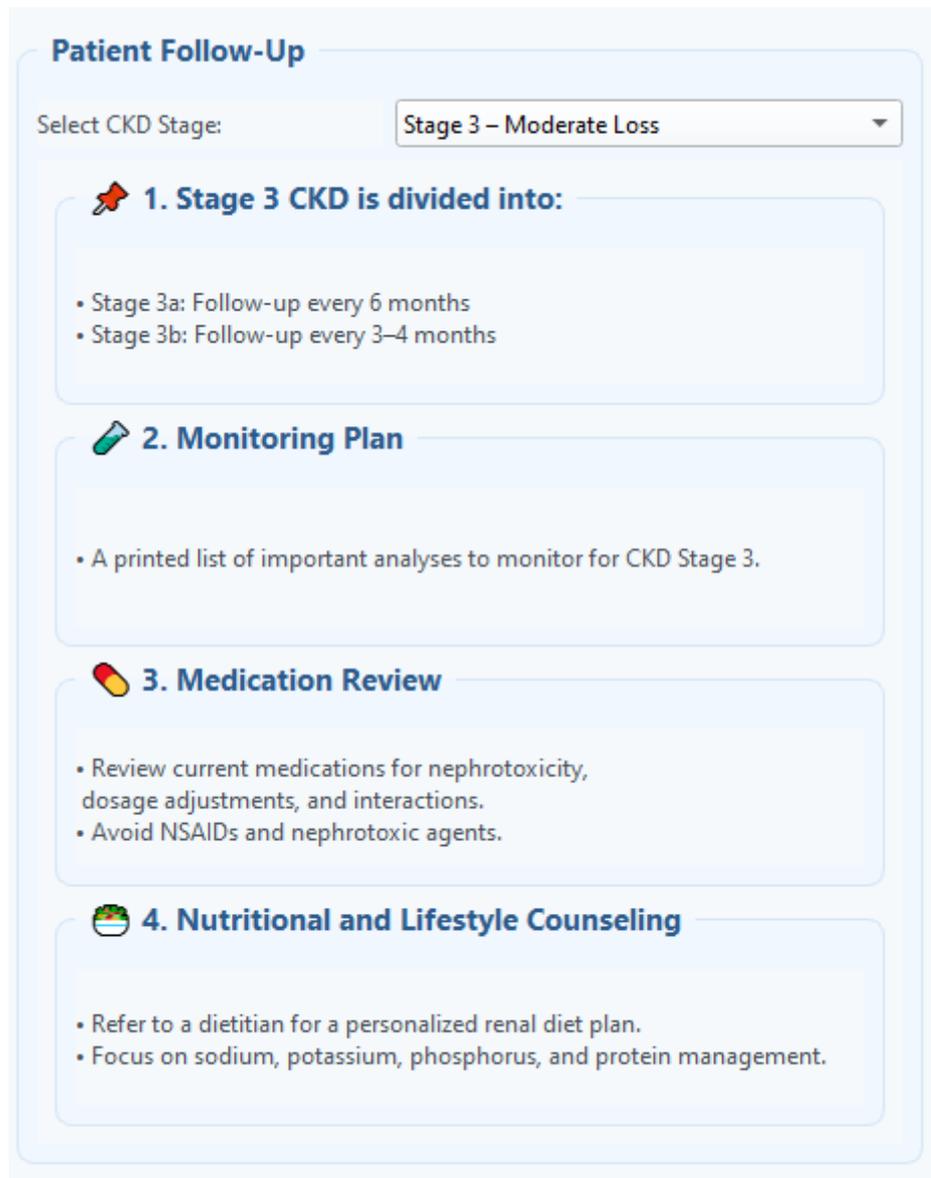


Figure 3.32 Interface «le suivi des patients au stade 3 (MRC)».

Le suivi des patients atteints des stades 4 et 5 de la maladie rénale chronique (MRC) nécessite une prise en charge spécialisée et multidisciplinaire, incluant :

- une surveillance biologique étroite,
- la préparation à un traitement de suppléance rénale (dialyse ou transplantation),
- la prise en charge des complications métaboliques,
- un suivi nutritionnel personnalisé,
- ainsi qu'un accompagnement psychologique visant à aider le patient à faire face à la chronicité de la maladie et à son impact émotionnel.

Patient Follow-Up

Select CKD Stage: Stage 4&5 Severe Loss and Kidney Failure

CKD Stage 4–5 Panel

- Choose Stage: Stage 4(1–3 months) Stage 5(1–4 weeks)
- Monitoring Plan • Medication Review
- Nutritional & Lifestyle Counseling (Refer to a dietitian)
- Dialysis Planning:
 - type of dialysis : Select Type
 - Center: Select Center Schedule Date: Select Day
- Symptoms After Dialysis:
 - Fatigue Muscle cramps Low blood pressure
 - Itching Shortness of breath Others
- Refer to the psychologist
- Transplant Status: Select Status
- CKD Patient Summary Table

Dialysis Type	Center	Schedule Date	Symptoms
Peritoneal	Center A	Mon Wed Fri	Fatigue

Figure 3.33 Interface «le suivi des patients aux stades 4et 5 (MRC)».

Le suivi des patients aux stades 4 et 5 de la maladie rénale chronique est assuré avec succès grâce à une prise en charge spécialisée, combinant surveillance biologique, préparation à la dialyse ou transplantation rénale gestion des complications, suivi nutritionnel et accompagnement psychologique;

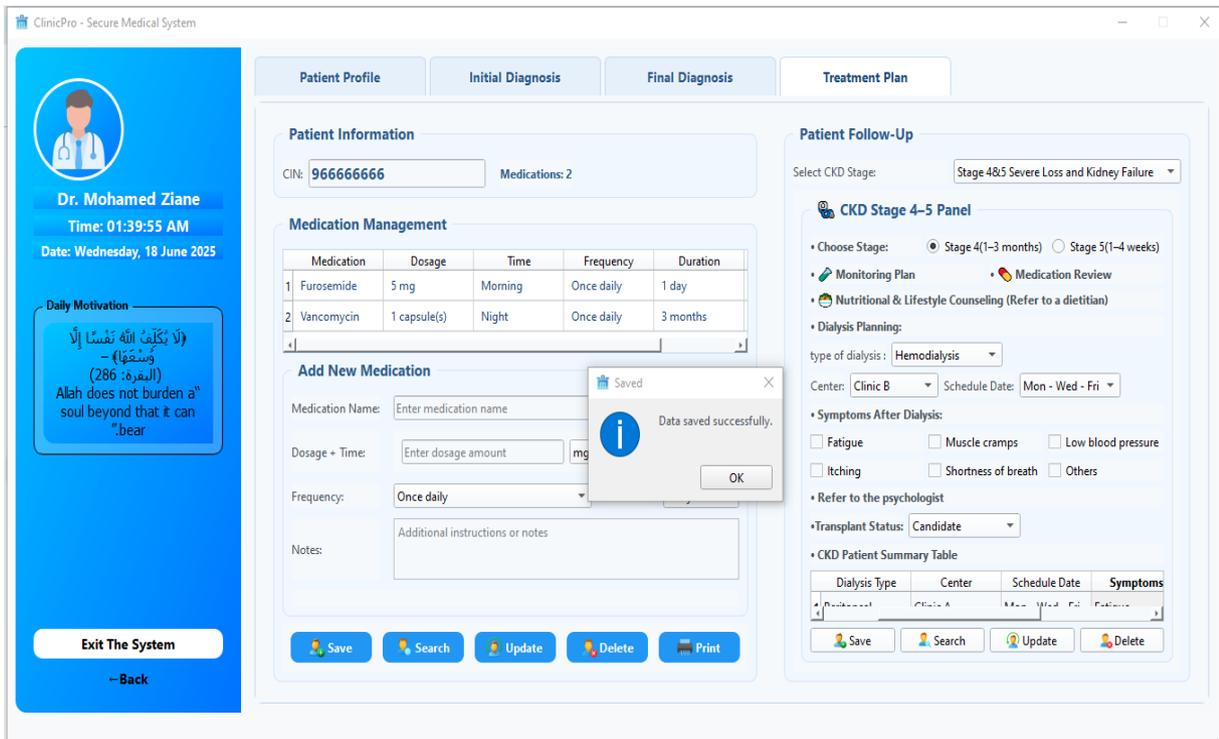


Figure 3.34 Interface «le suivi des patients aux stades 4 ou 5 (MRC) ajouté avec succès».

3.7 Discussion

Le développement du système d'aide à la décision s'est appuyé sur des choix technologiques judicieux, alliant **efficacité**, **simplicité de déploiement** et **pertinence médicale**. L'utilisation de **Python** a permis une intégration fluide des modules de traitement des données, de prédiction par intelligence artificielle, et de création d'interfaces graphiques via **PyQt6**.

Le recours à **l'intelligence artificielle**, notamment à travers les arbres de décision, a facilité l'interprétation des données médicales et la génération de prédictions fiables (malade/sain), tout en restant compréhensible pour les professionnels de santé. Cette transparence algorithmique est un atout essentiel dans le domaine médical.

Sur le plan de l'interface, le design a été pensé pour être **intuitif**, **professionnel** et **accessible**, permettant au personnel médical de saisir rapidement les données

cliniques, de consulter les prédictions, et d'assurer le suivi des patients. L'ajout du calcul automatique du **GFR** et de l'identification du **stade de la MRC** renforce la précision du diagnostic.

Cependant, certaines limites sont à souligner. Le modèle repose sur un ensemble de données limité, ce qui peut influencer la généralisation des résultats. De plus, l'intégration d'autres algorithmes (comme le Random Forest ou le SVM) pourrait améliorer les performances prédictives. Une extension vers des bases de données centralisées et sécurisées représenterait également un atout pour une utilisation en milieu clinique réel.

3.8 Conclusion

Dans ce chapitre, nous avons décrit le développement d'un système intelligent d'aide au diagnostic des maladies rénales. L'application a été conçue dans un environnement Python, en s'appuyant sur le langage PyQt6 pour l'interface graphique, et sur des algorithmes d'intelligence artificielle pour la prédiction médicale. Grâce à une interface conviviale et un moteur de décision performant, le système permet une prise en charge efficace des patients et constitue une base évolutive pour des améliorations futures.

Références

- [1] Laurène Mazeau. Intelligence artificielle et responsabilité civile : Le cas des logiciels d'aide à la décision en matière médicale. *Revue pratique de la prospective et de l'innovation*, 2018, 1, pp.38–43. fihal-01852548f
- [2] Darmoni, S., Griffon, N. et Massari, P. (2014). Les systèmes d'aide à la décision médicale. *Annales des Mines - Réalités industrielles*, Novembre 2014(4), 47-50. <https://doi.org/10.3917/rindu.144.0047>.
- [3] Bernelin, M. et Desmoulin-Canselier, S. (2021). Chapitre 2. L'intelligibilité des algorithmes dans les systèmes d'aide à la décision médicale. *Journal international de bioéthique et d'éthique des sciences*, . 32(2), 19-31. <https://doi.org/10.3917/jibes.322.0019>.
- [4] Hall KK, Fitall E. Artificial Intelligence and Diagnostic Errors. PSNet [internet]. Rockville (MD): Agency for Healthcare Research and Quality, US Department of Health and Human Services. 2020.
- [5] Séroussi, B., et Bouaud, J. Systèmes informatiques d'aide à la décision en médecine: panorama des approches utilisant les données et les connaissances. *Pratique Neurologique-FMC*, 2014, vol. 5, no 4, p. 303-316.
- [6] Moreno, M. Développement des systèmes d'aide à la décision dans les cabinets de médecine générale en France. Thèse pour obtenir le grade de Docteur en Médecine. <https://petale.univlorraine.fr/notice/view/univ-lorraine-ori-34885>
- [7] Lamine Benmimoune. Une approche pour la conception de systèmes d'aide à la décision médicale basés sur un raisonnement mixte à base de connaissance. *Ingénierie assistée par ordinateur*. Université de Technologie de Belfort-Montbéliard, 2016. Français. ffNNT : 2016BELF0307ff. fftel-02003476f
- [8] Chadia Ed-Driouch. Dialogue humain machine pour l'aide à la décision médicale. *Intelligence artificielle [cs.AI]*. Nantes Université, 2023. Français. ffNNT : 2023NANU4027ff. fftel-04334708f
- [9] Reed T. Sutton, David Pincock, Daniel C. Baumgart, Daniel C. Sadowski, Richard N. Fedorak, and Karen I. Kroeker. An overview of clinical decision support systems : benefits, risks, and strategies for success. *npj Digital Medicine*, 3(1) :1–10, 2 2020. ISSN 2398-6352. doi : 10.1038/s41746-020-0221-y.
- [10] A. T. M. Wasylewicz and A. M. J. W. Scheepers-Hoeks. Clinical decision support systems. In Pieter Kubben, Michel Dumontier, and Andre Dekker, editors, *Fundamentals of Clinical Data Science*. Springer, Cham (CH), 2019. ISBN 978-3-319-99712-4.
- [11]] Peng Zhang. Chapter 13 - human–machine interfaces. In Peng Zhang, editor, *Advanced Industrial Control Technology*, pages 527–555. William Andrew Publishing, Oxford, 1 2010. ISBN 978-1-4377-7807-6.
- [12] Evangelia Tsolaki, Patricia Svolos, Evanthia Kousi, Eftychia Kapsalaki, Ioannis Fezoulidis, Konstantinos Fountas, Kyriaki Theodorou, Constantine Kappas, and Ioannis Tsougos. Fast spectroscopic multiple analysis (fasma) for brain tumor classification : a clinical decision support system utilizing multi-parametric 3t mr data. *International Journal of Computer Assisted Radiology and Surgery*, 10(7) :1149–1166, 7 2015. ISSN 1861-6429. doi : 10.1007/s11548-014-1088-7.
- [13] Chadia Ed-Driouch, Florent Chéneau, Françoise Simon, Guillaume Pasquier, Benoit Combès, Anne Kerbrat, Emmanuelle Le Page, Sophie Limou, Nicolas Vince, David-Axel Laplaud, Franck Mars, Cédric Dumas, Gilles Edan, and Pierre-Antoine Gourraud. Multiple sclerosis clinical decision support system based on projection to reference datasets. *Annals of Clinical and Translational Neurology*, 9(12) :1863–1873, 12 2022. doi : 10.1002/acn3.51649.

- [14] Aron Simon Downie, Mark Hancock, Christina Abdel Shaheed, Andrew J. McLachlan, Ahmet Baki Kocaballi, Christopher M. Williams, Zoe A. Michaleff, and Chris G. Maher. An electronic clinical decision support system for the management of low back pain in community pharmacy : Development and mixed methods feasibility study. *JMIR medical informatics*, 8(5) :e17203, 5 2020. ISSN 2291-9694. doi : 10.2196/17203.
- [15] Guangyao Wu, Pei Yang, Yuanliang Xie, Henry C. Woodruff, and al. Development of a clinical decision support system for severity risk prediction and triage of covid19 patients at hospital admission : an international multicentre study. *The European Respiratory Journal*, 56(2) :2001104, 8 2020. ISSN 1399-3003. doi : 10.1183/13993003.01104-2020.
- [16] Sandrine Fosse Etorh. La surveillance du diabète en France : un outil d'aide à la décision et à l'évaluation des politiques publiques de prévention et de promotion de la santé. *Médecine humaine et pathologie*. Université Paris-Est, 2022. Français. ffNNT : 2022PESCO015ff. fftel-04391505f
- [17] Merzoug <https://theses-algerie.com/1888968805074810/memoire-de-master/universite-larbi-tebessi-tebessa/un-systeme-automatique-daide-a-la-detection-de-la-retinopathie-diabetique-base-sur-une-approche-dapprentissage-supervise>
- [18] <https://dspace.univ-adrar.edu.dz/jspui/handle/123456789/8864>
- [19] Zinnour, Lahcene (2023). *Évaluation et développement de service (API) d'aide à la décision clinique à l'aide de réseaux de neurones bayésiens*. Mémoire. Trois-Rivières, Université du Québec à Trois-Rivières, 159 p.
- [20] Ahmed, M. S. (2016). Fouille de données pour l'analyse des traces patients. Thèse de doctorat, Université Oran 1. <https://dspace.univ-oran1.dz/handle/123456789/2748>
- [21] ZEMOULI, M. (2023). Un système intelligent pour améliorer la prédiction des maladies cardiovasculaires. Thèse de doctorat, Université Guelma, <http://dspace.univ-guelma.dz/jspui/handle/123456789/15053>
- [22] Gangani Dharmarathne, Madhusha Bogahawaththa, Marion McAfee, Upaka Rathnayake, D.P.P. Meddage (2024) On the diagnosis of chronic kidney disease using a machine learning-based interface with explainable artificial intelligence, *Intelligent Systems with Applications*, Volume 22, 200397, ISSN 2667-3053, <https://doi.org/10.1016/j.iswa.2024.200397>.
- [23] Jiqian Liu, Bingzhen Sun, Jin Ye, Xixuan Zhao, Xiaoli Chu (2025), Hybrid deep-learning prediction model based on kernel multi-granularity fuzzy rough sets and its application in the diagnosis and treatment of chronic kidney disease, *Engineering Applications of Artificial Intelligence*, Volume 147, <https://doi.org/10.1016/j.engappai.2025.110297>.
- [24] Mansoul, A.. (2020). Aide à la décision médicale guidée par un processus d'extraction de connaissances. 10.13140/RG.2.2.22765.54249.
- [25] Das, T.K. Intelligent Techniques in Decision Making: A Survey. *Indian Journal of Science and Technology*, Vol 9, N 12, 2016.
- [26] N. Garcelon, Des données médicales à la connaissance : entrepôts et fouilles de données, *Annales de Dermatologie et de Vénéréologie*, Volume 142, Issue 12, Supplement, 2015, Pages S389-S390, ISSN 0151-9638, <https://doi.org/10.1016/j.annder.2015.10.171>. (<https://www.sciencedirect.com/science/article/pii/S015196381500719X>)
- [27] D. A. Zighed, Y. Kodratoff, and A Napoli, "Extraction de connaissance à partir d'une base de donnée," *Bulletin AFIA'01*, 2001.
- [28] Hand, D.J. (1998) Data mining: statistics and more ?, *The American Statistician*, 52, 112-118
- [29] Lajnef, M.-A., Ben Ayed, M., Kolski, C. (2005). Convergence possible des processus du data mining et de conception-évaluation d'Ihm : Adaptation du modèle en U, *IHM '05: Proceedings of the 17th Conference on l'Interaction Homme-Machine*, 243-6. <https://doi.org/10.1145/1148550.1148587> (PDF) *KNOWLEDGE EXTRACTION TECHNIQUES*

- FOR POWER TRANSFORMER MAINTENANCE DATA: REVIEW. Available from: https://www.researchgate.net/publication/375132081_KNOWLEDGE_EXTRACTION_TECHNIQUES_FOR_POWER_TRANSFORMER_MAINTENANCE_DATA_REVIEW [accessed Apr 30 2025].
- [30] Robardet, C. Data mining, available at: <http://prisma.insa-lyon.fr/~crobardet>
- [31] Mille A., Fuchs B. Et Herbeaux O., (1996) A unifying framework for Adaptation in Case-Based reasoning. In A. Voss, Ed., Proceeding of the ECAI'96 Workshop : Adaptation in Case-Based Reasoning, p. 22-28
- [32] Boukhobza, M., & Mimi, M. (2012). Détection automatique de la présence d'anomalie sur une mammographie par réseau de neurones artificiels.
- [33] Touzet, C. (1992). *les réseaux de neurones artificiels, introduction au connexionnisme*. Ec2.
- [34] Parizeau, M. (2004). Réseaux de neurones. *GIF-21140 et GIF-64326*, 124.
- [35] A. Balcerac, B. Tervil, N. Vayatis, D. Ricard, Principes fondamentaux de l'apprentissage automatique pour les neurologues, Pratique Neurologique – FMC, Volume 14, Issue 4, 2023, Pages 225-236, ISSN 1878-7762, <https://doi.org/10.1016/j.praneu.2023.10.005>.
- [36] Bachtarzi-Taleb, S., Selmi, H., & Chergui, S. (2015). *Insuffisance rénale en Algérie : Stratégies pour une meilleure prise en charge*. Journal Algérien de Médecine, 23(2), 61–68. <https://www.asjp.cerist.dz/en/article/102403>
- [37] Stevens LA, Coresh J, Greene T, Levey AS. Assessing kidney function—measured and estimated glomerular filtration rate. *N Engl J Med*. 2006; 354: 2473-83
- [38] Bénédicte Stengel, Cécile Couchoud, Catherine Helmer, Carole Loos-Ayav, Michèle Kessler, Épidémiologie de l'insuffisance rénale chronique en France, La Presse Médicale, Volume 36, Issue 12, Part 2, 2007, Pages 1811-1821, ISSN 0755-4982, <https://doi.org/10.1016/j.lpm.2007.04.039>.
- [39] Kidney Disease: Improving Global Outcomes (KDIGO). (2022). KDIGO 2022 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International Supplements*, 12(1), 1–115. <https://doi.org/10.1016/j.kisu.2022.02.001>
- [40] Kalantar-Zadeh, K., & Fouque, D. (2021). Nutritional management of chronic kidney disease. *The New England Journal of Medicine*, 385(18), 1761–1772. <https://doi.org/10.1056/NEJMra2026849>
- [41] Levin, A., Tonelli, M., Bonventre, J., Coresh, J., Donner, J.-A., Fogo, A. B., & Nally, J. V. (2017). Global kidney health 2017 and beyond: A roadmap for closing gaps in care, research, and policy. *The Lancet*, 390(10105), 1888–1917. [https://doi.org/10.1016/S0140-6736\(17\)30788-2](https://doi.org/10.1016/S0140-6736(17)30788-2)
- [42] Object Management Group. (2017). *OMG Unified Modeling Language (OMG UML), version 2.5.1*. <https://www.omg.org/spec/UML/2.5.1/>
- [43] Van Rossum, G., & Drake Jr, F. L. (2009). *Python 3 Reference Manual*. Create Space. Disponible via: <https://docs.python.org/3/>
- [44] Lutz, M. (2013). *Learning Python* (5th Edition), O'Reilly Media. ISBN: 9781449355739.
- [45] Riverbank Computing. *PyQt6 Documentation*. <https://www.riverbankcomputing.com/static/Docs/PyQt6/>
- [46] Hipp, D. R. (2024). *SQLite Documentation*. Lien : <https://www.sqlite.org/docs.html>
- [47] Microsoft. *Visual Studio Code Documentation*. <https://code.visualstudio.com/docs>
- [48] DB Browser for SQLite Team. (2024). *Official Documentation*. <https://sqlitebrowser.org/>
- [49] W3C. (2024). *HTML & CSS Standards*. World Wide Web Consortium. <https://www.w3.org/standards/>

- [50] Tom Mitchell. *Machine Learning*. McGraw Hill, 1997.
- [51] Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd Ed.). O'Reilly Media.
- [52] Breiman et al. (1984). *Classification and Regression Trees*: ouvrage fondamental détaillant l'algorithme CART.
- [53] *Pre-processed Chronic Kidney Disease Dataset*.
- [54] Quinlan, J. R. (1986). *Induction of decision trees*. *Machine Learning*, 1(1), 81–106. DOI: <https://doi.org/10.1007/BF00116251>
- [55] Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2006). *Machine learning: A review of classification and combining techniques*. *Artificial Intelligence Review*, 26(3), 159–190.
- [56] KDIGO. (2020). *Clinical Practice Guideline for the Evaluation and Management of Chronic* <https://kdigo.org/guidelines/ckd-evaluation-and-management/>
- [57] Sharma, S., Saini, R., & Kumar, D. (2020). *Prediction of chronic kidney disease using machine learning algorithms*. *Materials Today: Proceedings*, 33, 4256–4261. <https://doi.org/10.1016/j.matpr.2020.04.631>
- [58] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer.
- [59] Nguyen, H., Cooper, E., & Kamei, K. (2011). *Handling class imbalance in classification problems*. In *Proceedings of the International Conference on Advanced Data Mining and Applications*.
- [60] Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.
- [61] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). *Scikit-learn: Machine learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830.
- [62] Quinlan, J. R. (1986). *Induction of decision trees*. *Machine Learning*, 1(1), 81–106.
- [63] Shannon, C. E. (1948). *A mathematical theory of communication*. *Bell System Technical Journal*.
- [64] Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions*. In *Advances in Neural Information Processing Systems (NeurIPS)*.
- [65] Pedregosa, F., et al. (2011). *Scikit-learn: Machine learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830.
- [66] Chicco, D., & Jurman, G. (2020). *The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation*. *BMC Genomics*, 21(1), 6. <https://doi.org/10.1186/s12864-019-6413-7>
- [67] Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2e édition). O'Reilly Media.
- [68] Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. <https://christophm.github.io/interpretable-ml-book/>
- [69] Quinlan, J. R. (1986). *Induction of decision trees*. *Machine Learning*, 1(1), 81–106. Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830.
- [70] Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.

ملخص

في إطار الصحة الإلكترونية، تقدم هذه الدراسة نظامًا ذكيًا لدعم تشخيص أمراض الكلى، يجمع بين الاستدلال بالحالات السابقة وخوارزميات الذكاء الاصطناعي. الهدف هو تصنيف مراحل المرض الكلوي بدقة وسرعة، مع ضمان أمن البيانات والأرشفة الرقمية. تُسهّل المنصة عمل موظفي الاستقبال والمرضى وأطباء الكلى من خلال واجهات سهلة الاستخدام، خطط المراقبة، والتنبؤات الآلية. تُظهر النتائج نموذجًا موثوقًا يحسّن التشخيص المبكر وإدارة المرضى في أقصر وقت ممكن.

Abstract

In the context of e-health, this thesis proposes an intelligent system to help diagnose kidney disease, combining case-based reasoning (CBR) and artificial intelligence (AI) algorithms. The aim is to quickly and accurately determine the stage of kidney disease while guaranteeing data security and digital archiving. The platform supports receptionists, nurses, and nephrologists with intuitive interfaces, monitoring plans, and automated predictions. The results demonstrate a reliable model, optimizing early diagnosis and patient management in the shortest possible time.

Résumé

Dans le cadre de la e-santé, ce mémoire propose un système intelligent d'aide au diagnostic des maladies rénales, combinant le **raisonnement à partir de cas** (RàPC) et des **algorithmes d'intelligence artificielle** (IA). L'objectif est de préciser rapidement et avec exactitude le stade de la maladie rénale, tout en garantissant la **sécurité des données** et l'**archivage numérique**. La plateforme accompagne les réceptionnistes, infirmiers et néphrologues grâce à des interfaces intuitives, des plans de surveillance et des prédictions automatisées. Les résultats démontrent un modèle fiable, optimisant le diagnostic précoce et la prise en charge des patients dans un temps réduit.