



République Algérienne Démocratique et Populaire



Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Akli Mohand Oulhadj de Bouira

Faculté des Sciences et des Sciences Appliquées

Département d'Informatique

# Mémoire de Master

## en informatique

*Spécialité : Génie des systèmes informatiques*

*Ingénierie des systèmes d'information et des logiciels*

## Thème

---

Solution Big Data pour le domaine de la santé en  
Algérie

---

Encadré par

— MR BOUDJELABA HAKIM

Réalisé par

— BOUBOU Naim

— AYACHE Soumia

2022/2023

# *Remerciements*

Nous tenons à exprimer toute nos reconnaissance à notre encadrant , Monsieur Hakim Boudjelaba. Nous le remercions de nous avoir encadré, orienté, aidé ,conseillé, et aussi pour sa disponibilité inconditionné .

Nos vifs remerciements iront aussi aux membres de jury qui nous ont fait l'honneur de juger notre modeste travail.

Nos remerciements sont destinés de même à, tous nos enseignants du département informatique de l'université Akli Mohand Oulhadj Bouira, qui ont contribué à notre formation. Nous remercions tous nos amis (es) et camarades de promotion pour leurs aides, leurs remarques et leurs critiques qui nous ont fait évoluer.

# *Dédicaces*

Ce modeste mémoire est dédié à :

Tous les membres de nos précieuses familles.

Nous exprimons notre gratitude envers tous nos professeurs pour leur soutien ininterrompu et leurs conseils précieux tout au long de notre parcours académique.

À tous ceux qui nous ont aidés, nous sommes extrêmement reconnaissants.

*Naim , Soumia*

# Table des matières

<b>Table des matières</b>	<b>i</b>
<b>Table des figures</b>	<b>iv</b>
<b>Liste des tableaux</b>	<b>vii</b>
<b>Liste des abréviations</b>	<b>viii</b>
<b>Introduction générale</b>	<b>1</b>
<b>1 Big Data et la santé</b>	<b>2</b>
1.1 Introduction . . . . .	2
1.2 Historique du Big Data . . . . .	2
1.3 Définition du Big Data . . . . .	3
1.3.1 Caractéristiques du Big Data . . . . .	3
1.3.2 Fonctionnement du Big Data . . . . .	5
1.3.3 Cas d’usage du Big Data . . . . .	6
1.3.4 La sécurité dans le Big Data . . . . .	7
1.3.5 Les avantages du Big data . . . . .	8
1.3.6 Les Défis du Big Data . . . . .	9
1.3.7 Les inconvénients du Big Data . . . . .	9
1.4 Big Data et la santé . . . . .	10
1.4.1 Les applications du Big Data dans le domaine de la santé . . . . .	10
1.5 La santé en algérie . . . . .	11
1.5.1 Problématique . . . . .	12

1.5.2 Solutions proposées . . . . .	12
1.6 Conclusion . . . . .	13
<b>2 Etude préalable</b>	<b>14</b>
2.1 Introduction . . . . .	14
2.2 Définition de l'architecture du Big Data . . . . .	14
2.3 Les types d'architectures de Big Data . . . . .	14
2.3.1 L'architecture Lambda . . . . .	15
2.3.2 L'architecture Kappa . . . . .	15
2.4 Les principales technologies de Big Data . . . . .	16
2.4.1 Les technologies de stockage . . . . .	16
2.4.2 Apache Cassandra . . . . .	17
2.4.3 MongoDB . . . . .	19
2.4.4 Les technologies de traitement . . . . .	21
2.4.5 Hadoop . . . . .	21
2.4.6 Spark . . . . .	24
2.5 Intelligence artificielle, machine learning et le deep learning . . . . .	26
2.6 L'apprentissage automatique(Machine learning) . . . . .	27
2.6.1 Définition . . . . .	27
2.6.2 Domaine d'application du machine Learning . . . . .	27
2.6.3 Principe de fonctionnement de machine learning . . . . .	28
2.6.4 l'objectif de machine learning dans le domaine de la santé . . . . .	29
2.7 l'apprentissage en profondeur(Deep learning) . . . . .	29
2.7.1 Définition . . . . .	29
2.7.2 Domaine d'application du deep Learning . . . . .	29
2.7.3 Principe de fonctionnement de deep learning . . . . .	30
2.7.4 L'objectif de deep learning dans le domaine de la santé . . . . .	31
2.8 Comparaison entre machine learning et deep learning . . . . .	32
2.9 Conclusion . . . . .	33
<b>3 Analyse et conception</b>	<b>34</b>
3.1 Introduction . . . . .	34
3.2 Les diagramme utilisés . . . . .	34

3.2.1	Définition de l'UML . . . . .	34
3.2.2	Diagramme d'architecture de système . . . . .	35
3.2.3	Diagramme des acteurs . . . . .	35
3.2.4	Diagramme de cas d'utilisation . . . . .	36
3.2.5	Diagramme de séquence . . . . .	41
3.3	Conclusion . . . . .	49
<b>4</b>	<b>Réalisation</b>	<b>50</b>
4.1	Introduction . . . . .	50
4.2	Environnement de développement . . . . .	50
4.2.1	Environnement matériel . . . . .	50
4.2.2	Environnement logiciel . . . . .	51
4.2.3	Les Schema utilisés . . . . .	54
4.3	Implémentation . . . . .	55
4.3.1	Démonstration des interfaces . . . . .	55
4.4	Conclusion . . . . .	80
	<b>Conclusion générale</b>	<b>81</b>
	<b>Bibliographie</b>	<b>82</b>
	<b>Annexe 1</b>	<b>84</b>
	<b>Annexe 2</b>	<b>85</b>

# Table des figures

1.1	Les caractéristiques du Big Data . . . . .	4
2.1	L'architecture Lambda . . . . .	15
2.2	L'architecture Kappa . . . . .	16
2.3	L'architecture de Cassandra . . . . .	18
2.4	L'architecture de MongoDB . . . . .	20
2.5	L'architecture de hadoop . . . . .	23
2.6	L'architecture Apache spark . . . . .	25
3.1	Diagramme d'architecture de système . . . . .	35
3.2	Diagramme des acteurs . . . . .	37
3.3	Rôle de super admin . . . . .	37
3.4	Rôle de L'admin . . . . .	38
3.5	Rôle de medcin . . . . .	38
3.6	Rôle de réceptionniste . . . . .	39
3.7	Rôle technicien de laboratoire . . . . .	39
3.8	Rôle de Radiologue . . . . .	40
3.9	Rôle de L'infirmier . . . . .	40
3.10	L'Authentication . . . . .	41
3.11	Ajouter Modifier l'utilisateur . . . . .	42
3.12	Supprimer l'utilisateur . . . . .	42
3.13	Ajouter et modifier symptomes,diagnostique et traitement . . . . .	43
3.14	Supprimer symptomes,diagnostique et traitement . . . . .	43
3.15	Ajouter et Modifier analyse . . . . .	44

---

3.16 Supprimer analyse . . . . .	44
3.17 Ajouter Modifier radio . . . . .	45
3.18 Supprimer radio . . . . .	45
3.19 Ajouter & Modifier soins . . . . .	46
3.20 Supprimer soins . . . . .	46
3.21 Ajouter et modifier un patient . . . . .	47
3.22 Supprimer patient . . . . .	47
3.23 La recherche . . . . .	48
3.24 Changer le mot de passe . . . . .	48
3.25 insertion de radiographies et analyse par l'état privé . . . . .	49
4.1 schéma . . . . .	54
4.2 Page connexion . . . . .	55
4.3 Espace SuperAdmin . . . . .	55
4.4 Ajouter un nouvel hôpital . . . . .	56
4.5 Liste des admins des hôpitaux . . . . .	56
4.6 Les étapes visualisation et de téléchargement des analyse . . . . .	57
4.7 Les étapes visualisation et de téléchargement des radiographies . . . . .	58
4.8 Espace admin . . . . .	59
4.9 Ajouter un médecin par l'admin . . . . .	59
4.10 Ajouter un service par l'admin . . . . .	60
4.11 Liste des service . . . . .	60
4.12 Espace médecin . . . . .	61
4.13 Liste des patients en visite . . . . .	61
4.14 Ajouter des diagnostics et symptômes . . . . .	62
4.15 Liste des patients en attente . . . . .	62
4.16 le traitement suivant (continue le traitement) . . . . .	63
4.17 Liste des patients ayant terminé leur traitement . . . . .	63
4.18 Les choix Selectionement des types de treatment . . . . .	64
4.19 Espace radiologue . . . . .	65
4.20 Liste des patients concernés par la radio . . . . .	65
4.21 Ajouter une radio . . . . .	66
4.22 Espace infirmier . . . . .	66

4.23	Liste des patients concernés par un soin . . . . .	67
4.24	Faire le soins par l’infirmier . . . . .	67
4.25	Liste des patients ayant terminé leur soins . . . . .	68
4.26	Espace thechnicien de laboratoire . . . . .	68
4.27	Liste des patients concernés par l’analyse . . . . .	69
4.28	Ajoute et Modifier l’analyse . . . . .	69
4.29	Liste des patients ayant terminé leur analyses . . . . .	70
4.30	Espace respionniste . . . . .	70
4.31	Ajouter patients . . . . .	71
4.32	Liste des patients ayant terminé leur traitement par le médcin . . . . .	71
4.33	Modifier les informations des patients . . . . .	72
4.34	page de profil des users . . . . .	72
4.35	Changer le mot de passe . . . . .	73
4.36	Login pour des cliniques privé . . . . .	74
4.37	Liste des utilisateurs . . . . .	74
4.38	scanner le qr-code . . . . .	75
4.39	Ajouter des analyses . . . . .	75
4.40	Scanner le qr-code . . . . .	76
4.41	Uploader des radiographies . . . . .	76
4.42	Lancer l’exécution pour créer le modèle . . . . .	77
4.43	Fin d’exécution . . . . .	77
4.44	la première étape . . . . .	78
4.45	la deuxième étape . . . . .	78
4.46	la troisième étape . . . . .	78
4.47	lancer l’exécution pour créer les modèles . . . . .	79
4.48	Fin l’exécution pour créer le modèle . . . . .	79
4.49	Les étapes des tests Diabétique avec DEEP LEARNING . . . . .	80
50	Liste des analyses . . . . .	84
51	Liste des analyses . . . . .	85

# Liste des tableaux

- 2.1 Les technologies NoSQL . . . . . 17
- 2.2 Les avantages et les inconvénients de cassandra . . . . . 19
- 2.3 Les avantages et les inconvénients de mongoDB . . . . . 21
- 2.4 Les avantages et les inconvénients de hadoop . . . . . 24
- 2.5 Les avantages et les inconvénients de spark . . . . . 26
- 2.6 Comparaison entre machine learning et deep learning . . . . . 32
  
- 3.1 Description des acteurs . . . . . 36
  
- 4.1 Environnement matériel . . . . . 51

# Liste des abréviations

API	Application Programming Interface
ETL	extraire transformer charger)
DES	dossiers de santé électroniques
IA	Intelligence Artificielle
HDFS	Hadoop Distributed File System
RDDs	Resilient Distributed Datasets
UML	Unified Modeling Language
MERN Stack	MongoDB Express React Node stack

# Introduction générale

Le secteur de la santé est confronté à de nombreux défis, notamment en termes de gestion des données, de prise de décision clinique et d'amélioration des soins de santé en général. Avec l'évolution rapide de la technologie et l'explosion des données médicales, il devient impératif de trouver des solutions innovantes pour relever ces défis. C'est dans ce contexte que les solutions Big Data ont émergé comme un outil puissant pour transformer le domaine de la santé.

Le présent mémoire se concentre sur l'utilisation des solutions Big Data dans le domaine de la santé en Algérie. L'Algérie, en tant que pays en développement, est confrontée à des besoins croissants en matière de santé et à des défis spécifiques liés à la gestion des données et à l'amélioration des soins de santé. Ce mémoire explore comment les solutions Big Data peuvent être utilisées pour relever ces défis et améliorer les services de santé dans le pays.

L'objectif principal de cette étude est d'évaluer les avantages potentiels de l'adoption des solutions Big Data dans le domaine de la santé en Algérie. Nous examinerons en détail les différentes applications des solutions Big Data, telles que la gestion des données médicales, la prise de décision clinique basée sur les données et l'optimisation des ressources de santé. De plus, nous analyserons les défis auxquels l'Algérie est confrontée en matière de confidentialité des données, de formation du personnel de santé et d'infrastructure technologique.

# Big Data et la santé

## 1.1 Introduction

Le Big Data, également connu sous le nom de mégadonnées, fait référence à des ensembles de données caractérisés par leur volume, leur vitesse et leur variété, qui nécessitent l'utilisation de technologies et de méthodes analytiques spécifiques pour les transformer en informations utilisables. Dans le domaine de la santé, les big data englobent toutes les données disponibles concernant la santé dans son sens le plus large, recueillies auprès de diverses sources. Exploiter ces données présente de nombreux avantages, tels que l'acquisition d'une meilleure compréhension du système de soins, l'identification de facteurs de risque de maladies, l'aide au diagnostic, au choix et au suivi de l'efficacité des traitements, la pharmacovigilance, l'épidémiologie, etc. De nos jours, la gestion médico-économique des établissements de santé, les décisions de santé publique et même la recherche biomédicale reposent sur l'exploitation de ces données massives. Cependant, la collecte et l'exploitation de ces données présentent encore de nombreux défis techniques, ainsi que des questions éthiques . [11].

## 1.2 Historique du Big Data

Bien que le concept de Big Data soit un concept relativement récent, l'utilisation de vastes ensembles de données remonte aux années 60 et 70. À cette époque, le domaine des données était encore émergent, marqué par l'établissement des premiers datacenters et le développement des bases de données relationnelles .[2]

## 1.3 Définition du Big Data

Le Big Data fait référence à l'énorme volume de données informatiques qui circulent à travers le monde. Ces données deviennent si massives qu'elles ne peuvent plus être stockées efficacement dans des bases de données de taille standard. Cela conduit à l'utilisation croissante d'outils informatiques performants pour gérer ces données. Le Big Data est caractérisé par trois (voire cinq) principales dimensions, toutes commençant par la lettre "V" : le Volume, qui représente la quantité massive de données générées ; la Vitesse, qui fait référence à la vitesse à laquelle les données sont produites et doivent être traitées ; et la Variété, qui englobe la diversité des types et des sources de données. De plus, la Valeur et la Véracité des données peuvent également être considérées comme des dimensions essentielles du Big Data . [8]

### 1.3.1 Caractéristiques du Big Data

Les principales caractéristiques du Big Data, souvent résumées par les "5V", sont les suivantes : [3] :

#### **Variété**

Le Big Data englobe des données provenant de différentes sources et présentant des formats variés tels que des données numériques, des adresses, des numéros de téléphone, etc. Il est essentiel de les unifier pour faciliter leur analyse.

#### **Volume**

Le Big Data concerne des quantités massives de données. Il permet de collecter et d'analyser de grands ensembles de données comprenant des millions d'enregistrements.

#### **Véracité**

L'objectif du Big Data est de s'appuyer sur des données réelles pour prendre des décisions. Il est important de veiller à la fiabilité et à la qualité des données, en éliminant celles qui ne sont pas valables ou conformes aux normes établies.

## Vitesse

Le flux de données en ligne est souvent rapide et constant. Les données doivent être mises à jour en temps réel pour maintenir leur pertinence et permettre une analyse en temps opportun.

## Valeur

Le Big Data vise à identifier les données les plus pertinentes et utiles pour l'entreprise. Cela implique de sélectionner les enregistrements appropriés et de tenir compte de leurs relations réelles. Les données peuvent apporter de la valeur à l'entreprise en améliorant ses produits, en personnalisant ses services ou en offrant de meilleurs prix.



FIGURE 1.1 – Les caractéristiques du Big Data

### 1.3.2 Fonctionnement du Big Data

Le Big Data offre de nouvelles perspectives qui ouvrent des opportunités et favorisent de nouveaux modèles économiques. Son adoption repose sur trois actions principales :

#### **Intégrer**

Le Big Data rassemble des données provenant de sources et d'applications diverses. Les méthodes d'intégration de données traditionnelles, telles que l'extraction, la transformation et le chargement (ETL), ne sont généralement pas adaptées. Pour analyser des ensembles de données Big Data à l'échelle du téraoctet voire du pétaoctet, de nouvelles stratégies et technologies doivent être adoptées. L'intégration implique l'importation des données, leur traitement et leur mise à disposition dans un format accessible aux analystes.

#### **Gérer**

Le Big Data nécessite du stockage. Vous pouvez choisir de stocker vos données dans le cloud, sur site ou une combinaison des deux. Vous avez la liberté de stocker vos données dans le format de votre choix et d'imposer les exigences de traitement et les moteurs de traitement nécessaires. De nombreux acteurs optent pour une solution de stockage basée sur l'emplacement où leurs données sont hébergées. Le cloud est de plus en plus adopté car il prend en charge les besoins informatiques actuels et offre la possibilité d'augmenter les ressources selon les besoins.

#### **Analyser**

La valeur de votre investissement dans le Big Data réside dans l'analyse et l'action sur vos données. Vous pouvez développer de nouvelles perspectives grâce à une analyse visuelle de vos différents ensembles de données. Explorez plus en profondeur les données pour faire de nouvelles découvertes. Partagez vos conclusions avec d'autres utilisateurs. Créez des modèles de données à l'aide du machine learning et de l'intelligence artificielle. Exploitez pleinement vos données.

### 1.3.3 Cas d'usage du Big Data

Le Big Data peut être utilisé dans de nombreux domaines pour résoudre des problèmes complexes et prendre des décisions éclairées en se basant sur l'analyse de données massives. Voici quelques exemples de cas d'utilisation du Big Data :

#### **Marketing et publicité**

Les entreprises peuvent utiliser le Big Data pour analyser les données des réseaux sociaux, des emails, des publicités en ligne, etc. afin de mieux comprendre les préférences des consommateurs, d'optimiser les campagnes publicitaires, de personnaliser les offres et d'augmenter les ventes.

#### **Santé**

Le Big Data peut aider les professionnels de la santé à analyser les données des dossiers médicaux électroniques, des résultats d'analyse de laboratoire, des images médicales, etc. pour améliorer le diagnostic, la prévention et le traitement des maladies.

#### **Transport**

Les données du trafic routier, des capteurs de véhicules, des applications de navigation, etc. peuvent être utilisées pour optimiser la gestion du trafic, réduire les temps de trajet, améliorer la sécurité routière et prévoir les congestions.

#### **Finance**

Les institutions financières peuvent utiliser le Big Data pour l'analyse des données transactionnelles, la détection des fraudes, la prévision des risques financiers, la segmentation de la clientèle et la personnalisation des offres.

#### **Énergie**

Les données des compteurs intelligents, des capteurs de réseau électrique, des éoliennes, des panneaux solaires, etc. peuvent être utilisées pour optimiser la production d'énergie, réduire les coûts, prévoir les pannes et améliorer la durabilité environnementale.

Ces exemples montrent comment le Big Data peut être utilisé pour améliorer l'efficacité, réduire les coûts, améliorer la qualité et transformer les industries.

### 1.3.4 La sécurité dans le Big Data

La sécurité semble être l'un des principaux obstacles du big data . Avec l'augmentation de la variabilité et la quantité de données échangées, ces barrières n'ont pas vraiment été éliminées. De nombreuses entreprises spécialisées dans la sécurité des données ont vu le jour, mais elles ne se concentrent que sur un ou plusieurs éléments. En revanche, le progrès technologique est plus rapide que le développement de ces solutions. Mais afin de minimiser le risque, vous devez au moins respecter les cinq règles de sécurité suivantes :

#### **Authentification**

La vérification d'identité consiste à garantir l'identité de l'utilisateur, c'est-à-dire à s'assurer que le partenaire de chaque correspondant est bien l'identité qu'il revendique. Afin de protéger le Big Data, il est essentiel d'intégrer un outil permettant d'identifier les personnes ou les processus avant d'autoriser sa connexion. Il existe plusieurs techniques d'authentification qui peuvent être divisées en deux catégories : — L'authentification de base est appelée authentification simple : Qui suis-je : identifiant (login). Preuve-le : Nom d'utilisateur (mot de passe). — L'authentification forte : Renforcez la vérification d'identité simple Pour renforcer la vérification de L'authentification qui n'est pas suffisamment sécurisée, vous devez ajouter des "verrous" : empreinte biométrique.

#### **Autorisation**

Une fois connecté, la personne n'est pas autorisée à effectuer des opérations ou à accéder à toutes les ressources. Plusieurs techniques peuvent être utilisées. Le plus ancien est l'utilisation d'ABAC (Attribute-Based Access Control) : contrôle d'accès basé sur les attributs. Une autre technologie la plus largement utilisée est RBAC (Role-Based Access Control), qui rend possible le contrôle d'accès basé sur les rôles.

#### **Audit**

Bien qu'une personne ait été identifiée, authentifiée et autorisée à accéder à certaines ressources, un système sûr et fiable doit suivre toutes les opérations de cette personne. Il

s'agit de la traçabilité, qui est un processus essentiel.

## Cryptage

Données Afin de répondre efficacement à l'évolution des menaces, une approche centrée sur les données est également nécessaire pour protéger les informations sensibles (telles que les informations de carte de crédit ou les dossiers confidentiels des patients). Vous pouvez également y parvenir en chiffrant certaines données, qui sont stockées en toute sécurité et deviennent difficiles à comprendre en raison de l'utilisation d'algorithmes de chiffrement pour les garder secrètes. Afin de rendre ce processus efficace, la clé de chiffrement doit être strictement protégée [21]. Les quatre règles décrites doivent former la base de votre stratégie Big Data. Si vous ne suivez pas ces règles, vos données seront menacées. C'est pourquoi vous devez suivre strictement ces règles.

## Vie privé

La protection de la vie privée est une préoccupation importante dans l'utilisation de Big Data. Les données stockées dans les systèmes Big Data peuvent inclure des informations personnelles telles que des noms, des adresses, des numéros de téléphone, des numéros de sécurité sociale, des informations médicales, des données financières et des informations de localisation. Si ces données sont mal protégées, cela peut entraîner des violations de la vie privée et des risques pour la sécurité des personnes concernées.

### 1.3.5 Les avantages du Big data

- Le Big Data permet d'obtenir des réponses plus complètes, car le volume d'informations est plus important.
- Des réponses plus complètes signifient plus de confiance dans les données, ce qui signifie une approche complètement différente de la résolution des problèmes.
- L'architecture Big Data peut apporter certains avantages, tels que :  
Évolutivité, Performance, Faible coût, Disponibilité.
- L'architecture Big Data fournit son propre mécanisme de haute disponibilité.[5]

### 1.3.6 Les Défis du Big Data

Le Big Data présente certains défis malgré les perspectives intéressantes qu'il offre.

Tout d'abord, le volume des données constitue un défi majeur. Malgré les avancées technologiques en matière de stockage des données, les volumes de données doublent approximativement tous les deux ans. Les entreprises éprouvent encore des difficultés à gérer leur croissance et à trouver des moyens efficaces de stockage.

Cependant, le simple stockage des données ne suffit pas. Pour être utiles, les données doivent être exploitées et préalablement organisées. Le nettoyage des données ou leur pertinence pour le client, ainsi que leur organisation en vue d'une analyse significative, nécessitent un travail considérable. Les experts en données consacrent entre 50 et 80 % de leur temps à l'organisation et à la préparation des données avant de pouvoir les utiliser.

Enfin, la technologie du Big Data évolue rapidement. Il y a quelques années, Apache Hadoop était la technologie la plus utilisée pour le traitement du Big Data. Ensuite, Apache Spark est apparu en 2014. Actuellement, l'association de ces deux infrastructures semble constituer l'approche la plus prometteuse. La maîtrise de la technologie du Big Data représente donc un défi continu. [4].

### 1.3.7 Les inconvénients du Big Data

Le Big Data est une technologie puissante qui présente de nombreux avantages, mais elle comporte également certains inconvénients. Voici quelques-uns des principaux inconvénients du Big Data :

#### Coût élevé

La mise en place d'une infrastructure de big data peut être très coûteuse.

#### Complexité technique

Les systèmes de big data peuvent être très complexes, avec de nombreuses technologies différentes à maîtriser.

## Risques de sécurité

Le big data implique la collecte, le stockage et l'analyse de vastes quantités de données, ce qui peut présenter des risques de sécurité importants.

## Protection des données

Le big data peut également soulever des problèmes de confidentialité et de protection des données.

## Éthique et responsabilité

Enfin, le big data soulève des questions éthiques et de responsabilité.

# 1.4 Big Data et la santé

Dans le domaine de la santé, les big data, également appelées données massives, regroupent toutes les informations disponibles sur la santé dans son ensemble, provenant de diverses sources. L'exploitation de ces données présente de multiples avantages : une meilleure compréhension du système de soins, l'identification de facteurs de risque de maladies, une aide au diagnostic, à la sélection et au suivi de l'efficacité des traitements, la pharmacovigilance, ainsi que l'épidémiologie. De nos jours, la gestion médico-économique des établissements de santé, les décisions de santé publique et même la recherche biomédicale s'appuient sur l'exploitation de ces données massives. Néanmoins, leur collecte et leur utilisation continuent de présenter de nombreux défis techniques, ainsi que des questions éthiques [6].

### 1.4.1 Les applications du Big Data dans le domaine de la santé

Le Big Data est utilisé dans le domaine de la santé pour améliorer la qualité des soins, la prévention, le diagnostic et le traitement des maladies. L'augmentation exponentielle de la quantité de données non structurées, la capacité de stockage et d'analyse. Voici quelques exemples d'applications du Big Data dans le domaine de la santé :

## **Analyse des données de santé**

Les données des dossiers médicaux électroniques, des résultats d'analyse de laboratoire, des images médicales, des capteurs portables, des réseaux sociaux, etc. peuvent être analysées pour identifier les tendances, les corrélations et les modèles de santé.

## **Prévention des maladies**

Les données du Big Data peuvent être utilisées pour prévoir les risques de maladies et pour prévenir les maladies. Les modèles d'apprentissage automatique peuvent être utilisés pour identifier les patients à haut risque et pour recommander des interventions préventives.

## **Diagnostic des maladies**

Les données du Big Data peuvent être utilisées pour améliorer le diagnostic des maladies en fournissant des informations supplémentaires aux médecins. Les algorithmes d'apprentissage automatique peuvent être utilisés pour analyser les images médicales et aider à identifier les anomalies.

## **Personnalisation des traitements**

Les données du Big Data peuvent être utilisées pour personnaliser les traitements en fonction des caractéristiques individuelles des patients. Les modèles d'apprentissage automatique peuvent être utilisés pour prédire la réponse d'un patient à un traitement spécifique.

## **Gestion des soins de santé**

Les données du Big Data peuvent être utilisées pour améliorer la gestion des soins de santé en fournissant des informations sur les coûts, l'utilisation des services, la qualité des soins et les résultats pour les patients.

## **1.5 La santé en algérie**

Le domaine de la santé en Algérie génère une quantité importante de données, provenant de diverses sources, telles que les hôpitaux, les cliniques, les laboratoires, les centres

de recherche, les programmes de santé publique, les assurances maladie, etc.

Ces données peuvent inclure des informations sur les patients, les fournisseurs de soins de santé, les installations de santé, les médicaments et les dispositifs médicaux, ainsi que d'autres aspects du système de santé.

Ces données peuvent être utilisées pour améliorer la qualité des soins de santé, optimiser l'utilisation des ressources, prévoir les tendances de santé, évaluer l'efficacité des programmes de santé publique, développer de nouveaux traitements et médicaments, et aider les décideurs à prendre des décisions éclairées en matière de santé.

Cependant, en Algérie, il ya des défis en matière de collecte, d'analyse et de partage des données de santé. Il peut y avoir des lacunes dans l'infrastructure de données, des problèmes de confidentialité et de sécurité, et des obstacles juridiques et réglementaires à l'utilisation des données de santé.

### 1.5.1 Problématique

Les applications de santé qui utilisent SQL peuvent rencontrer plusieurs problématiques, notamment :

- Limitation de la scalabilité lorsqu'il s'agit de gérer des volumes massifs de données.
- La Complexité de la gestion des données ce qui peut augmenter la complexité de conception.
- Les données de santé peuvent inclure des informations sensibles cela peut les rendre vulnérables aux violations de la sécurité des données.
- Problème des coûts élevés qui peuvent augmenter avec la taille des données et le nombre d'utilisateurs qui accèdent à la base de données.

### 1.5.2 Solutions proposées

Les hôpitaux peuvent utiliser le Big Data pour améliorer la qualité des soins, optimiser les opérations et réduire les coûts. Voici quelques solutions proposées :

- Collecter et stocker les données des patients dans un système DES(dossiers de santé électroniques),qui permet aux professionnels de santé d'accéder rapidement et facilement aux informations pertinentes. Et qui aider à réduire les erreurs médicales et à améliorer la coordination des soins.
- La Surveillance en temps réel les données des patients permet aux professionnels de santé de détecter rapidement les changements dans l'état du patient et de prendre des mesures appropriées.
- Analyser l'utilisation des ressources, telles que les salles d'opération,. Cela permet aux hôpitaux de planifier les ressources de manière plus efficace et d'optimiser les coûts.
- Suivre la performance de leur personnel, Cela peut aider les hôpitaux à identifier les domaines à améliorer et à mettre en place des mesures correctives pour améliorer la qualité des soins.

## 1.6 Conclusion

La prolifération massive des données numériques a poussé les chercheurs à explorer de nouvelles approches pour observer et analyser le monde qui nous entoure. Cette expansion demande une exploration de nouveaux horizons en ce qui concerne la recherche, le partage, le stockage, l'analyse et la présentation des données. Le concept de Big Data se présente alors comme une solution prometteuse pour résoudre ces problèmes de manière novatrice.

## Etude préalable

### 2.1 Introduction

Après la présentation de quelques généralités sur le big data, ainsi que leurs caractéristiques, avantages et inconvénients, et après avoir défini les problématiques de la santé en algérie nous passerons à la phase d'étude préalable de notre système, là où nous commencerons tous d'abord par définir les types d'architectures du big data. Par la suite, nous présenterons Les principales technologies du Big Data, aussi que les sous-discipline de l'IA .

### 2.2 Définition de l'architecture du Big Data

L'architecture de Big Data est la structure organisationnelle et technique qui permet de stocker, gérer et analyser de grandes quantités de données variées provenant de différentes sources. Cette architecture est conçue pour répondre aux besoins de traitement de données volumineuses, complexes et souvent en temps réel [10].

### 2.3 Les types d'architectures de Big Data

Il existe plusieurs architectures couramment utilisées pour les applications de big data, chacune offrant des avantages et des inconvénients en fonction des besoins spécifiques de l'application. Voici quelques exemples :

### 2.3.1 L'architecture Lambda

L'architecture Lambda est une approche en deux couches qui traite les données en temps réel et en batch. Elle utilise deux pipelines de traitement de données parallèles, un pour le traitement en temps réel et un pour le traitement en batch. Les données en temps réel sont traitées par un flux de données en continu, tandis que les données en batch sont traitées périodiquement en utilisant des travaux de traitement de données en batch. Les résultats des deux pipelines sont combinés pour fournir des données cohérentes [9].

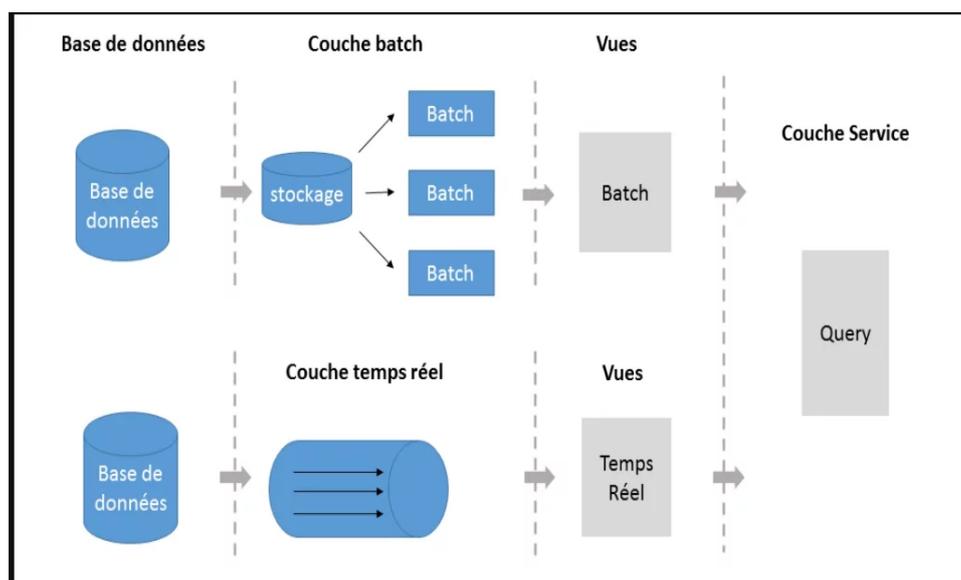


FIGURE 2.1 – L'architecture Lambda

### 2.3.2 L'architecture Kappa

L'architecture Kappa, quant à elle, est une approche en une seule couche qui ne traite que les données en temps réel. Elle utilise un pipeline de traitement de données en continu pour ingérer, traiter et stocker les données en temps réel. Les données sont stockées dans un seul système de stockage, généralement Un système de stockage décentralisé , ce qui permet de traiter les données plus rapidement et plus efficacement [7].

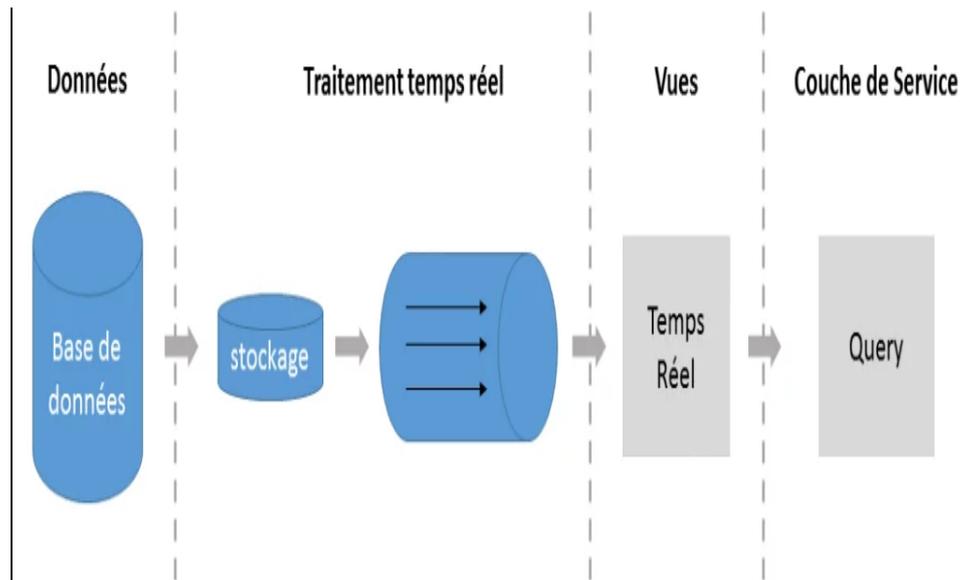


FIGURE 2.2 – L'architecture Kappa

## 2.4 Les principales technologies de Big Data

Il existe différentes options technologiques permettant d'optimiser les temps de traitement des bases de données massives. De manière générale, ces options peuvent être regroupées en deux grandes familles

### 2.4.1 Les technologies de stockage

Type de stockage	Description	Exemples courants
Systèmes de stockage de documents	Stockent des données dans des formats semi-structurés de documents	MongoDB, Couchbase
Systèmes de stockage de colonnes	Organisent les données par colonnes plutôt que par lignes	Apache Cassandra, HBase
Systèmes de stockage de graphes	Permettent le stockage et l'interrogation de données basées sur des relations graphiques	Neo4j, Amazon Neptune
Systèmes de stockage clé-valeur	Stockent les données en utilisant des paires clé-valeur	Redis, Amazon DynamoDB

TABLE 2.1 – Les technologies NoSQL

## 2.4.2 Apache Cassandra

### Definition

Apache Cassandra est une solution de gestion de bases de données NoSQL distribuée et libre, conçue pour traiter efficacement de larges volumes de données répartis sur plusieurs serveurs, sans dépendre d'un point de défaillance unique. Ce système a été initialement développé par Facebook avant d'être publié sous forme de projet open-source par la Fondation Apache. Programmé en Java, Cassandra est renommé pour son aptitude à gérer avec performance et évolutivité d'importantes quantités de données.

### Caractéristiques

Apache Cassandra est une base de données de type NoSQL. distribuée, open-source, haute performance et hautement évolutive. Voici quelques-unes de ses caractéristiques clés :

1. **Architecture distribuée** : Cassandra est conçu pour fonctionner sur des clusters de serveurs, permettant une haute disponibilité et une répartition de charge efficace.
2. **Pas de point de défaillance unique** : Cassandra est conçu pour éviter les points

de défaillance uniques en répartissant les données sur plusieurs nœuds.

3. **Évolutivité linéaire** :Cassandra excelle dans le traitement de volumineuses quantités de données et offre une capacité d'adaptation transparente lors de l'ajout de nouveaux nœuds, sans aucune interruption de service.
4. **Tolérance aux pannes** :Cassandra est conçu pour être résistant aux pannes, ce qui signifie qu'il peut continuer à fonctionner même si certains nœuds du cluster ne sont pas disponibles.

### Architecture de Cassandra

Le fonctionnement central de Cassandra repose sur une configuration où des nœuds indépendants sont interconnectés pour constituer un cluster.

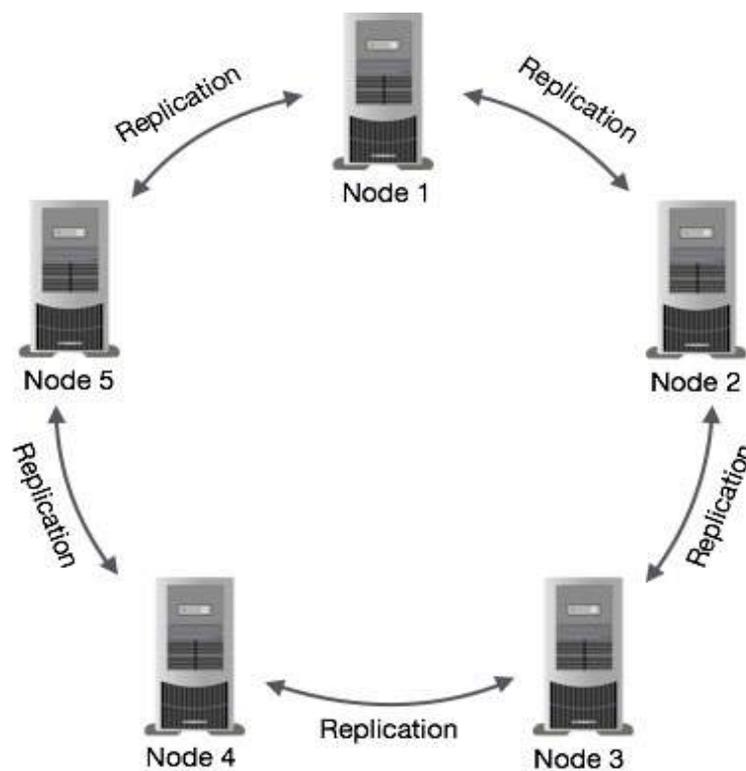


FIGURE 2.3 – L'architecture de Cassandra

## Les avantages & Les inconvénients de Cassandra

Les inconvénients	Les avantages
<ul style="list-style-type: none"> <li>-Compétences spécifiques requises pour la configuration, la maintenance et le développement d'applications.</li> <li>- Modélisation de données plus complexe que dans les bases de données relationnelles.</li> <li>- Coût de stockage élevé en raison de son modèle de stockage basé sur les colonnes.</li> <li>- Coût de stockage élevé en raison de son modèle de stockage basé sur les colonnes.</li> </ul>	<ul style="list-style-type: none"> <li>- Haute évolutivité grâce à sa conception distribuée</li> <li>-Performance élevée grâce à son stockage en mémoire et son écriture optimisée.</li> <li>-Réplication des données pour améliorer la disponibilité et la tolérance aux pannes.</li> <li>-Large écosystème d'outils, de frameworks et de communautés.</li> </ul>

TABLE 2.2 – Les avantages et les inconvénients de cassandra

### 2.4.3 MongoDB

#### Définition

MongoDB est considéré comme la pionnière parmi les bases de données NoSQL orientées document, sans schéma, utilisant le format BSON pour stocker les données. Le BSON, basé sur le JSON, offre une intégration rapide et facile des données avec certains types d'applications, grâce à son format binaire. En plus de sa capacité à évoluer horizontalement, MongoDB n'a pas de point de défaillance unique. Un cluster MongoDB diffère d'un cluster Cassandra en incluant un arbitre, un nœud maître et plusieurs nœuds esclaves. Depuis 2009, MongoDB est un projet open-source détenu par la société 10gen.companny et est sous licence AGPL.

## Caractéristique de MongoDB

Voici les principales caractéristiques de MongoDB :

1. **Modèle de données orienté document** : MongoDB exploite un modèle de données axé sur les documents, ce qui offre une souplesse remarquable dans la gestion des données.
2. **Pas de schéma fixe** : Contrairement aux bases de données relationnelles, MongoDB ne nécessite pas de schéma fixe, ce qui facilite le développement et la maintenance de l'application.
3. **Haute évolutivité** : MongoDB est conçu pour être hautement évolutif, capable de gérer de grandes quantités de données et de s'adapter facilement à l'ajout de nouveaux nœuds..
4. **Haute disponibilité** : MongoDB est conçu pour être hautement disponible, avec des options de réplication et de tolérance de panne intégrées.

## Architecture MongoDB

L'architecture de MongoDB est conçue pour être hautement évolutive, distribuée et résiliente aux pannes. Les éléments clés de l'architecture comprennent les clusters, la réplication, le sharding, l'indexation, un langage de requête flexible, des drivers et API, ainsi que des outils de gestion. MongoDB est capable de gérer de très grands volumes de données et est utilisé dans une grande variété de domaines, y compris les applications web, les applications mobiles, les jeux vidéo et les systèmes de géolocalisation.

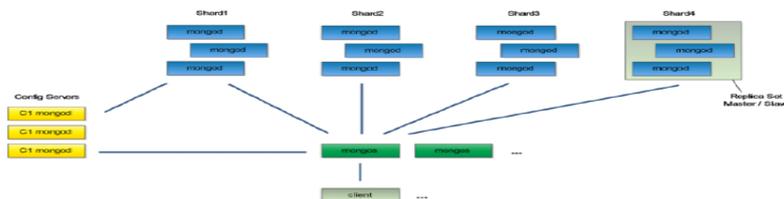


FIGURE 2.4 – L'architecture de MongoDB

## Les avantages et les inconvénients de MongoDB

Les inconvénients	Les avantages
<p>MongoDB, en raison de son manque de support pour les transactions ACID, peut présenter des défis pour certaines applications.</p> <ul style="list-style-type: none"> <li>- La distribution sur plusieurs serveurs peut entraîner des problèmes de cohérence dans MongoDB.</li> <li>- MongoDB nécessite des compétences spécifiques pour travailler efficacement avec la base de données.</li> <li>- MongoDB ne prend pas en charge les jointures, ce qui peut rendre plus difficile l'extraction de données à partir de collections différentes.</li> </ul>	<ul style="list-style-type: none"> <li>- MongoDB est flexible et permet de stocker des données non structurées ou semi-structurées.</li> <li>- MongoDB est évolutif et distribué sur plusieurs serveurs, ce qui permet une meilleure scalabilité et une meilleure tolérance aux pannes.</li> <li>- MongoDB offre de bonnes performances grâce à son architecture distribuée, son indexation efficace et son langage de requête flexible.</li> <li>- MongoDB est compatible avec de nombreux langages de programmation et dispose de nombreux outils pour faciliter le développement et la maintenance des applications.</li> </ul>

TABLE 2.3 – Les avantages et les inconvénients de mongoDB

## 2.4.4 Les technologies de traitement

### 2.4.5 Hadoop

#### Definition

Hadoop est un framework open source conçu spécifiquement pour le stockage, la gestion et le traitement de vastes ensembles de données répartis sur des clusters de serveurs. Il repose sur le système de fichiers distribué HDFS (Hadoop Distributed File System) et tire parti du modèle de programmation MapReduce pour le traitement des données. Les applications d'Hadoop sont diverses, allant de l'analyse de données à l'apprentissage automatique, en passant par la recherche web, la surveillance et la sécurité. À l'origine développé par l'Apache Software Foundation, Hadoop est largement adopté par les grandes

entreprises et les organisations gouvernementales en vue de gérer et d'analyser des volumes considérables de données.

## Caractéristiques

Les principales caractéristiques de Hadoop sont les suivantes :

1. **Stockage distribué** : Hadoop permet de stocker de grandes quantités de données distribuées sur des clusters de serveurs, ce qui permet d'assurer la disponibilité et la tolérance aux pannes des données.
2. **Traitement distribué** : Hadoop utilise le modèle de programmation MapReduce pour traiter de grandes quantités de données distribuées sur des clusters de serveurs. Cela permet d'accélérer considérablement le traitement des données.
3. **Évolutivité horizontale** : Hadoop est conçu pour être évolutif horizontalement, ce qui signifie qu'il peut être facilement étendu pour prendre en charge de plus grandes quantités de données en ajoutant simplement plus de serveurs à un cluster.
4. **Flexibilité** : Hadoop est flexible en ce qui concerne les types de données qu'il peut stocker et traiter, ce qui permet de gérer des données semi-structurées et non structurées.

## L'architecture de hadoop

L'architecture de Hadoop est distribuée et basée sur un modèle maître-esclave, avec un serveur maître (NameNode) qui coordonne la distribution des données et des tâches de traitement sur les différents nœuds esclaves (DataNodes) qui composent le cluster. Les principaux composants de Hadoop sont HDFS, YARN, MapReduce et Hadoop Common.

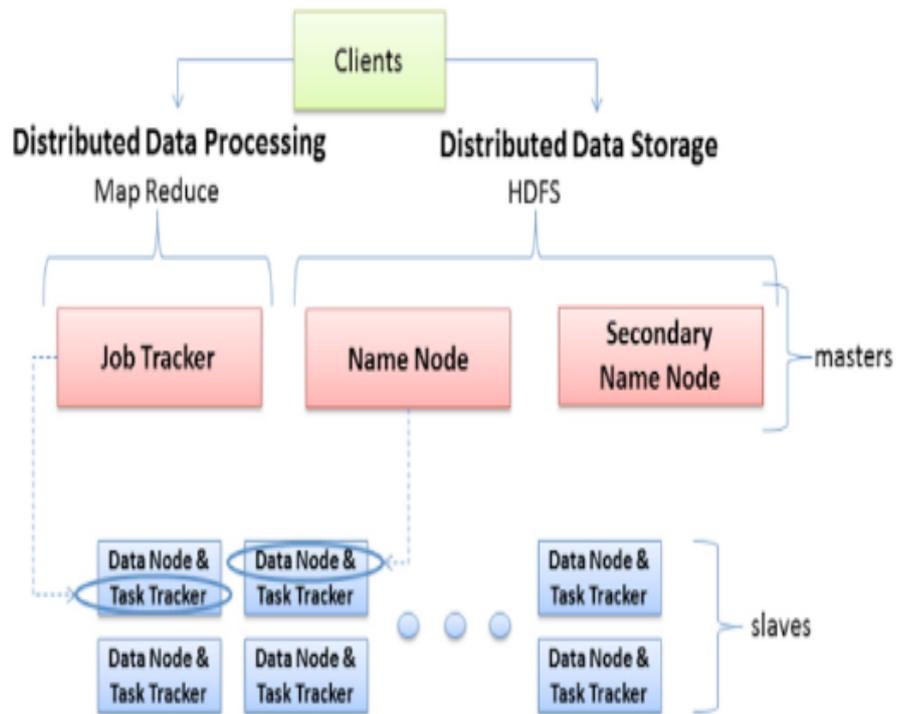


FIGURE 2.5 – L'architecture de hadoop

### Les avantages & Les inconvénients de hadoop

Les inconvénients	Les avantages
<ul style="list-style-type: none"> <li>- Hadoop est un système complexe et nécessite une expertise technique pour être configuré et géré.</li> <li>- Besoins de stockage supplémentaires.</li> <li>- Besoin de compétences en programmation</li> <li>- Problèmes potentiels de sécurité</li> </ul>	<ul style="list-style-type: none"> <li>- Évolutivité</li> <li>- Flexibilité.</li> <li>- Tolérance aux pannes</li> <li>- Coût abordable .</li> <li>- Traitement de données non structurées.</li> <li>- Haute performance.</li> </ul>

TABLE 2.4 – Les avantages et les inconvénients de hadoop

## 2.4.6 Spark

### Definition

Apache Spark est un système de traitement des données en cluster open source conçu pour être rapide, flexible et facile à utiliser. Il fournit une interface de programmation unifiée pour le traitement de données en batch, en temps réel, par flux, et d'autres types de tâches de traitement de données. Spark est souvent utilisé pour le traitement de grands ensembles de données distribuées sur un cluster de serveurs et peut être utilisé avec diverses langues de programmation, telles que Java, Scala, Python et R.

### Caractéristiques

Voici les principales caractéristiques de Spark :

1. **Rapidité** : Spark est conçu pour être très rapide grâce à son architecture de traitement en mémoire, qui permet de réduire considérablement les temps de traitement.
2. **Évolutivité** : Spark est conçu pour s'adapter facilement à l'ajout de nœuds supplémentaires pour traiter de plus grands ensembles de données.
3. **Flexibilité** : Spark prend en charge divers types de tâches de traitement de données, notamment le traitement en batch, en temps réel, par flux, le traitement de graphes, le traitement de données SQL, etc.

4. **interface de programmation unifiée** : Spark fournit une interface de programmation unifiée pour traiter divers types de tâches de traitement de données.

### L'architecture de spark

Spark est une plate-forme de traitement de données distribuée qui utilise une architecture en cluster pour le traitement parallèle des données. Son architecture est basée sur le concept de Resilient Distributed Datasets (RDDs) qui sont des ensembles de données immuables et distribués sur un cluster de machines. Les nœuds de travail exécutent les tâches de traitement de données et stockent les RDDs en mémoire ou sur le disque. Les transformations sont des opérations qui sont appliquées aux RDDs pour les transformer en un autre ensemble de données. Les actions sont des opérations qui sont appliquées aux RDDs pour renvoyer les résultats au programme principal. En utilisant cette architecture, Spark est capable de traiter des quantités massives de données de manière efficace et rapide, ce qui en fait une plate-forme populaire pour les analyses de données et le traitement de données distribuées.

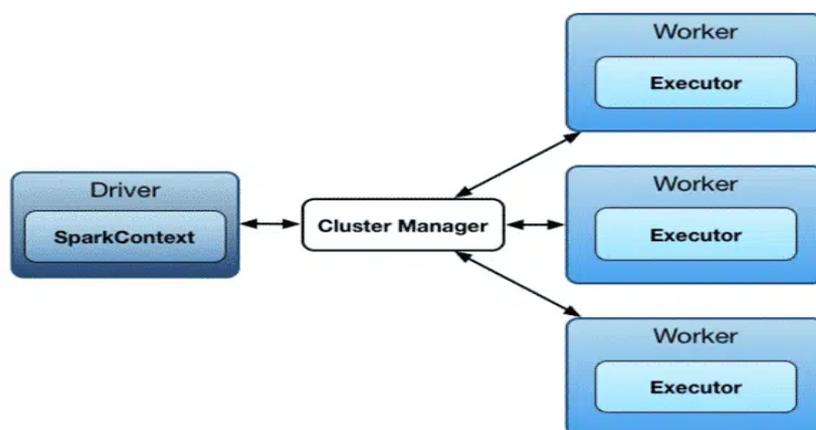


FIGURE 2.6 – L'architecture Apache spark

## Les avantages & Les inconvénients de spark

Les inconvénients	Les avantages
Besoin de compétences en programmation	Traitement distribué
Gestion de la mémoire	Flexibilité
Besoin d'un environnement de cluster	Traitement distribué .
Complexité	Polyvalence
Dépendance à la distribution Hadoop	Évolutivité.
Besoin d'un volume de données important.	Facilité d'utilisation.

TABLE 2.5 – Les avantages et les inconvénients de spark

## 2.5 Intelligence artificielle, machine learning et le deep learning

L'intelligence artificielle (IA) représente un domaine informatique qui se focalise sur la création de systèmes capables d'accomplir des tâches qui nécessitent habituellement l'intelligence humaine. Le machine learning, quant à lui, se situe dans le cadre plus large de l'IA et utilise des données pour former des modèles mathématiques afin de prendre des décisions ou effectuer des prédictions. Il existe deux formes principales de machine learning : supervisé et non supervisé. Parallèlement, le deep learning, une technique avancée du machine learning, exploite des réseaux de neurones artificiels profonds pour apprendre à partir des données. Les réseaux de neurones artificiels sont des modèles mathématiques qui reproduisent le fonctionnement du cerveau humain. Grâce au deep learning, des progrès significatifs ont été réalisés dans des domaines tels que la reconnaissance d'images, la reconnaissance vocale et la traduction automatique. En résumé, l'IA englobe le machine learning en tant que sous-discipline, et le deep learning, en tant que technique avancée du machine learning, utilise des réseaux de neurones artificiels pour l'apprentissage à partir des données.

## 2.6 L'apprentissage automatique(Machine learning)

### 2.6.1 Définition

L'apprentissage automatique, ou machine learning, est une branche de l'intelligence artificielle qui permet aux ordinateurs d'apprendre à partir de données sans être explicitement programmés. Il leur permet de détecter des modèles dans les données et de prendre des décisions en fonction de ces modèles. Il existe deux types principaux : l'apprentissage supervisé, où le modèle est entraîné avec des données étiquetées, et l'apprentissage non supervisé, où le modèle découvre des structures et des relations dans les données non étiquetées.

### 2.6.2 Domaine d'application du machine Learning

Le domaine d'application du machine learning est très vaste et couvre de nombreux secteurs d'activité, tels que :

#### **La santé**

pour la détection précoce des maladies, la personnalisation des traitements, la prédiction de résultats cliniques, etc.

#### **Les services financiers**

pour la détection de fraudes, l'analyse de crédit, la gestion de portefeuille, la prévision de tendances, etc.

#### **Le marketing**

pour la segmentation de la clientèle, la prédiction des comportements d'achat, la recommandation de produits, etc.

#### **Les médias sociaux**

pour l'analyse des sentiments, la détection des tendances, la recommandation de contenu, etc.

## Les transports

pour l'optimisation des trajets, la prédiction des pannes, la gestion du trafic, etc.

Ces exemples ne sont pas exhaustifs et il existe de nombreux autres domaines où le machine learning peut être appliqué pour résoudre des problèmes complexes et aider à prendre des décisions plus éclairées.

### 2.6.3 Principe de fonctionnement de machine learning

Le principe de fonctionnement du machine learning est basé sur l'apprentissage à partir de données. Voici les étapes principales :

#### Collecte de données

On rassemble différentes données provenant de sources variées, comme des images, du texte, des vidéos, etc.

#### Préparation des données

Les données sont nettoyées et préparées pour l'analyse, en les rendant exploitables par le modèle.

#### Entraînement du modèle

On utilise un algorithme spécifique pour entraîner le modèle à partir des données préparées, afin qu'il apprenne à reconnaître des schémas ou des relations.

#### Validation du modèle

On évalue les performances du modèle en le testant sur un ensemble de données distinct.

#### Utilisation du modèle

Une fois le modèle entraîné et validé, on peut l'utiliser pour effectuer des prédictions ou prendre des décisions sur de nouvelles données.

Ce processus peut être répété de manière itérative pour améliorer la précision et la performance du modèle de machine learning.

#### **2.6.4 l'objectif de machine learning dans le domaine de la santé**

Le but du machine learning dans le domaine de la santé est d'utiliser des algorithmes pour extraire des connaissances à partir des données médicales. Ces connaissances peuvent être utilisées pour améliorer les soins de santé, prédire des résultats médicaux, identifier des risques pour la santé, développer des traitements personnalisés et améliorer la recherche médicale en général. Le machine learning permet de trouver des relations entre les symptômes, les traitements et les résultats de santé, ainsi que d'aider à détecter des anomalies et à prédire des résultats médicaux pour les patients.

### **2.7 l'apprentissage en profondeur(Deep learning)**

#### **2.7.1 Définition**

Le deep learning est une technique d'apprentissage automatique qui utilise des réseaux de neurones artificiels profonds pour effectuer des tâches complexes comme la reconnaissance vocale, la vision par ordinateur et la traduction automatique. Les réseaux de neurones profonds sont des modèles qui ont plusieurs couches de neurones, ce qui leur permet d'analyser des données non structurées comme des images, des vidéos et des textes. Le deep learning est utilisé dans de nombreux domaines tels que la reconnaissance d'image, la conduite autonome, la recommandation de produits et la détection de fraudes.

#### **2.7.2 Domaine d'application du deep Learning**

Le deep learning est utilisé dans une grande variété de domaines et a des applications potentielles dans presque tous les secteurs qui traitent de grandes quantités de données non structurées. Voici quelques exemples de domaines d'application du deep learning :

**Vision par ordinateur**

pour reconnaître des images, détecter des objets, et faire de la reconnaissance faciale.

**Traitement du langage naturel**

pour la reconnaissance vocale, la traduction automatique, et la classification de texte.

**Santé**

pour l'analyse d'images médicales, le diagnostic de maladies, et la prédiction de résultats cliniques.

**Finance**

pour détecter les fraudes, prévoir les prix des actions, et évaluer la solvabilité des entreprises.

**Jeux vidéo**

pour améliorer l'intelligence artificielle dans les jeux, offrant des expériences de jeu plus réalistes.

**2.7.3 Principe de fonctionnement de deep learning**

Le deep learning est une technique d'apprentissage automatique basée sur l'utilisation de réseaux de neurones artificiels profonds. Le principe de fonctionnement du deep learning peut être décrit en quatre étapes principales :

**Apprentissage**

La troisième étape consiste à entraîner le modèle de deep learning à partir des données préparées. Le modèle de deep learning utilise des algorithmes de rétropropagation pour ajuster les poids des neurones dans le réseau afin de minimiser l'erreur de prédiction.

## Collecte des données

Le premier pas consiste à collecter les données nécessaires à l'apprentissage. Ces données peuvent être des images, des vidéos, des textes ou d'autres types de données non structurées.

## Préparation des données

La deuxième étape consiste à préparer les données pour l'entrée dans le modèle de deep learning. Cela peut inclure le redimensionnement des images, le nettoyage des données, la normalisation des données, etc.

## Prédiction

Une fois le modèle entraîné, il peut être utilisé pour prédire les résultats pour de nouvelles données en entrée. Le modèle est alimenté avec les nouvelles données, et il effectue une prédiction en utilisant les relations apprises pendant la phase d'apprentissage.

### 2.7.4 L'objectif de deep learning dans le domaine de la santé

Le deep learning dans le domaine de la santé vise à améliorer les soins aux patients en permettant un diagnostic plus précis, une prédiction des résultats cliniques, la découverte de nouveaux médicaments et une analyse approfondie des données de santé. En utilisant des modèles de deep learning entraînés sur des données médicales, il est possible d'améliorer les traitements et d'identifier les patients à risque de complications. En résumé, le deep learning contribue à améliorer les soins et les traitements médicaux.

## 2.8 Comparaison entre machine learning et deep learning

	machine learning	deep learning
Approche	Utilise des algorithmes pour l'apprentissage automatique	Utilise des réseaux de neurones profonds
Complexité des tâches	Convient à différentes tâches	Convient aux tâches complexes
Représentation des données	Extraction manuelle des caractéristiques des données	Apprentissage autonome des représentations des données
Besoins en données et en calcul	Peut fonctionner avec moins de données et des ressources de calcul moins importantes	Nécessite plus de données et des ressources de calcul puissantes
Performances	Bonnes performances	Performances exceptionnelles dans des tâches complexes

TABLE 2.6 – Comparaison entre machine learning et deep learning

## 2.9 Conclusion

En conclusion de ce chapitre d'étude préalable, nous avons examiné les différents types d'architectures utilisées dans le domaine des Big Data, ainsi que les technologies associées, telles que Hadoop, Spark, et les bases de données NoSQL, qui offrent des capacités de stockage et de traitement massives pour les données volumineuses. Nous avons également découvert certaines sous-disciplines de l'Intelligence Artificielle (IA) comme l'apprentissage en profondeur et l'apprentissage automatique, qui jouent un rôle clé dans le traitement des données massives.

# Analyse et conception

## 3.1 Introduction

Au cours de cette section, nous avons employé le langage UML afin de concevoir notre application web. Nous avons élaboré divers schémas pour décrire les caractéristiques de l'application, les échanges entre les utilisateurs et le système, ainsi que la structure interne de celle-ci. Ces schémas nous permettent de prévoir et de visualiser le fonctionnement de notre application.

## 3.2 Les diagramme utilisés

### 3.2.1 Définition de l'UML

UML (Unified Modeling Language) est un langage de modélisation graphique utilisé pour spécifier, visualiser, concevoir et documenter des systèmes logiciels. UML est un langage de modélisation standardisé qui permet aux développeurs de logiciels de communiquer et de comprendre les différents aspects d'un système logiciel.

UML se compose d'un ensemble de diagrammes qui représentent différents aspects du système logiciel, tels que les exigences, la structure, le comportement, les interactions entre les composants [1].

### 3.2.2 Diagramme d'architecture de système

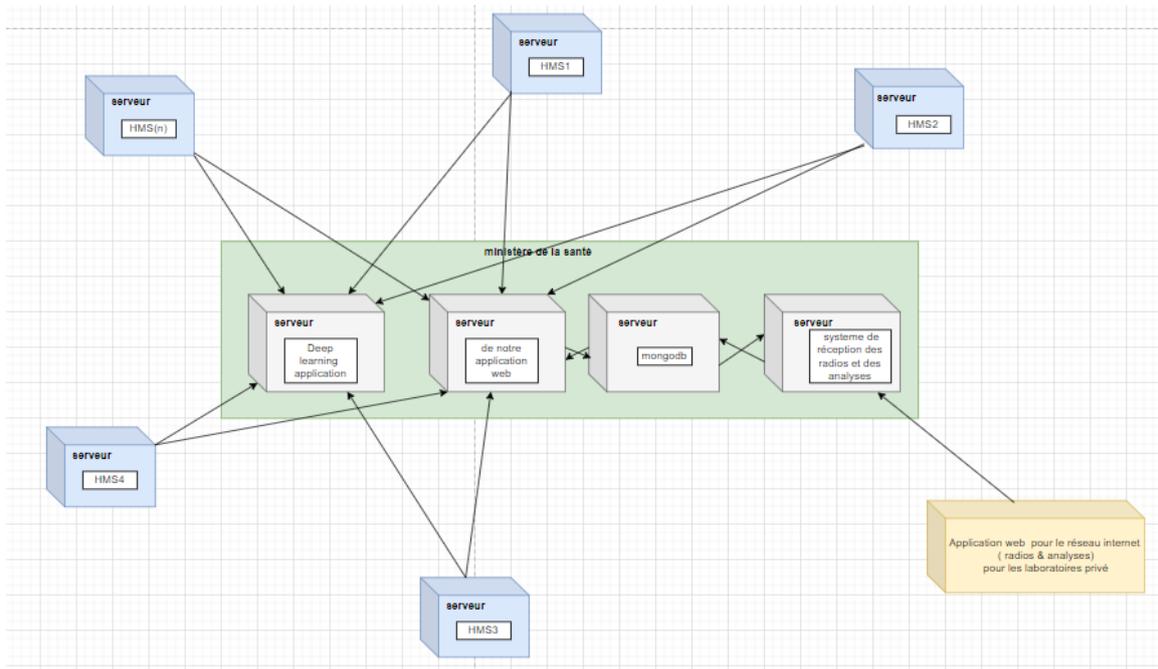


FIGURE 3.1 – Diagramme d'architecture de système

### 3.2.3 Diagramme des acteurs

#### Identification des acteurs

Lorsqu'il s'agit de concevoir des diagrammes de cas d'utilisation en UML, l'étape d'identification des acteurs revêt une grande importance dans le processus de modélisation des systèmes. Les acteurs représentent les entités externes qui interagissent avec le système et qui ont des rôles ou des responsabilités dans son utilisation ou sa gestion. 3.2

Nous décrivons dans ce Tableau les Sept acteurs principaux de notre système.

<b>Acteur</b>	<b>Description</b>
super admin	- Il/Elle gère les administrateurs hospitaliers du système et leurs profils. Il a la capacité d'accéder au système pour saisir, valider et modifier des informations spécifiques.
Admin	- Il/Elle gère les utilisateurs du système et administre le système.
Médecin	- Il/Elle s'occupe des dossiers médicaux des patients et gère les résultats d'examens médicaux de radiologie et d'analyses.
Réceptionniste	- Responsable de la gestion des patients et de leurs inscriptions. Il peut accéder au système pour saisir, valider et mettre à jour les informations des patients, ainsi que les orienter vers les médecins.
Infirmier	- Au sein d'une équipe de professionnels de la santé, cette personne assure la prestation de soins infirmiers aux patients.
Radiologue	- Le rôle du radiologue est de gérer, réaliser et interpréter des examens d'imagerie médicale (scanner, IRM, etc.) pour aider au diagnostic de maladies.
Technicien de laboratoire	- Il/Elle réalise des analyses et des tests sur des échantillons biologiques pour aider au diagnostic de maladies.

TABLE 3.1 – Description des acteurs

### 3.2.4 Diagramme de cas d'utilisation

Est un type de diagramme UML (Unified Modeling Language) qui décrit les interactions entre un système et ses utilisateurs ou acteurs. Il montre les fonctionnalités du système sous forme de cas d'utilisation et comment ces fonctionnalités sont utilisées par les acteurs du système.

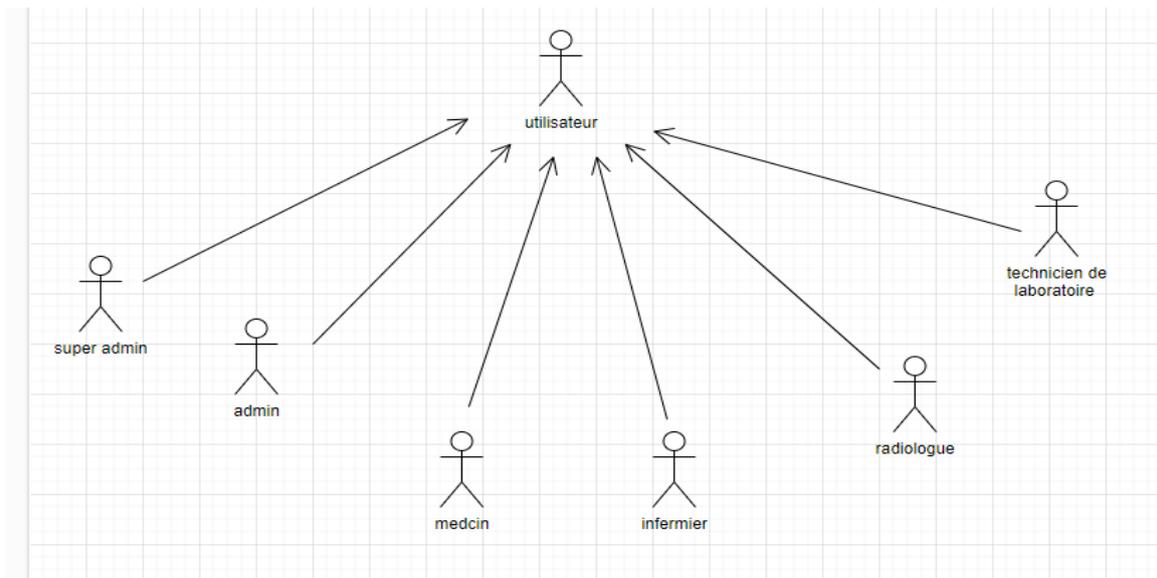


FIGURE 3.2 – Diagramme des acteurs

### A-Le rôle de super admin

Son rôle consiste à gérer les utilisateurs du système (à savoir les administrateurs hospitaliers) et leurs profils. Il a la capacité d’ajouter des hôpitaux, d’accéder au système pour saisir, valider et modifier des informations spécifiques.

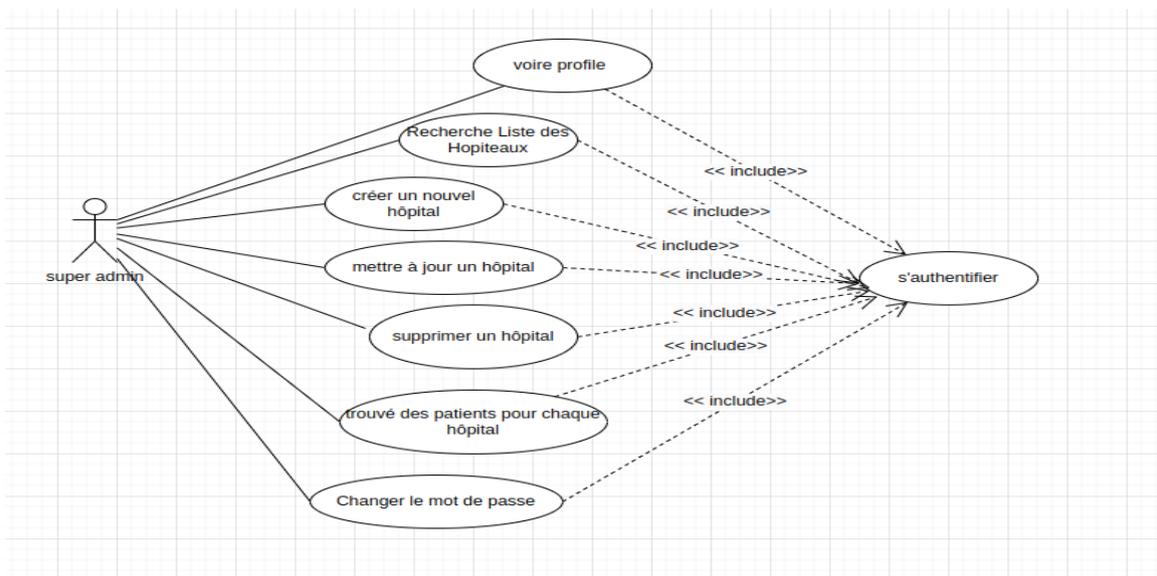


FIGURE 3.3 – Rôle de super admin

### B-Le rôle de L'admin

Gère les utilisateurs du système (médecins,radiologues,infirmiers et techniciens de laboratoire) et leurs profils ,Comme il peut acceder au systeme pour remplir et acher et modifier certaines informations.

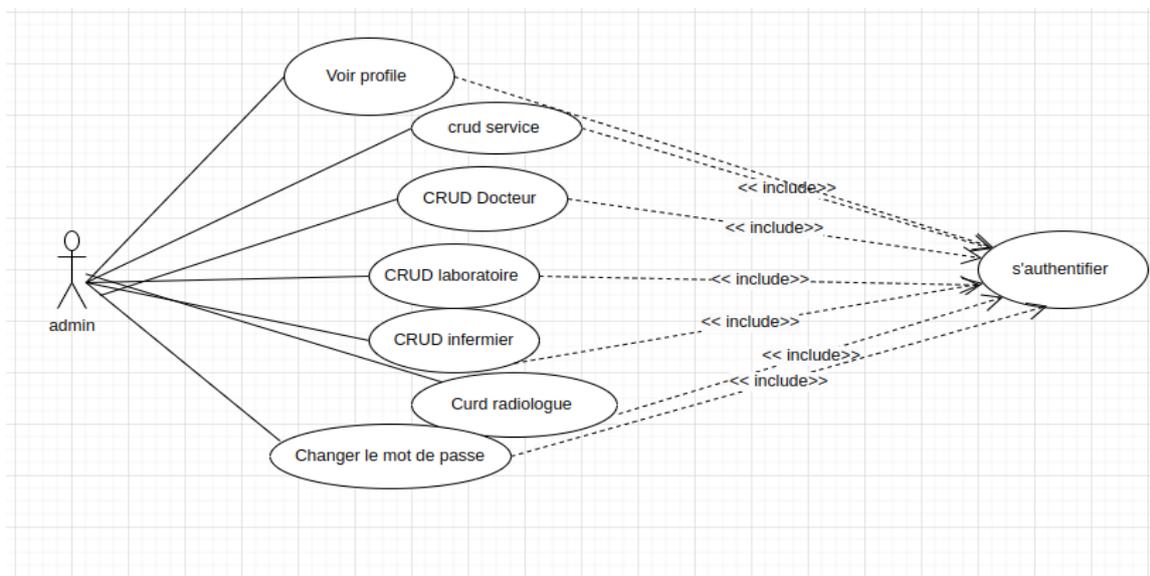


FIGURE 3.4 – Rôle de L'admin

### C-Le rôle de medcin

Son rôle est de diagnostiquer les patients ,prescrire les traitements appropriés ( médicaments,radio,analyses ou des soins) .

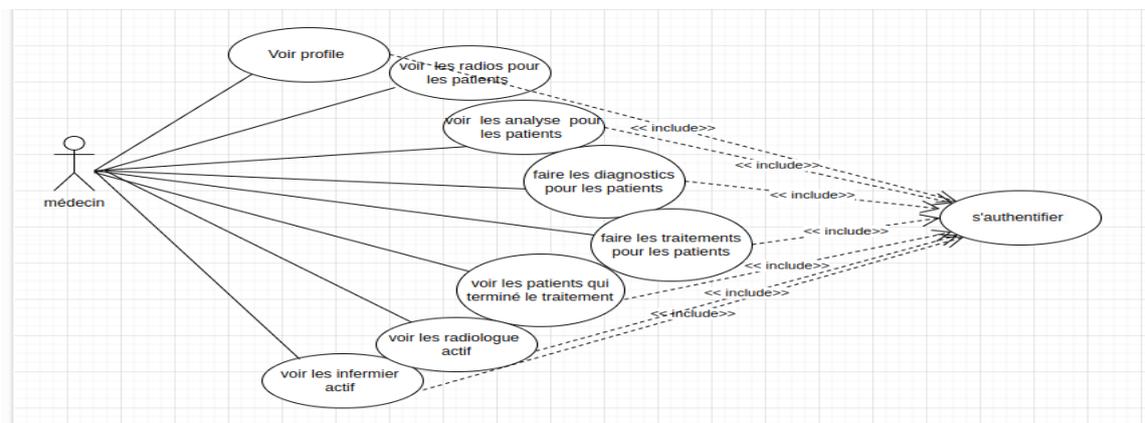


FIGURE 3.5 – Rôle de medcin

### D-Le rôle de réceptionniste

Il assume de charger la gestion des patients et leurs inscriptions. Comme il peut accéder au système pour saisir, valider et mettre à jour les informations des patients ainsi que leurs orientations chez les médecins.

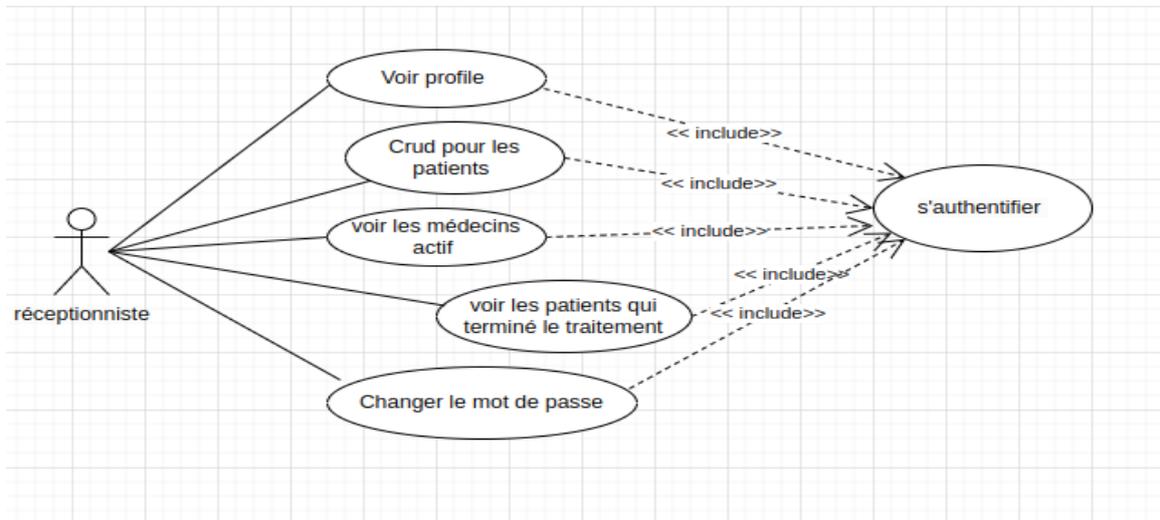


FIGURE 3.6 – Rôle de réceptionniste

### E-Le rôle de technicien de laboratoire

Il réalise des analyses et des tests sur des échantillons biologiques pour aider à diagnostiquer des maladies.

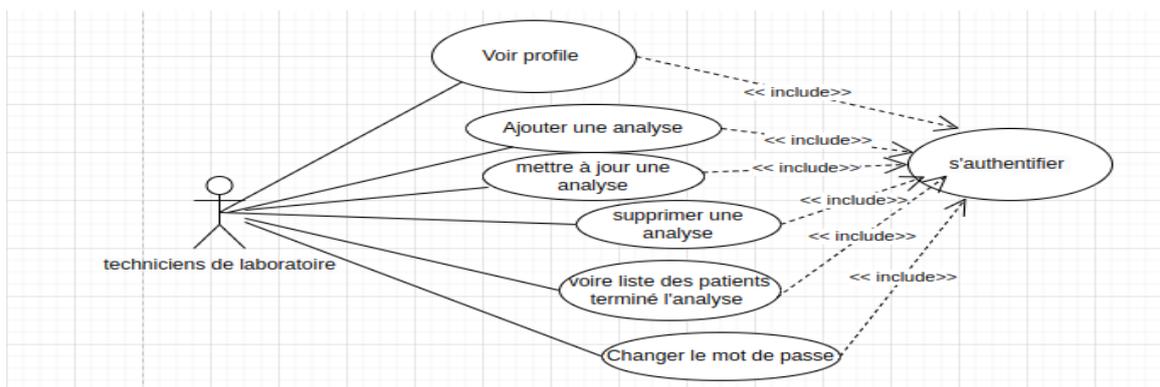


FIGURE 3.7 – Rôle technicien de laboratoire

### F-Le rôle de radiologue

Le rôle du radiologue est de gérer et réaliser et d'interpréter des examens d'imagerie médicale(scanner,IRM,...) pour aider à diagnostiquer des maladies .

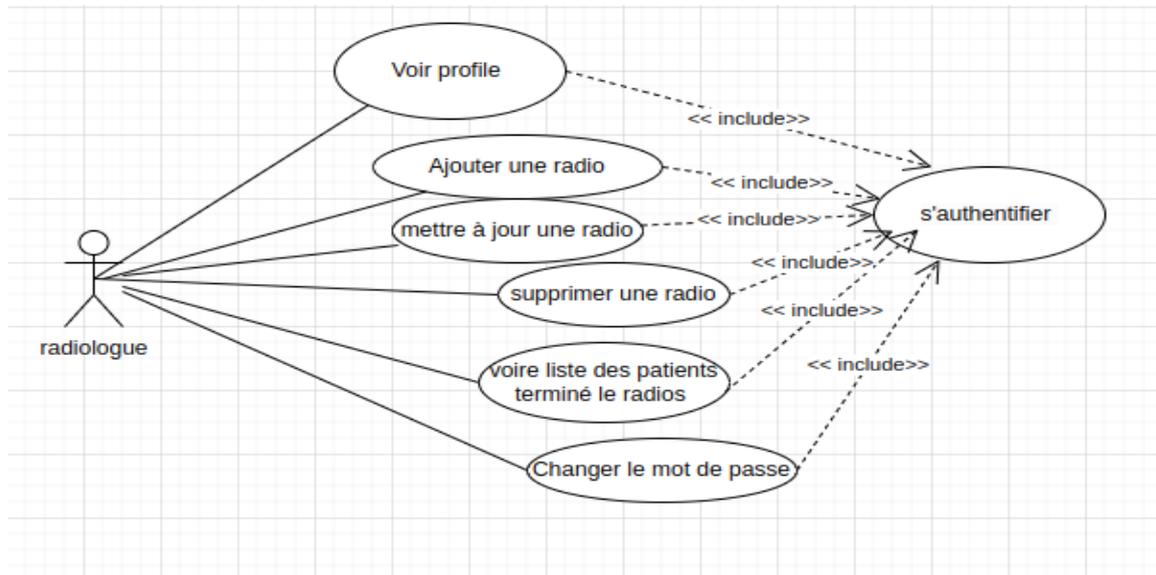


FIGURE 3.8 – Rôle de Radiologue

### G-Le rôle de l'infirmier

Le rôle de l'infirmier est d'administrer des soins infirmiers aux patients en collaboration avec une équipe de professionnels de la santé.

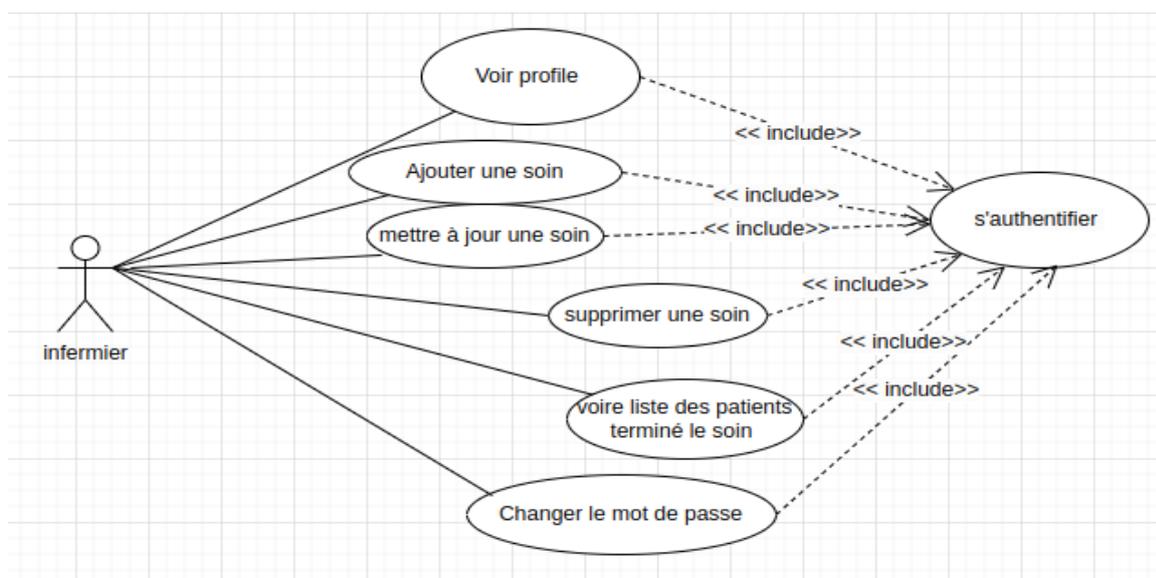


FIGURE 3.9 – Rôle de L'infirmier

### 3.2.5 Diagramme de séquence

Un diagramme de séquence, relevant de l'Unified Modeling Language (UML), permet de visualiser la séquence des actions qui se déroulent dans un système ou une application donnés. Son but est de décrire de manière graphique l'interaction entre divers objets ou acteurs au sein du système, en présentant de façon chronologique l'ordre dans lequel les messages sont échangés entre eux.

#### A-Diagramme de séquence pour le cas «L'Authentification »

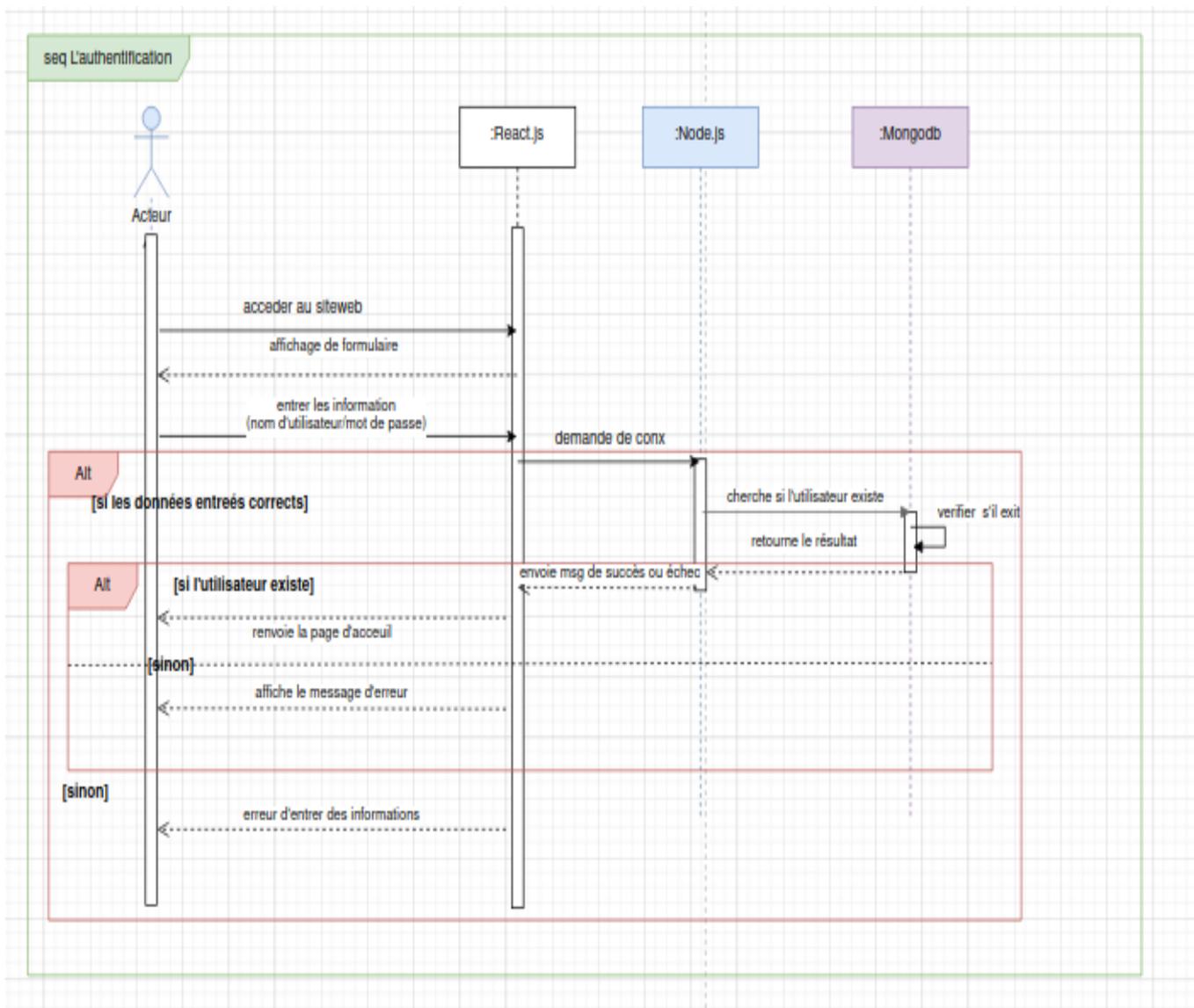


FIGURE 3.10 – L'Authentification

B-Diagramme de séquence pour le cas «Ajouter & Modifier l'utilisateur »

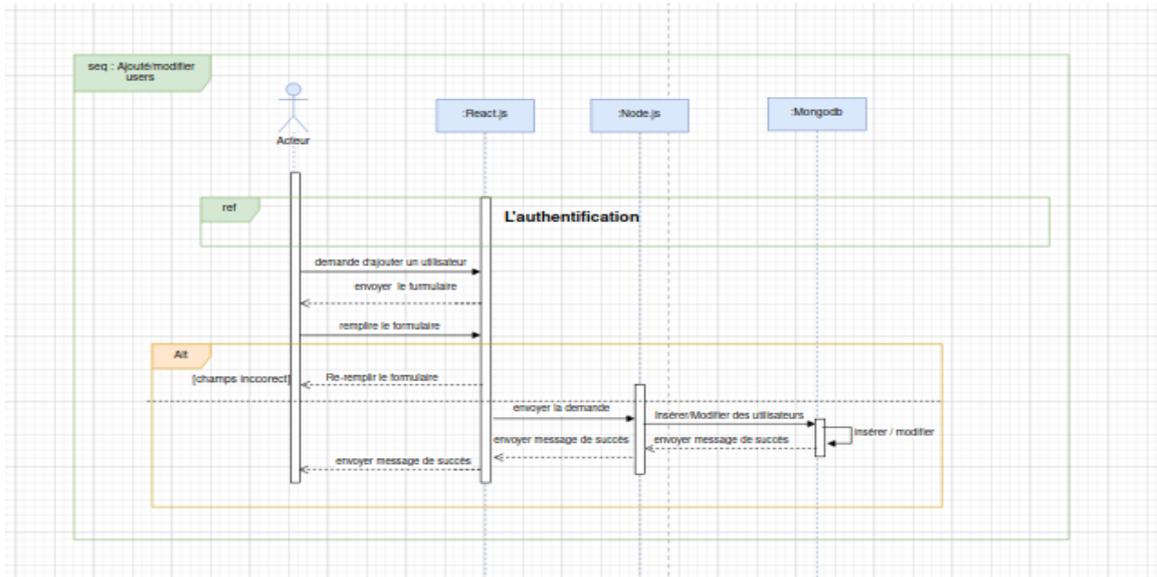


FIGURE 3.11 – Ajouter Modifier l'utilisateur

C- «Supprimer l'utilisateur»

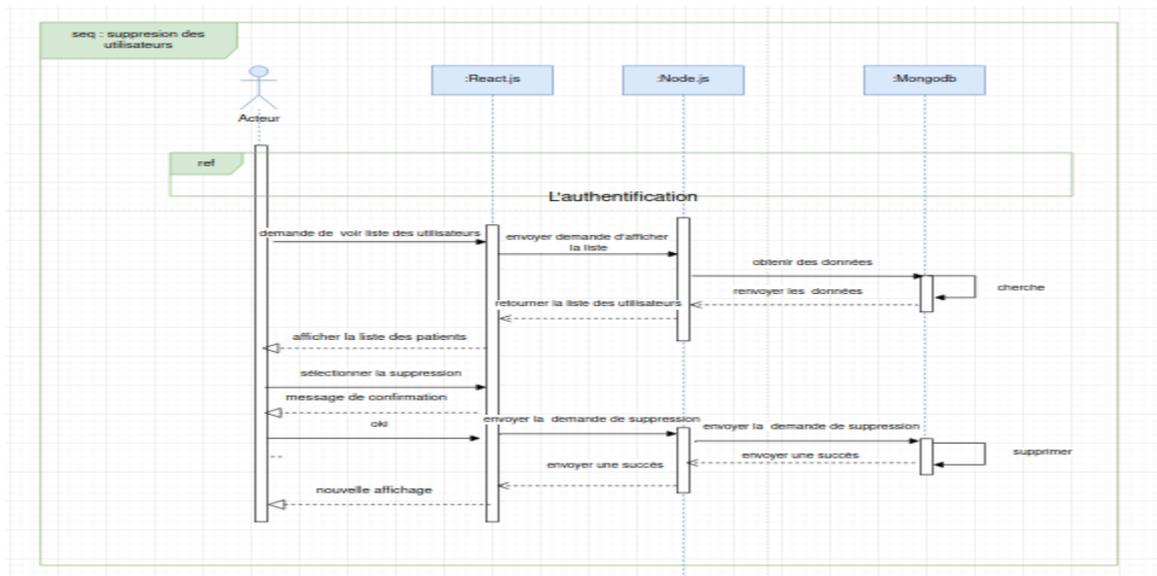


FIGURE 3.12 – Supprimer l'utilisateur

D-Diagramme de séquence pour le cas «Ajouter & Modifier symptômes,diagnostique et traitement »

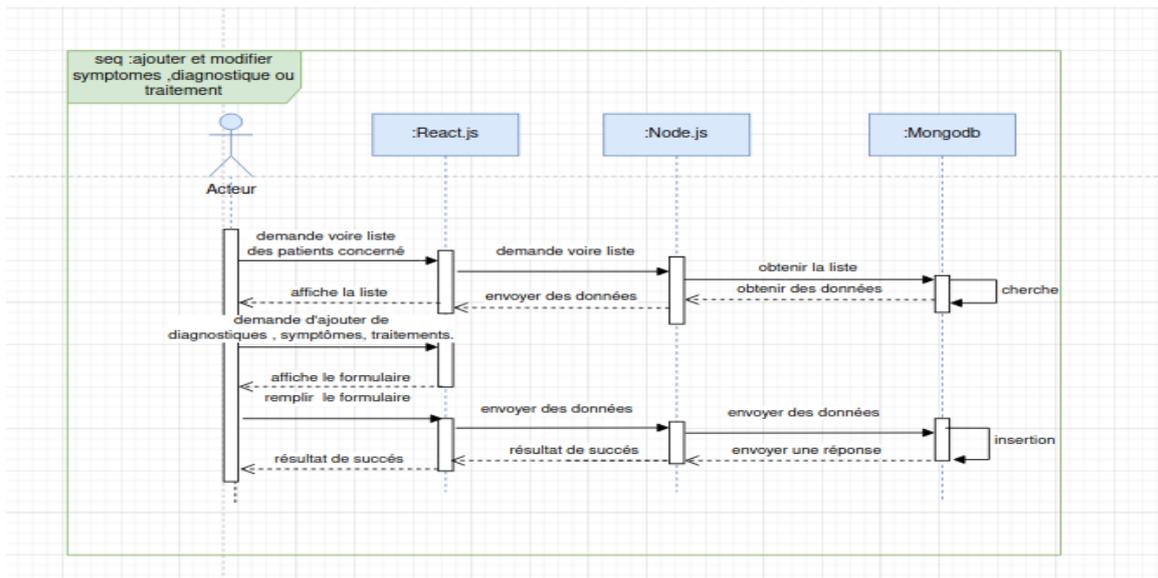


FIGURE 3.13 – Ajouter et modifier symptômes,diagnostique et traitement

E- «Supprimer symptômes,diagnostique et traitement »

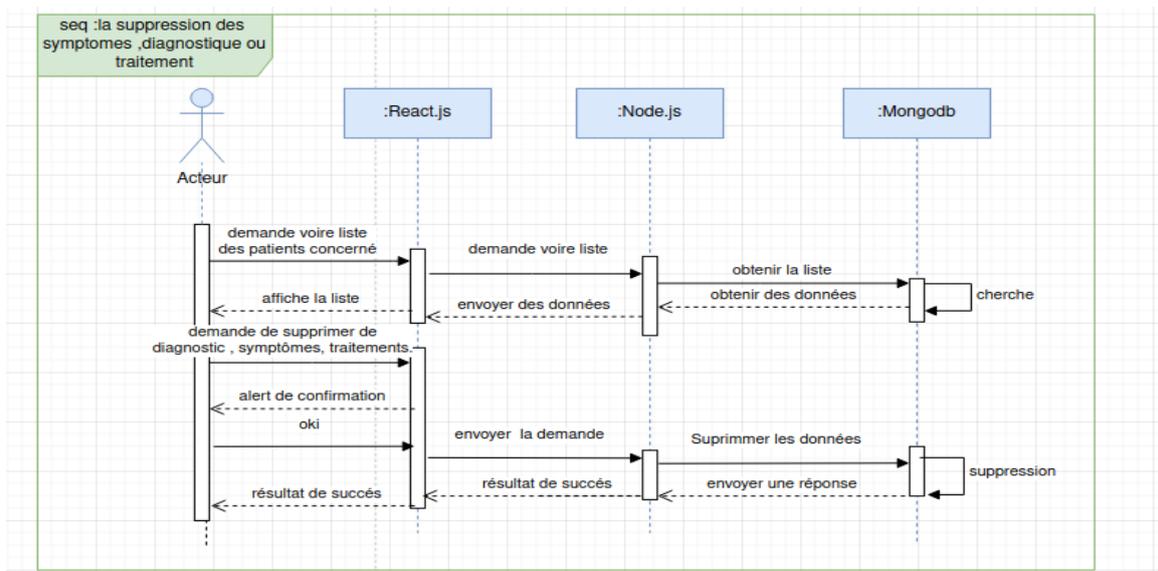


FIGURE 3.14 – Supprimer symptômes,diagnostique et traitement

F-Diagramme de séquence pour le cas «Ajouter & Modifier analyse »

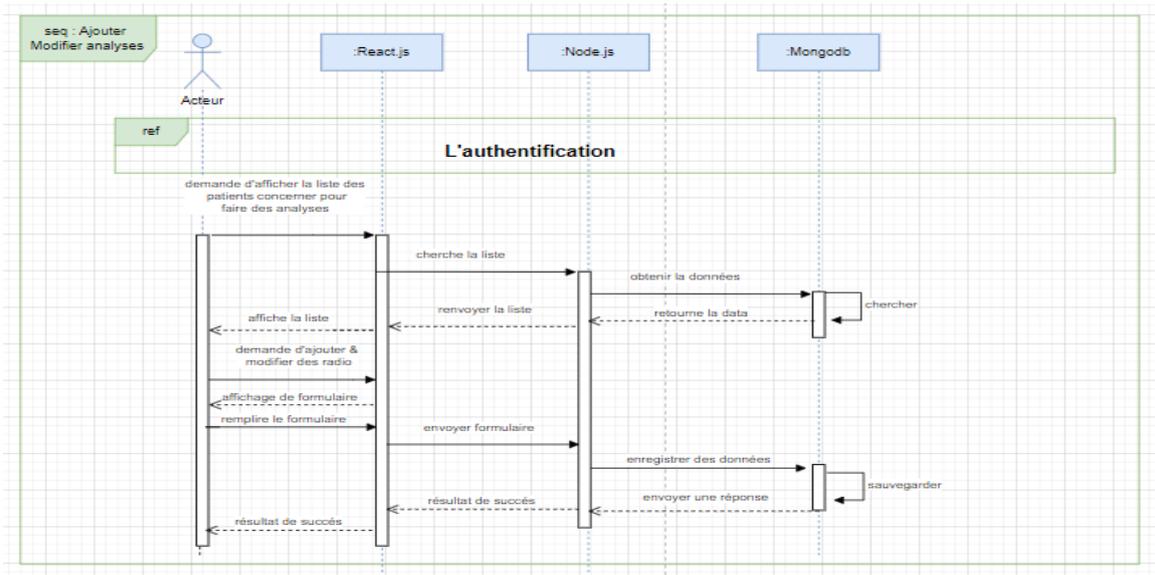


FIGURE 3.15 – Ajouter et Modifier analyse

G-«Supprimer analyse »

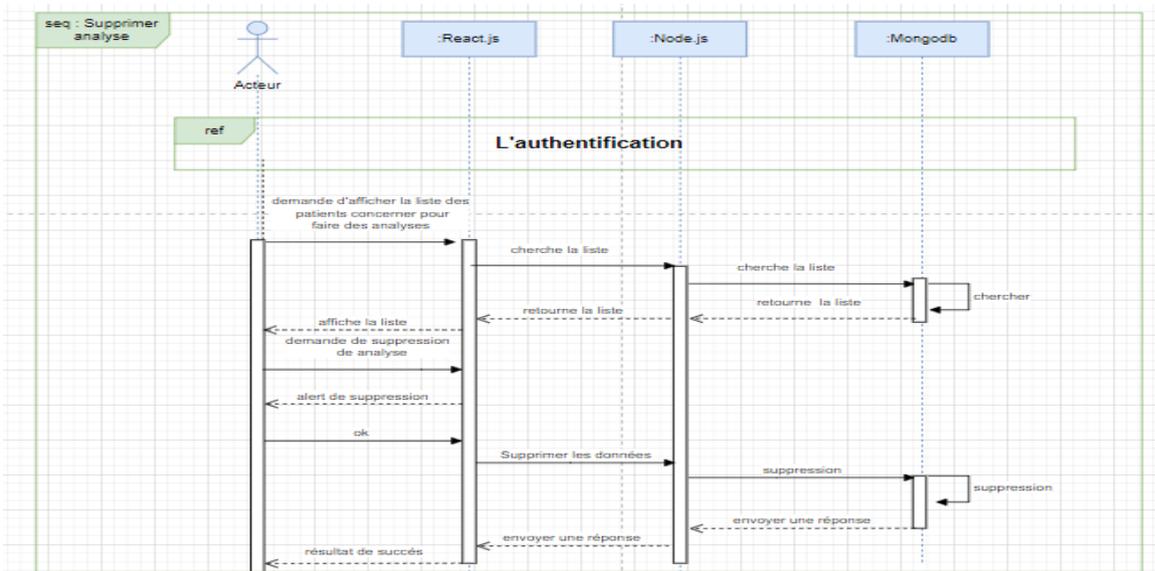


FIGURE 3.16 – Supprimer analyse

H-Diagramme de séquence pour le cas «Ajouter & Modifier la radiographie »

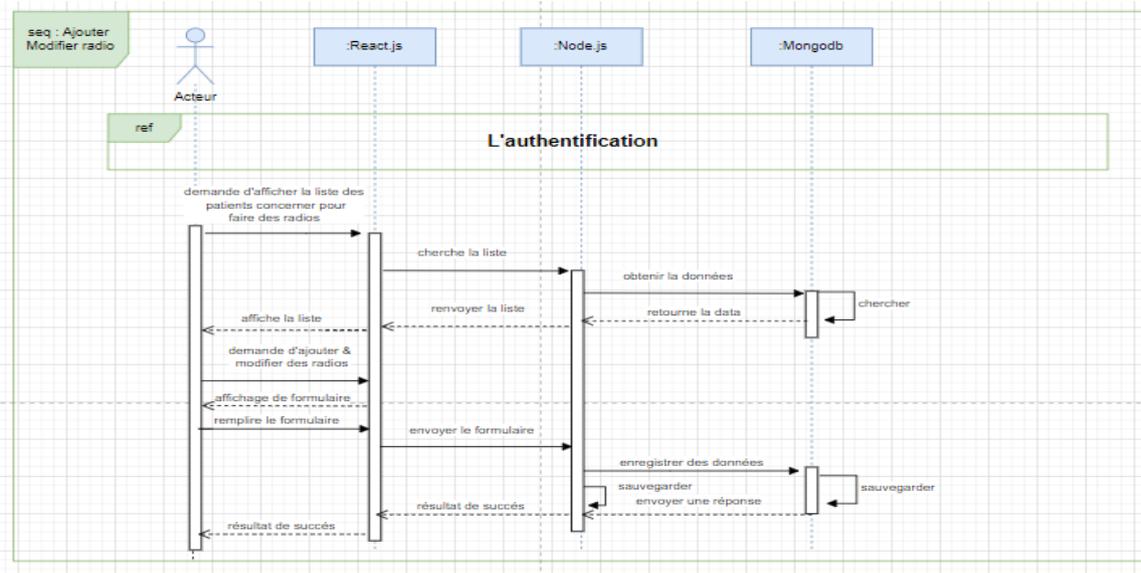


FIGURE 3.17 – Ajouter Modifier radio

I-«Supprimer la radiographie »

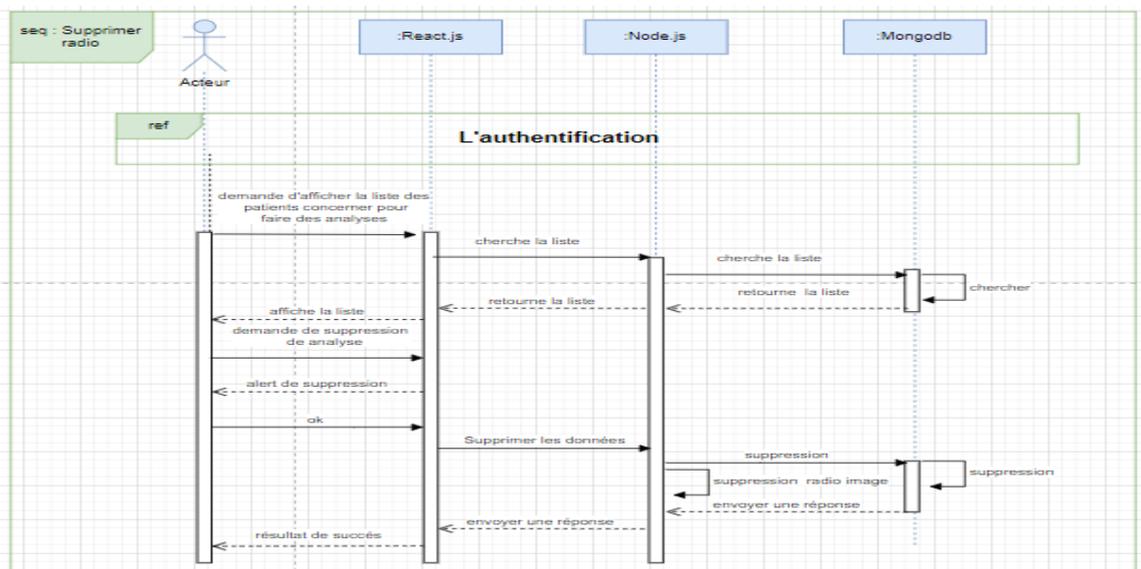


FIGURE 3.18 – Supprimer radio

J-Diagramme de séquence pour le cas «Ajouter et modifier soins »

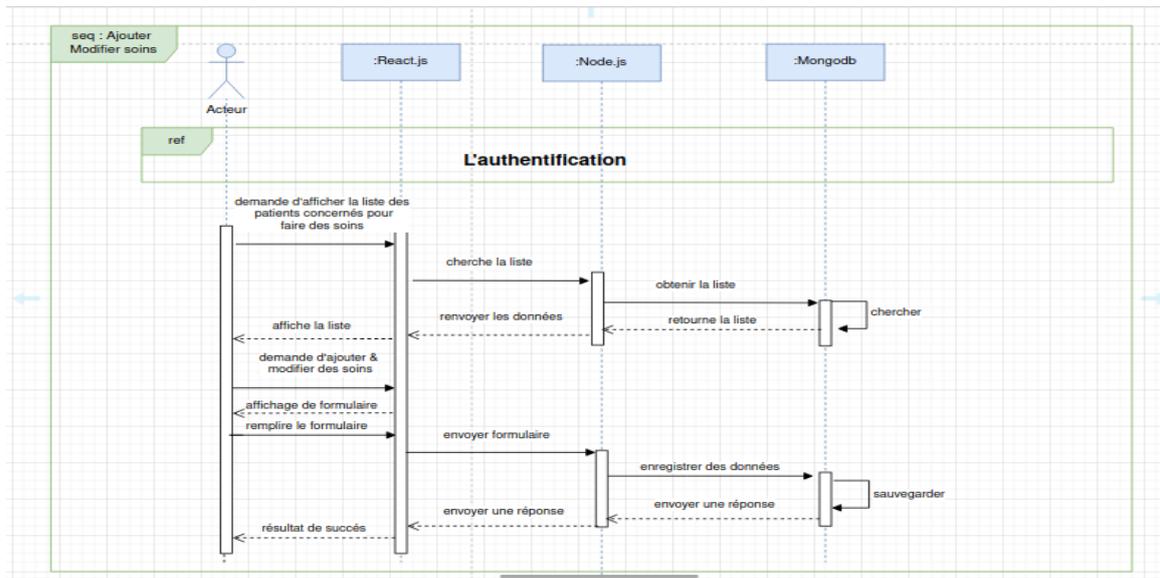


FIGURE 3.19 – Ajouter & Modifier soins

K-«Supprimer soins »

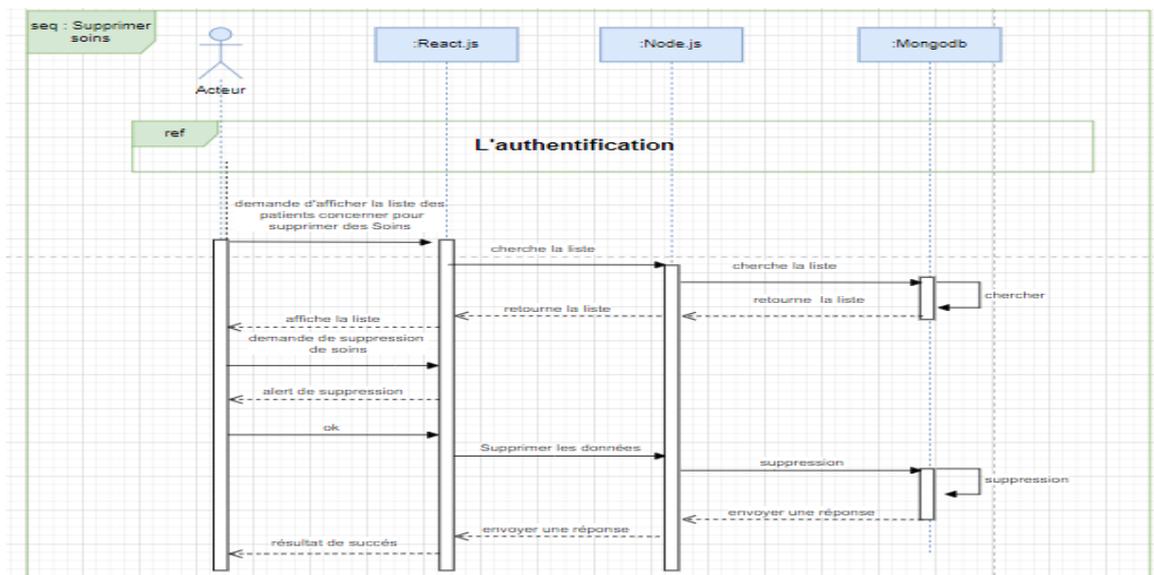


FIGURE 3.20 – Supprimer soins

L-Diagramme de séquence pour le cas «Ajouter et modifier un patient »

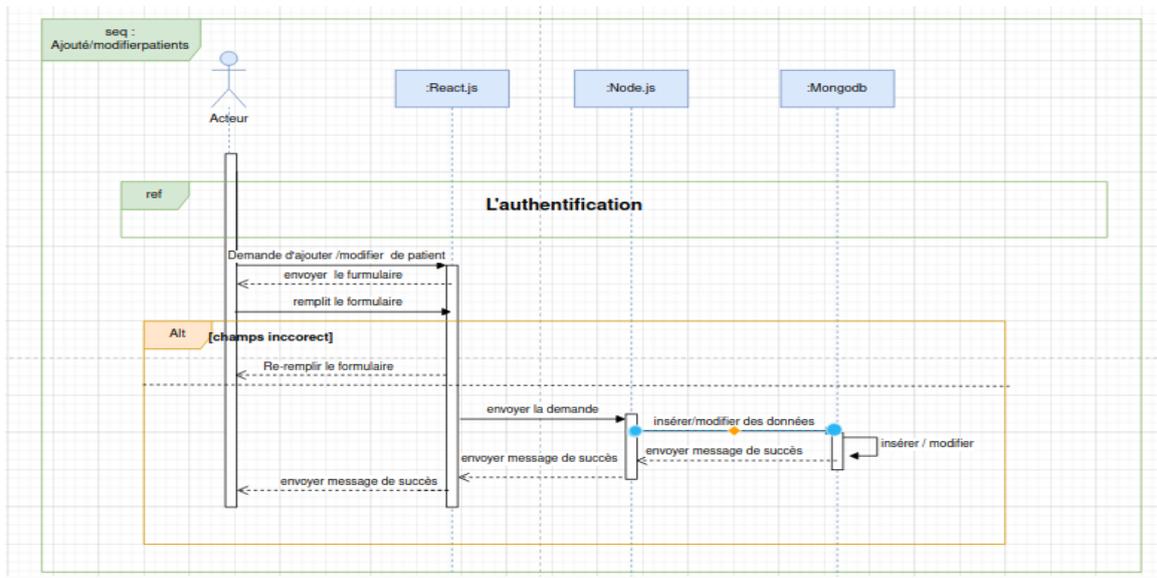


FIGURE 3.21 – Ajouter et modifier un patient

M-«Supprimer patient »

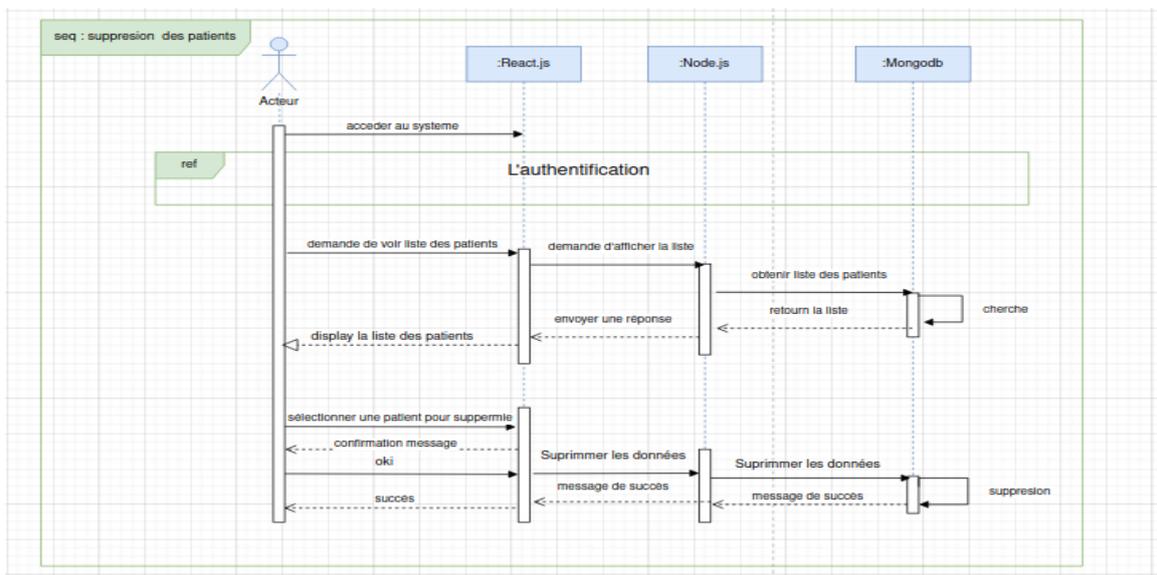


FIGURE 3.22 – Supprimer patient

N-Diagramme de séquence pour le cas «La recherche»

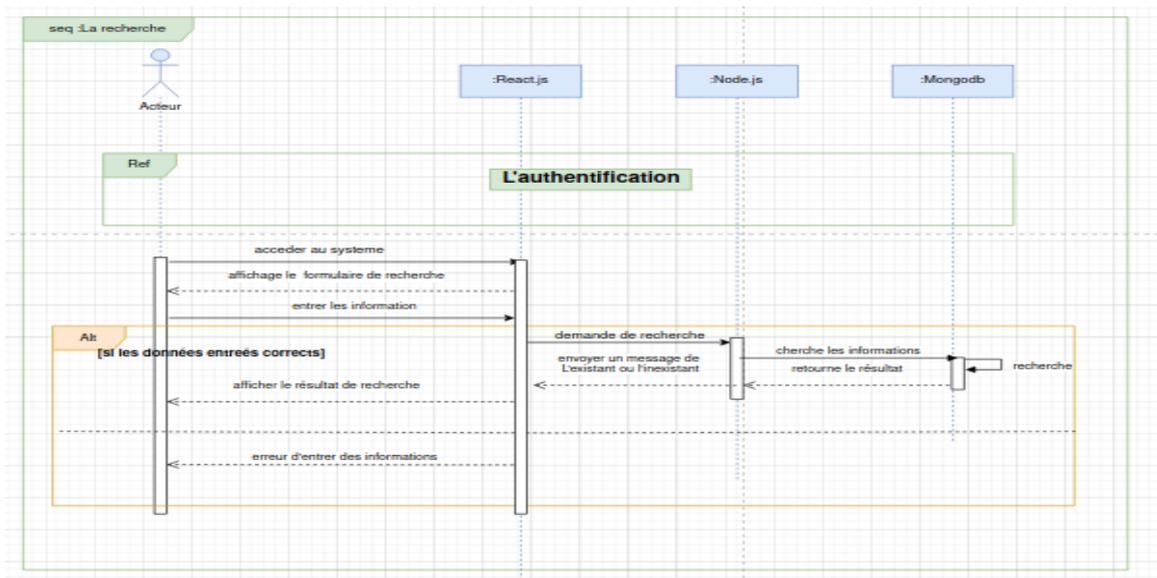


FIGURE 3.23 – La recherche

O-Diagramme de séquence pour le cas «Changer le mot de passe»

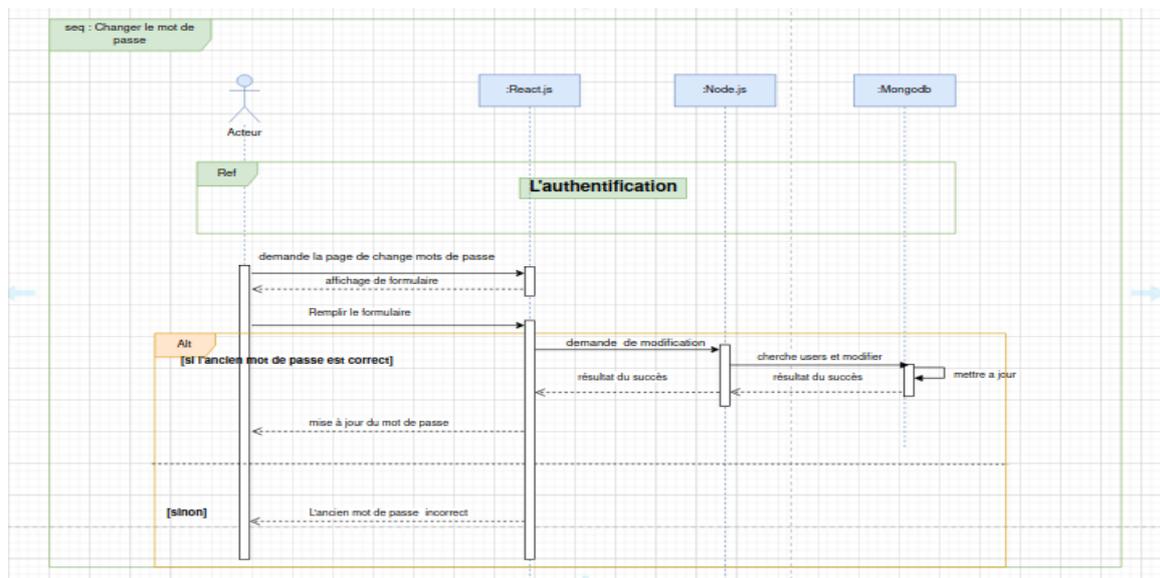


FIGURE 3.24 – Changer le mot de passe

## P-Diagramme de séquence pour le cas «insertion de radiographies et analyse par l'état privé»

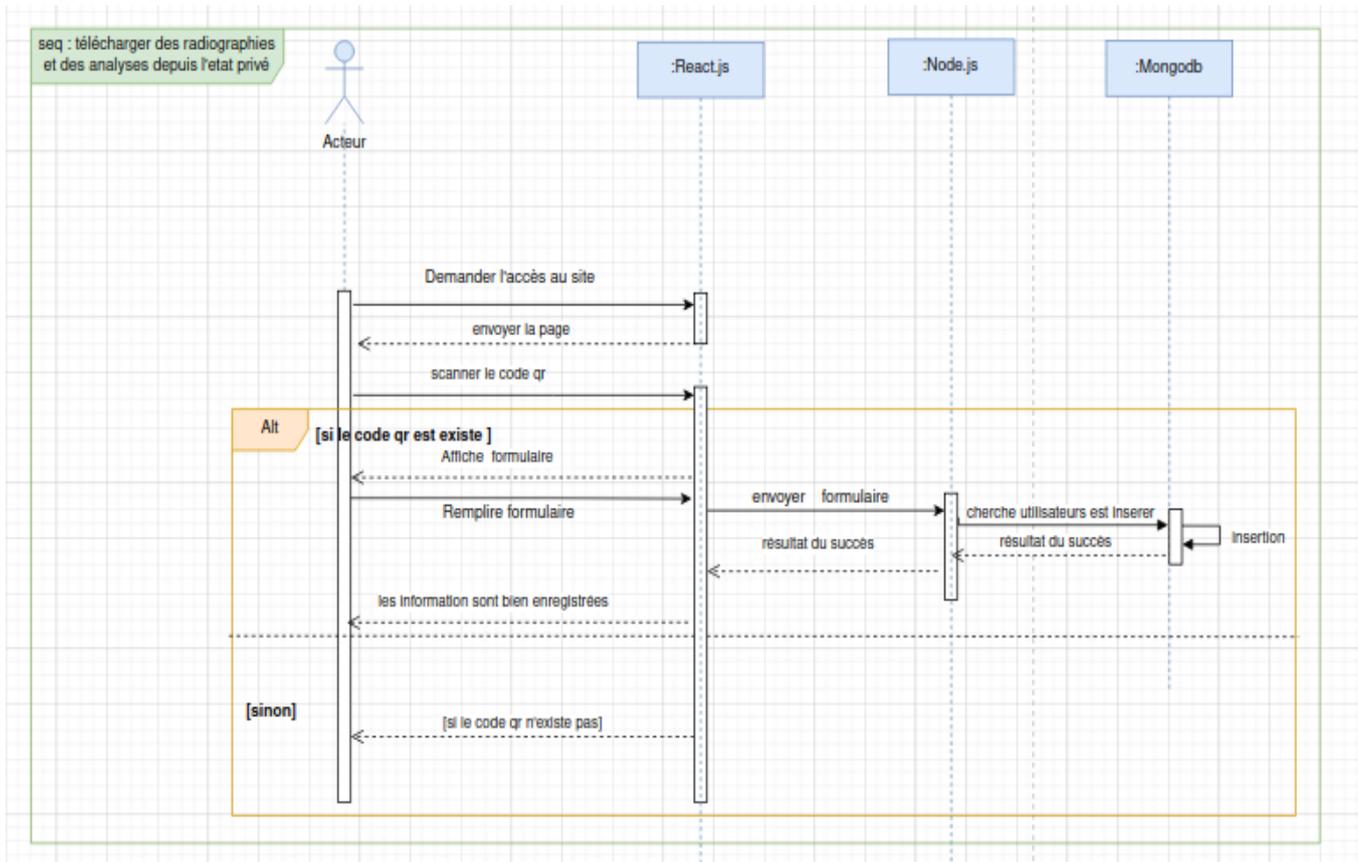


FIGURE 3.25 – insertion de radiographies et analyse par l'état privé

### 3.3 Conclusion

Au cours de ce chapitre, nous avons procédé à la modélisation de notre application en utilisant une variété de diagrammes et d'outils de modélisation. Cette approche nous a permis d'établir les schémas de la base de données, d'obtenir une vue d'ensemble du projet et de concevoir des interfaces répondant aux besoins identifiés. Le chapitre suivant se concentrera sur une exploration plus approfondie de ces aspects.

# Réalisation

## 4.1 Introduction

Au cours de cette dernière section, nous nous pencherons sur l'exploration des diverses ressources mobilisées pour concevoir et mettre en œuvre l'application. Nous allons présenter les langages de programmation, frameworks, outils de développement et la base de données créée. Nous examinerons également l'arborescence de l'application pour comprendre comment les différentes parties sont organisées et connectées. Nous présenterons quelques extraits de code clés pour mieux comprendre le fonctionnement de l'application. nous présenterons ensuite les différentes interfaces de notre système.

## 4.2 Environnement de développement

### 4.2.1 Environnement matériel

Caractéristiques	PC1 : Lenovo	PC2 : Hp
Disque dur	240 GigaO SSD , 500 GigaO HDD	500 GigaO HDD
Type de processeur	Intel® Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz	Intel® Prentium(R) CPU N3710 @ 1.60GHz 1.60GHz
Fréquence de processeur	2.71GHz	1.71 GHz
Mémoire mortCore i5 Core i3 S.E	Core i5 linux mint	Prentium(R) Windows 10 .

TABLE 4.1 – Environnement matériel

## 4.2.2 Environnement logiciel

Au cours du développement de cette application, nous avons exploité les logiciels et outils suivants pour faciliter le processus.

### Langage utilisé

#### 1. Mernstack

Le MERN Stack est un acronyme qui fait référence à une pile technologique pour le développement d'applications web. Chaque lettre représente une technologie spécifique : M :MongoDB est la première base de NoSQL, orientée document, sans schéma, qui stocke les données de l'application.

E : Express.js est un framework web léger spécialement conçu pour Node.js. Sa principale fonction est de simplifier la création de serveurs web et de routes afin de gérer efficacement les requêtes et les réponses.

R : React est une bibliothèque JavaScript utilisée pour construire l'interface utilisateur de l'application. Elle permet de créer des composants réutilisables et dynamiques pour une expérience utilisateur interactive.

N : Node.js est un environnement d'exécution côté serveur qui offre la possibilité de faire fonctionner du code JavaScript en dehors du navigateur. Son utilisation principale réside dans la création du serveur web dans la pile technologique MERN.

## 2. Python

Python est un langage de programmation généraliste et très riche avec de nombreuses fonctionnalités. C'est l'un des principaux objectifs d'un style de programmation connu sous le nom de programmation orientée objet. Il s'agit d'un langage de programmation interprété de haut niveau.

### Logiciel de développement

#### 1. VSCode

VSCode, également connu sous le nom de Visual Studio Code, est un éditeur de code léger et largement utilisé, développé par Microsoft. Il est spécialement conçu pour l'écriture et la modification de code dans divers langages de programmation. VSCode se distingue par son interface conviviale, ses capacités intégrées de débogage, ses suggestions intelligentes, sa gestion de projet efficace et sa possibilité d'extension pour étendre ses fonctionnalités.

#### 2. Overleaf

C'est un IDE LATEX en ligne, les utilisateurs ont la possibilité de rédiger et de partager des documents scientifiques en toute simplicité.

#### 3. Mongo Compass

MongoDB Compass est une interface graphique conviviale pour interagir avec les bases de données MongoDB. Cet outil facilite la visualisation des données, la création de requêtes et la gestion des bases de données grâce à une approche visuelle.

## Frameworks utilisés

### 1. Bootstrap

Bootstrap est un framework web largement utilisé pour créer des sites web réactifs et attrayants. Il fournit des composants prédéfinis, des styles CSS et des scripts JavaScript qui permettent aux développeurs de construire rapidement et facilement des interfaces utilisateur modernes. Avec Bootstrap, vous pouvez créer des mises en page adaptables à différents appareils et tailles d'écran sans avoir à écrire beaucoup de code CSS personnalisé. Cela rend le processus de développement plus efficace et vous permet d'obtenir un résultat esthétique et fonctionnel en peu de temps.

## 4.2.3 Les Schema utilisés

```

const mongoose = require("mongoose");

const userSchema = new mongoose.Schema({
  nonDeHopital: {
    type: String,
    required: true,
  },
  UserAd: {
    type: String,
    required: true,
  },
  Nom: {
    type: String,
    required: true,
  },
  Prenom: {
    type: String,
    required: true,
  },
  password: {
    type: String,
    required: true,
  },
  Date_de_naissance: {
    type: Date,
  },
  Numero_telephone: {
    type: String,
    default: "pas de numero ",
  },
  Email: {
    type: String,
  },
  Actif: {
    type: Boolean,
    required: true,
    default: false,
  },
  service: {
    type: String,
    default: true,
  },
  isSuperAdmin: {
    type: Boolean,
    default: false,
  },
  isAdmin: {
    type: Boolean,
    default: false,
  },
  isDoctor: {
    type: Boolean,
    default: false,
  },
  isNurse: {
    type: Boolean,
    default: false,
  },
  isRadiologue: {
    type: Boolean,
    default: false,
  },
  isResptionests: {
    type: Boolean,
    default: false,
  },
  isPharmacis: {
    type: Boolean,
    default: false,
  },
  isLaboratory: {
    type: Boolean,
    default: false,
  },
});
const userModel = mongoose.model("users", userSchema);

```

[utilisateur]

```

const userSchema = new Schema({
  //isResptionests
  ResptionestsId: {
    type: String,
    required: true,
  },
  Nss: {
    type: String,
    required: true,
  },
  nom: {
    type: String,
    required: true,
  },
  Prenom: {
    type: String,
    required: true,
  },
  Age: {
    type: String,
    required: true,
  },
  Phone: {
    type: String,
    required: true,
  },
  date: {
    type: Date,
    default: Date.now
  },
  traitementtermine: {
    type: Boolean,
    default: false,
    required: true,
  },
  radio: {
    type: Boolean,
    default: false,
    required: true,
  },
  radiofait: {
    type: Boolean,
    default: false,
    required: true,
  },
  analyse: {
    type: Boolean,
    default: false,
    required: true,
  },
  analysefait: {
    type: Boolean,
    default: false,
    required: true,
  },
  diagnostique: {
    type: Boolean,
    default: false,
    required: true,
  },
  Soins: {
    type: Boolean,
    default: false,
    required: true,
  },
  Soinsfait: {
    type: Boolean,
    default: false,
    required: true,
  },
  Attendre: {
    type: Boolean,
    default: false,
    required: true,
  },
  finish: {
    type: Boolean,
    default: false,
    required: true,
  },
  email: {
    type: String,
    validate: {

```

[patient]

FIGURE 4.1 – schéma

## 4.3 Implémentation

### 4.3.1 Démonstration des interfaces

interface de connexion

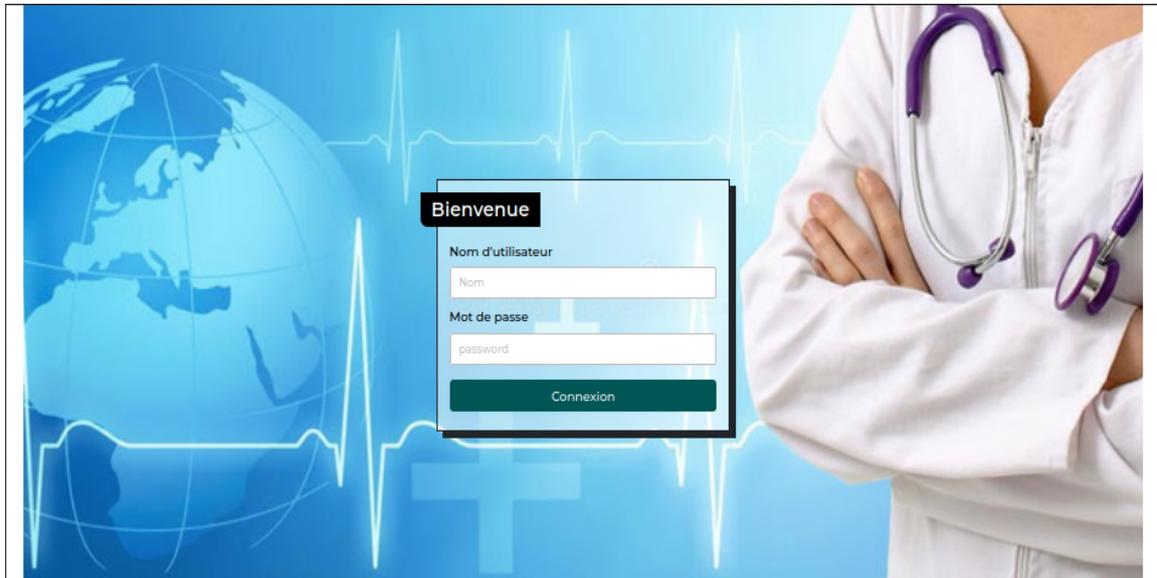


FIGURE 4.2 – Page connexion

Espace SuperAdmin



FIGURE 4.3 – Espace SuperAdmin

## Ajouter un nouvel hôpital

The screenshot shows a web form titled 'Ajouter un nouvel hôpital' within the HMS interface. The form contains the following fields:

- Le Nom de L'hopital
- ID de l'utilisateur
- Nom
- Prenom
- Mot de passe
- Date De Naissance (format: mm/dd/yyyy)
- Email
- Numero Telephone

At the bottom of the form is a blue button labeled 'Ajouté hopital'.

FIGURE 4.4 – Ajouter un nouvel hôpital

## Liste des admins des hôpitaux

The screenshot shows the 'Liste des admins' page in the HMS interface. It displays a table with the following data:

le nom de l'hopital	ID de l'utilisateur	Nom	Prenom	Date de naissance	Numero Telephone	Email	Actions
HMS1	Admin	Admin	Admin	2023-06-16T00:00:00.000Z	pas de numero	true	 
HMS2	adminHMS2	adminHMS2	adminHMS2	2023-06-16T00:00:00.000Z	0553196777	true	 

FIGURE 4.5 – Liste des admins des hôpitaux

Les étapes de visualisation et de téléchargement des analyse par le Ministère afin de les soumettre au deep learning



FIGURE 4.6 – Les étapes visualisation et de téléchargement des analyse

Les étapes de visualisation et de téléchargement des radiographies par le Ministère afin de les soumettre au deep learning



FIGURE 4.7 – Les étapes visualisation et de téléchargement des radiographies

## Espace Admin



FIGURE 4.8 – Espace admin

## Ajouter un médecin par l'admin



FIGURE 4.9 – Ajouter un médecin par l'admin

## Ajouter services par l'admin

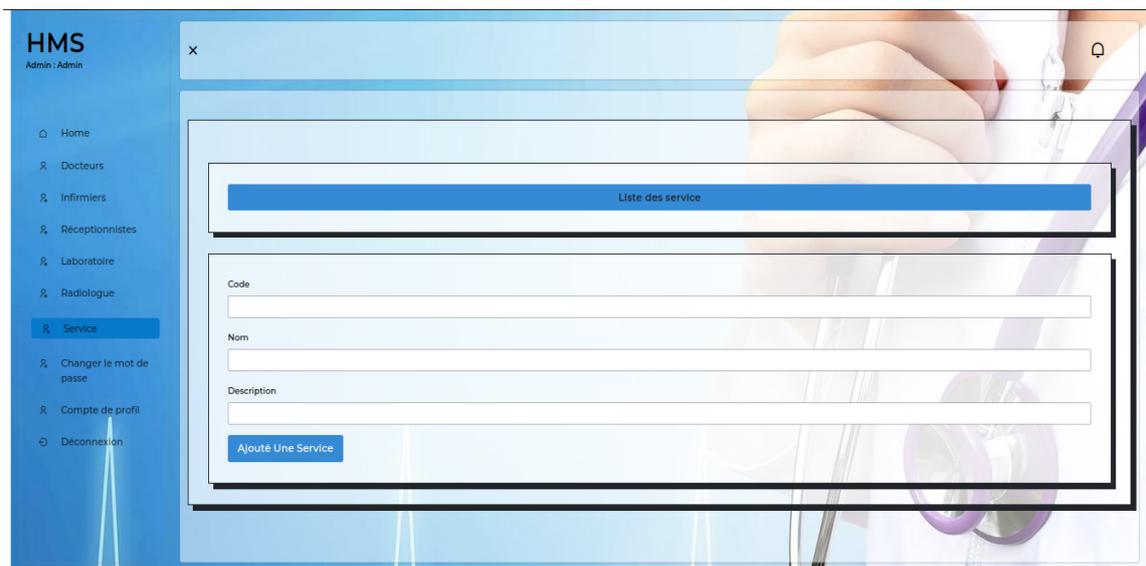


FIGURE 4.10 – Ajouter un service par l'admin

## Liste des services

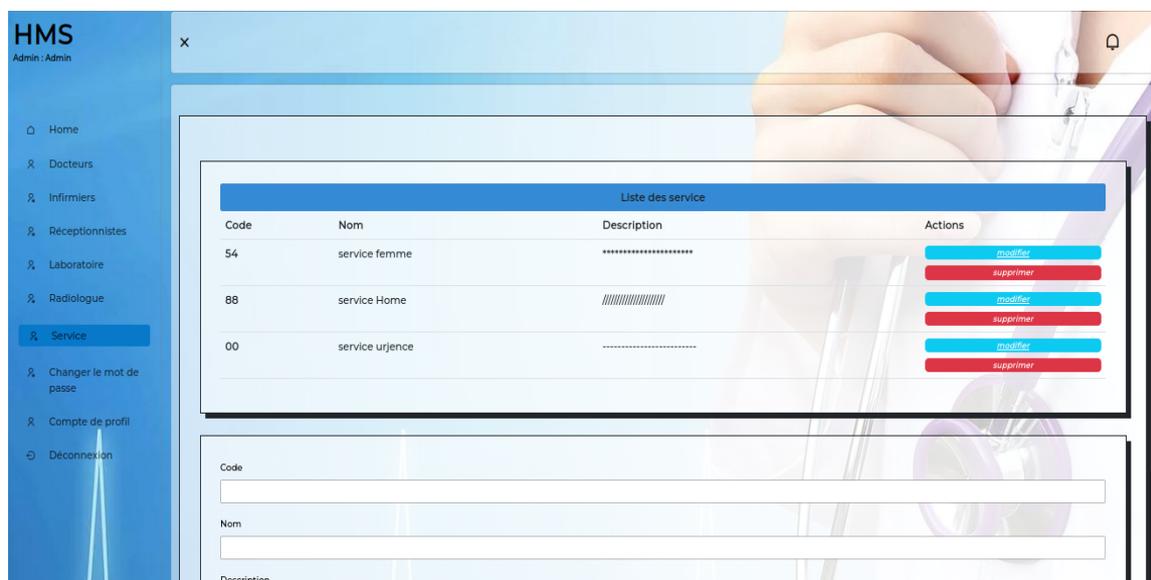


FIGURE 4.11 – Liste des service

## Espace médecin



FIGURE 4.12 – Espace médecin

## Liste des patients en visite

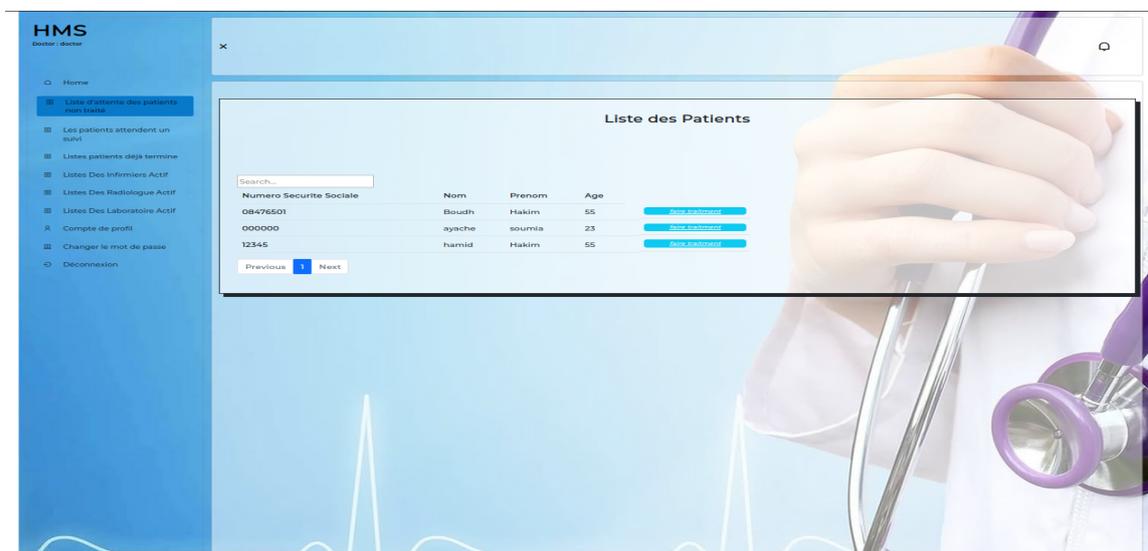


FIGURE 4.13 – Liste des patients en visite

## Ajouter des diagnostics et symptômes

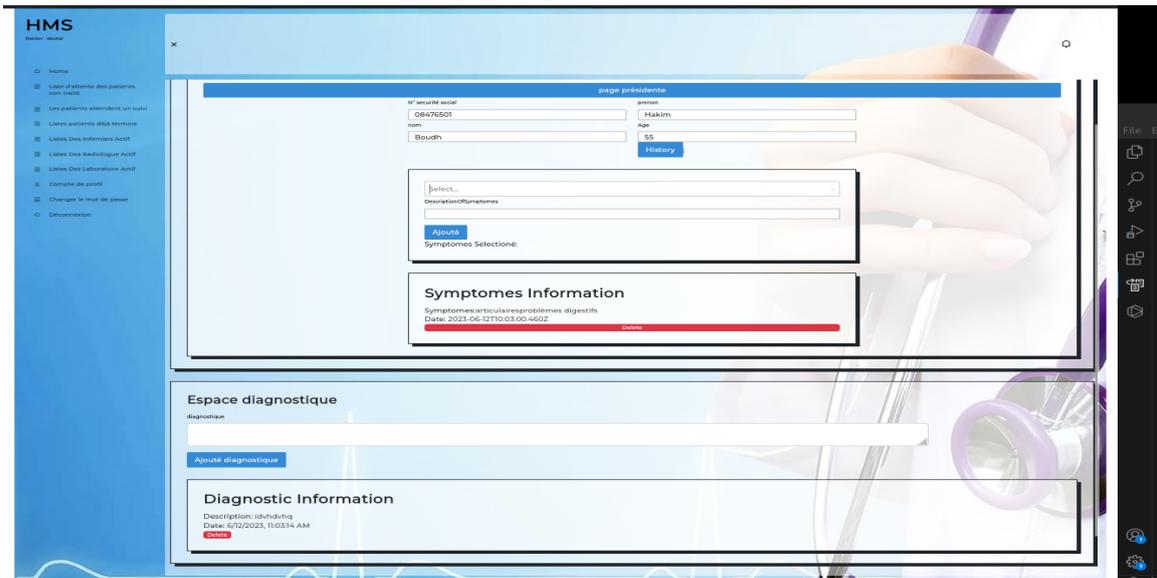


FIGURE 4.14 – Ajouter des diagnostics et symptômes

## Liste des patients en attente

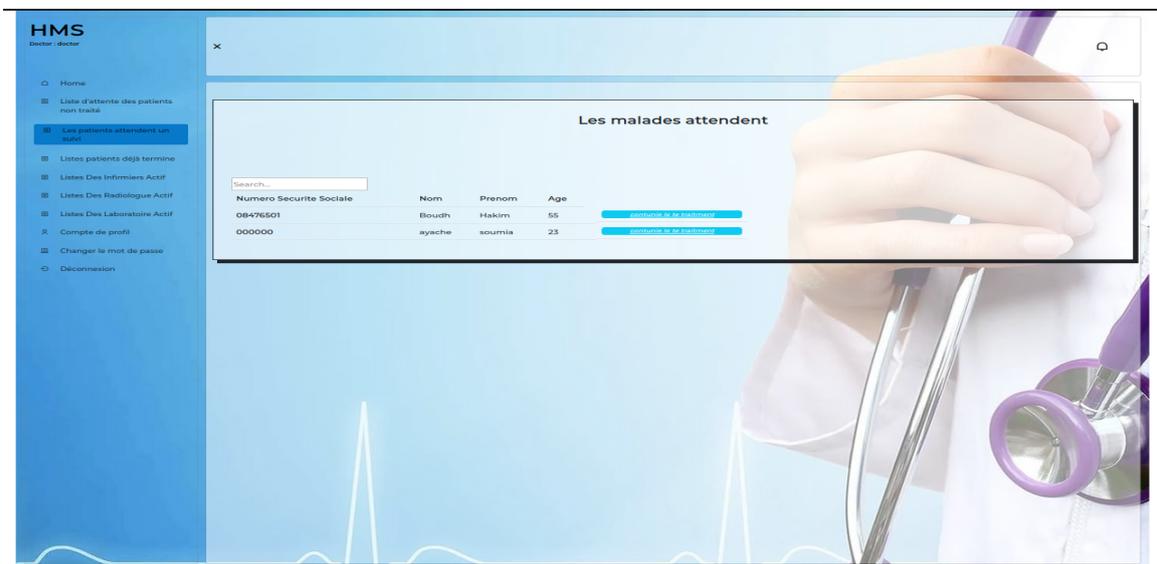


FIGURE 4.15 – Liste des patients en attente

## Continuer traitement

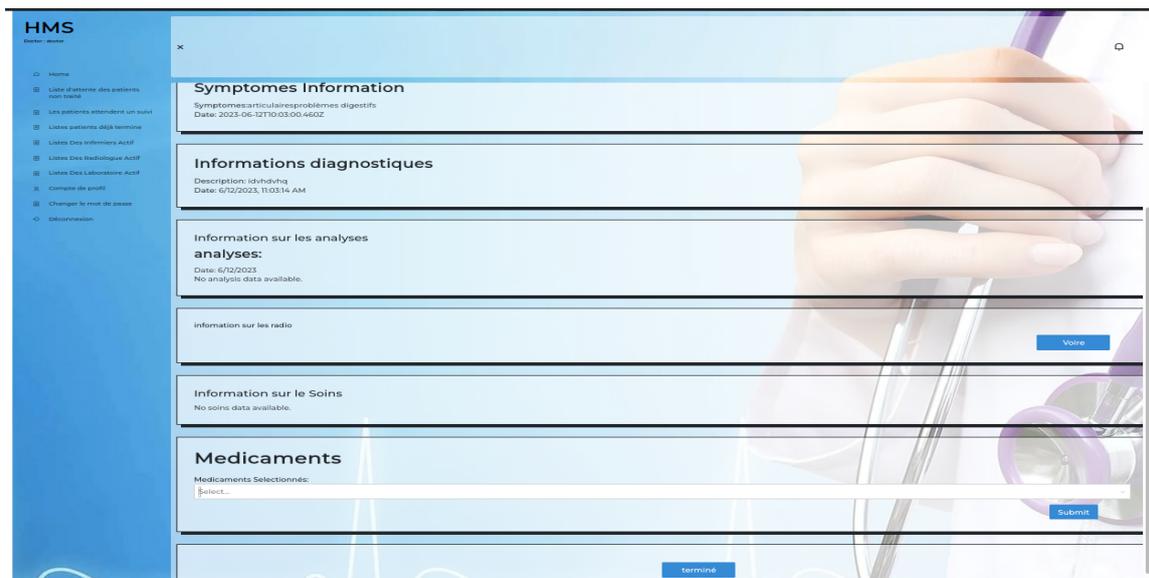


FIGURE 4.16 – le traitement suivant (continue le treatment)

## Liste des patients ayant terminé leur traitement



FIGURE 4.17 – Liste des patients ayant terminé leur traitement

## Sélection des types de traitement



FIGURE 4.18 – Les choix Selectionement des types de traitement

## Espace radiologue



FIGURE 4.19 – Espace radiologue

## Liste des patients concernés par la radio



FIGURE 4.20 – Liste des patients concernés par la radio

## Ajouter une radio par le radiologue

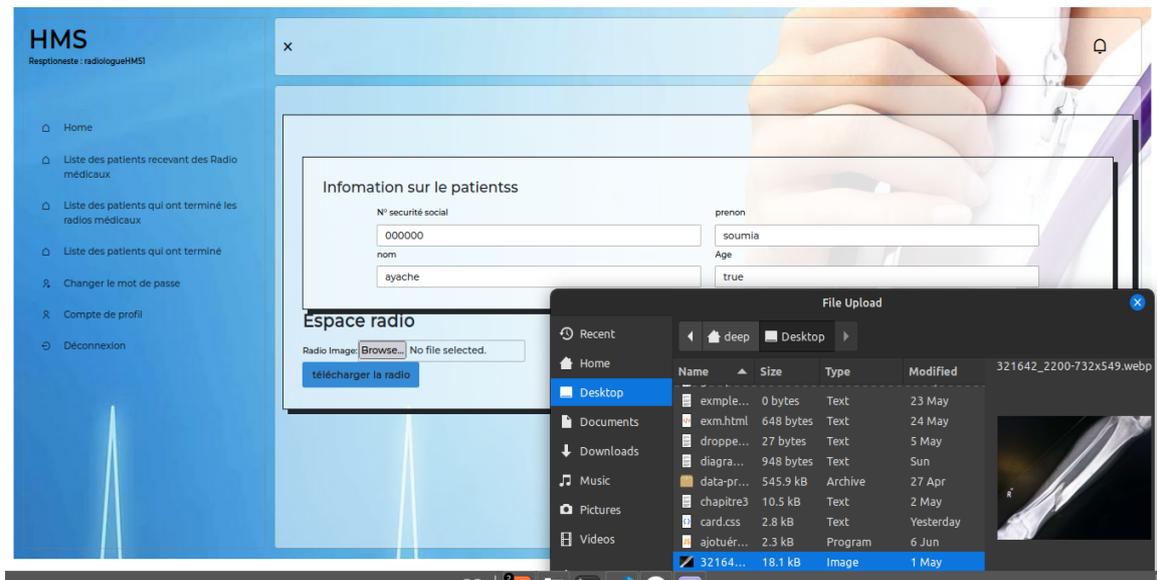


FIGURE 4.21 – Ajouter une radio

## Espace infirmier

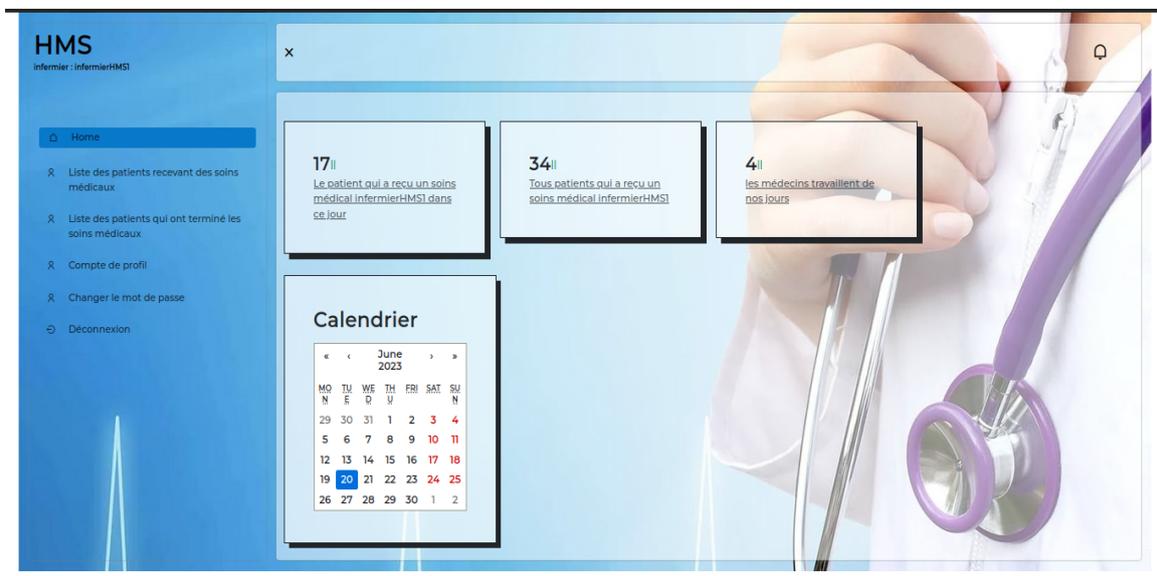


FIGURE 4.22 – Espace infirmier

## Liste des patients concernés par un soin

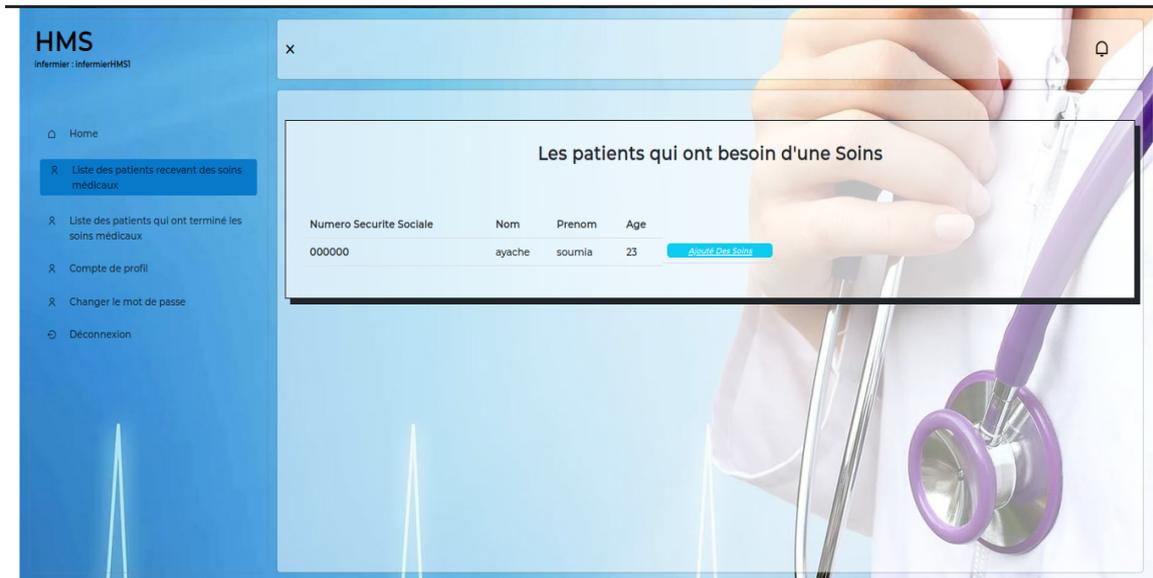


FIGURE 4.23 – Liste des patients concernés par un soin

## Faire le soins par l’infirmier

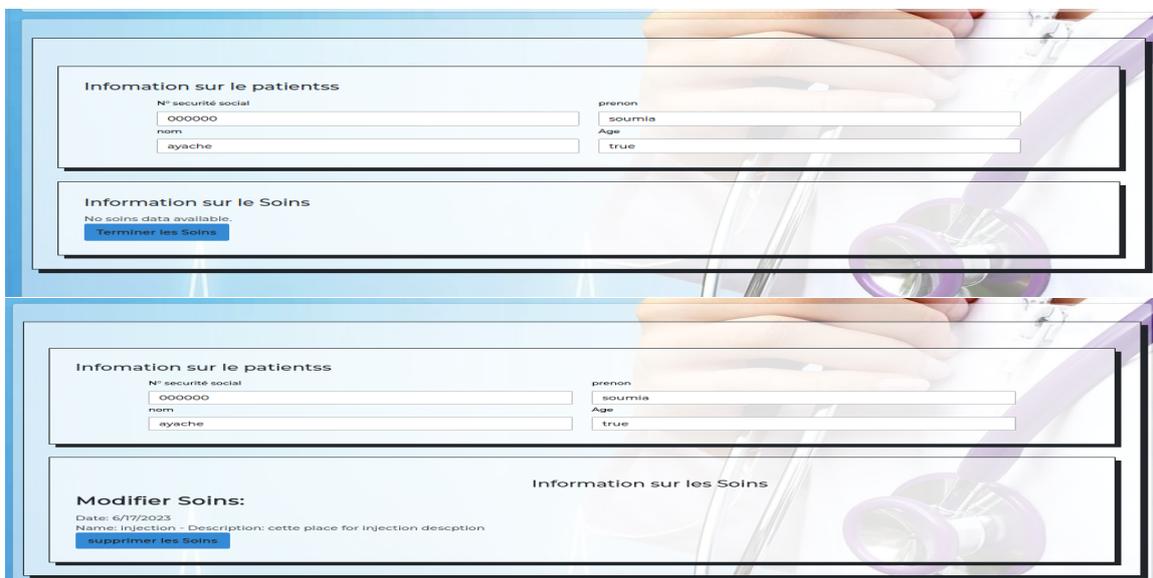


FIGURE 4.24 – Faire le soins par l’infirmier

### Liste des patients ayant terminé leur soins

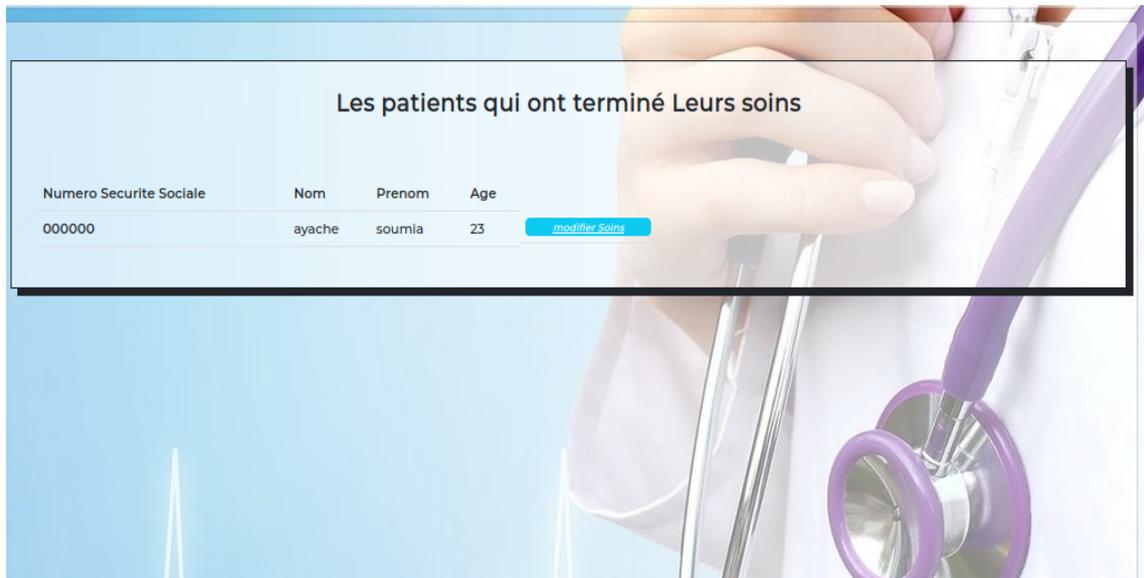


FIGURE 4.25 – Liste des patients ayant terminé leur soins

### Espace thechnicien de laboratoire

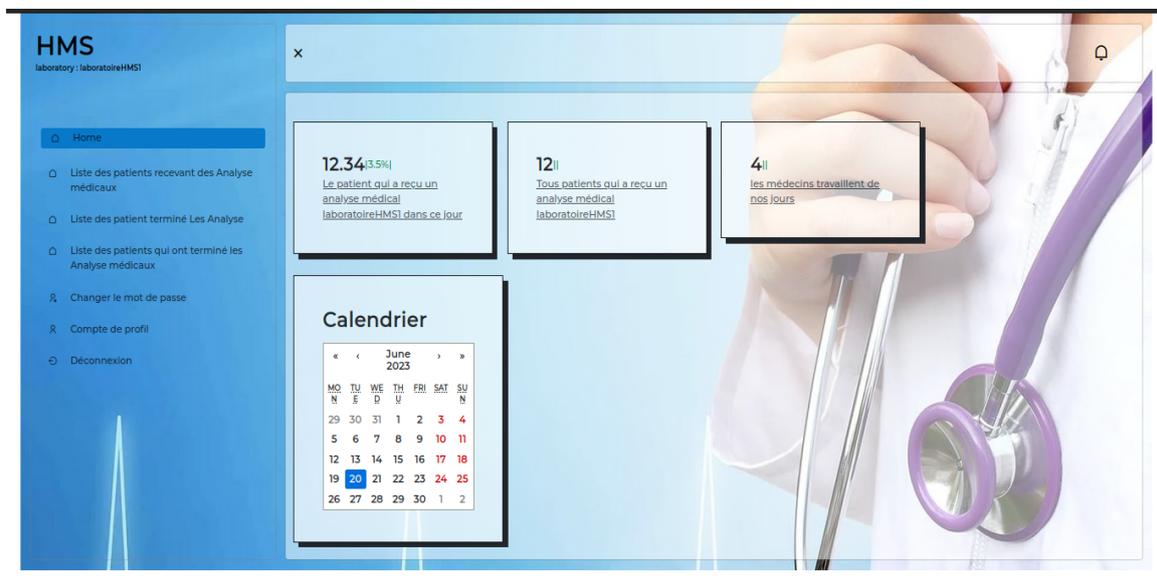


FIGURE 4.26 – Espace thechnicien de laboratoire

## Liste des patients concernés par l'analyse

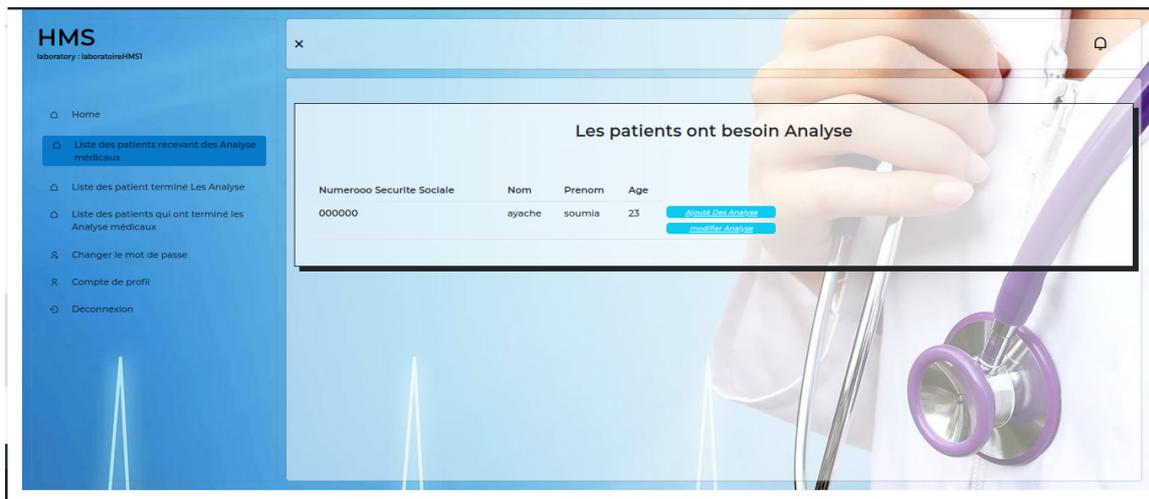


FIGURE 4.27 – Liste des patients concernés par l'analyse

## Ajoute et Modifier l'analyse



FIGURE 4.28 – Ajoute et Modifier l'analyse

### Liste des patients ayant terminé leur analyses

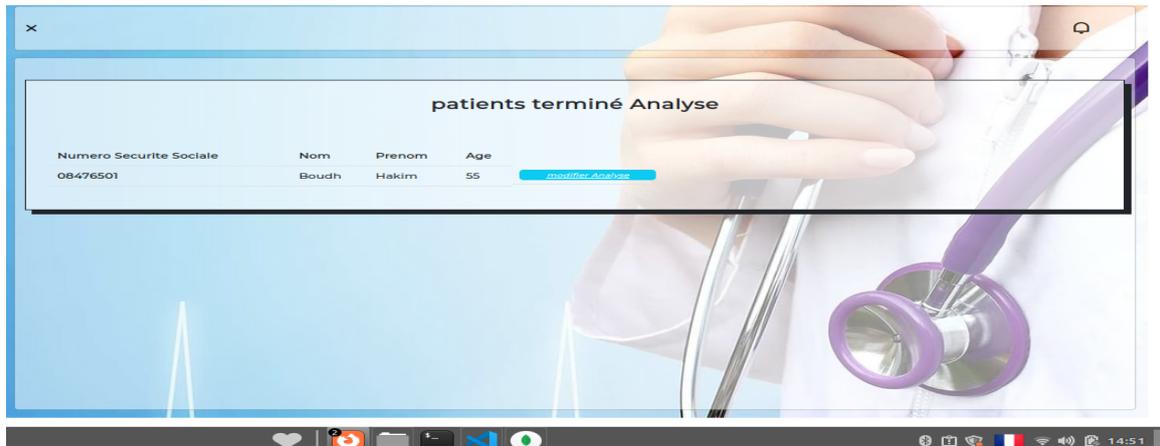


FIGURE 4.29 – Liste des patients ayant terminé leur analyses

### Espace réceptionniste

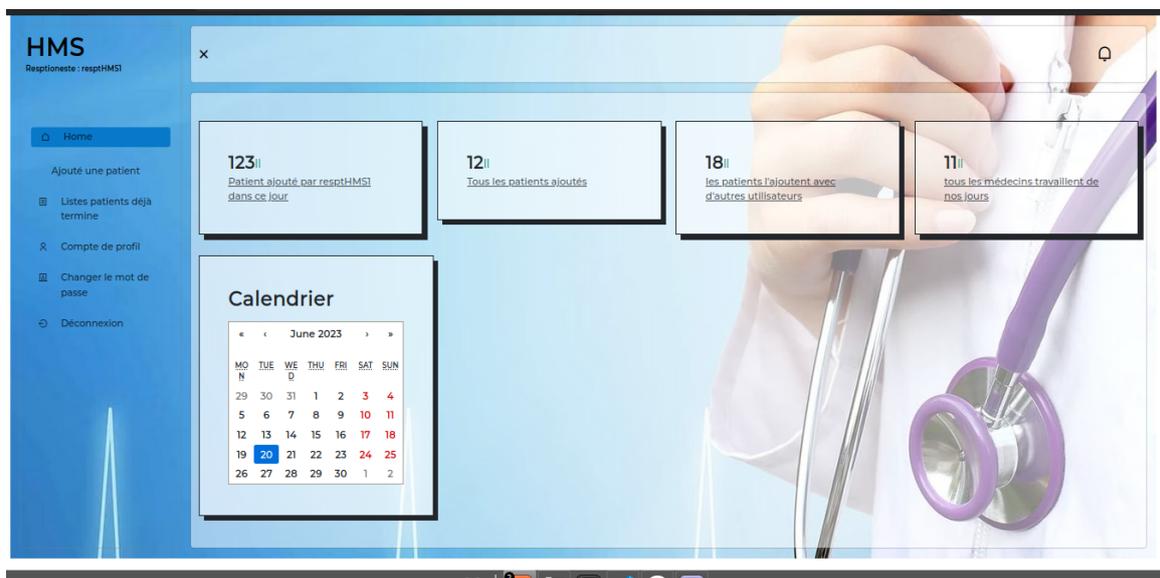


FIGURE 4.30 – Espace resptionniste

## Ajouter et Voir la liste patients

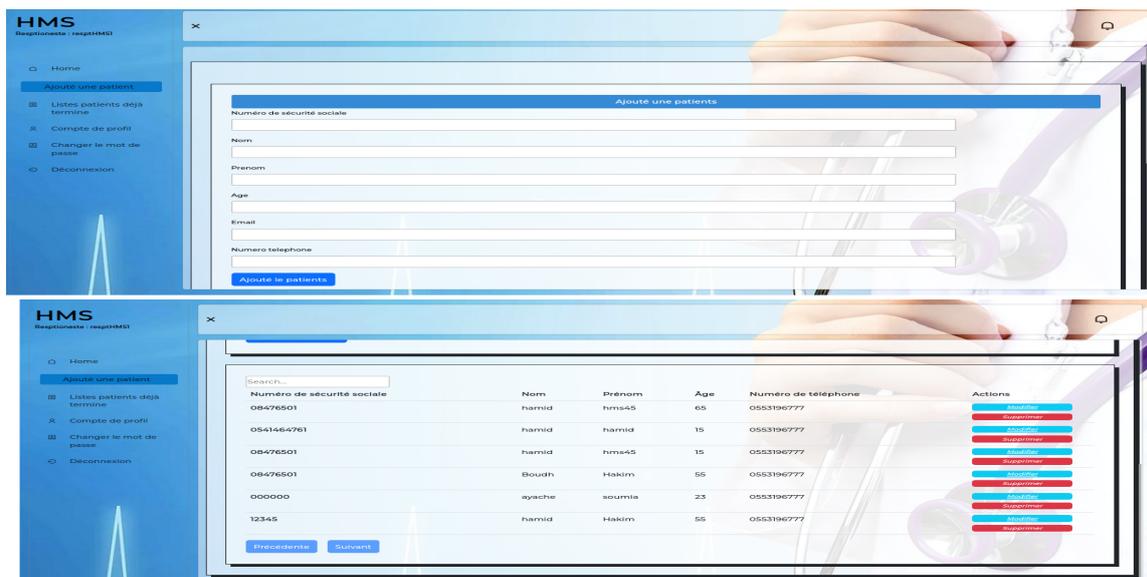


FIGURE 4.31 – Ajouter patients

## Liste des patients ayant terminé leur traitement par le médecin

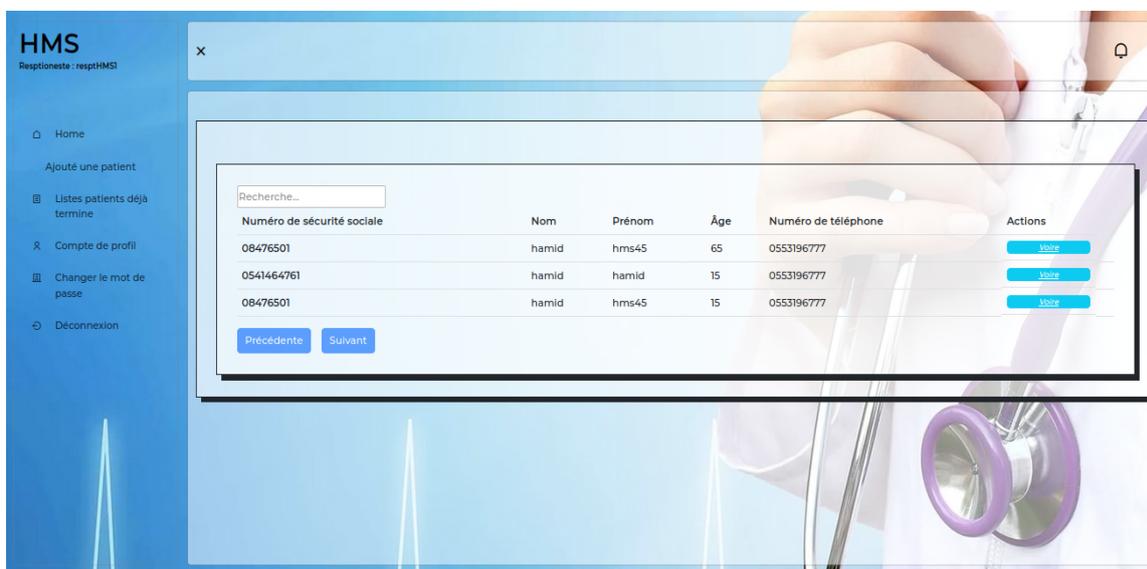


FIGURE 4.32 – Liste des patients ayant terminé leur traitement par le médecin

## Modifier les informations des patients

The screenshot shows a web application interface for HMS. On the left is a navigation menu with options: Home, Ajouté une patient, Listes patients déjà termine, Compte de profil, Change le mot de passe, and Déconnexion. The main content area displays a form for editing patient information. The form fields are as follows:

Numero Securite Sociale	08476501
Nom	hamid
Prenom	hms45
Age	65
email	booynai88@gmail.com
Phone	0553196777

At the bottom of the form is a blue button labeled "modifier et ferme".

FIGURE 4.33 – Modifier les informations des patients

## Afficher le profil de l'utilisateur

The screenshot shows the user profile page in the HMS application. The navigation menu on the left includes: Home, Ajouter un hospital, Liste des admins, Radios data, Analyse data, change le mot de passe, and Compte de profil (which is highlighted in blue). The main content area displays the user's profile information in a table:

Information de l'utilisateur	
Nom	Prenom
SuperAdmin	SuperAdmin
Date de Naissance	Numero de telephone
2023-06-15T00:00:00.000Z	0553196777
service	true

FIGURE 4.34 – page de profil des users

## Changer le mot de passe

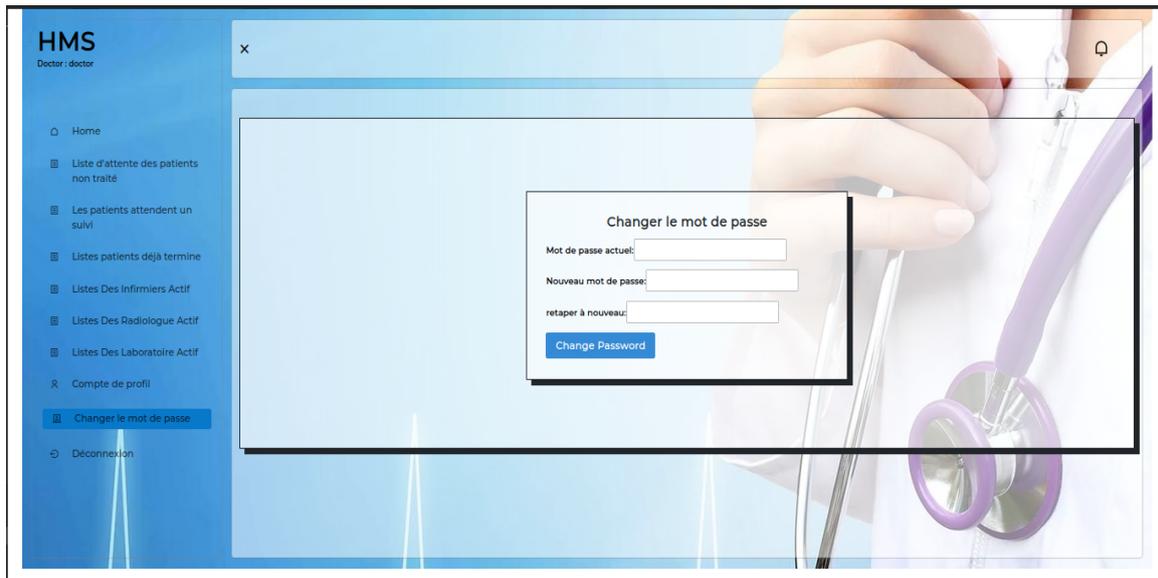


FIGURE 4.35 – Changer le mot de passe

## Integration des cliniques privé

### Interface de login pour les clinique privé



FIGURE 4.36 – Login pour des cliniques privé

### Liste des utilisateurs

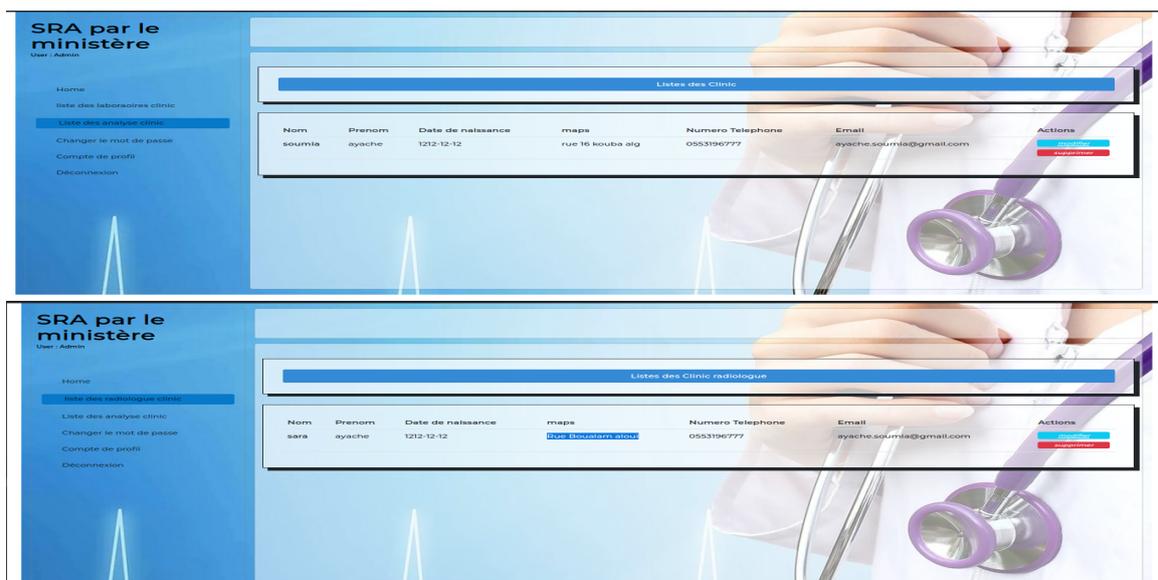


FIGURE 4.37 – Liste des utilisateurs

Espace pour scanner le qr-code pour ajouter les analyses des patients

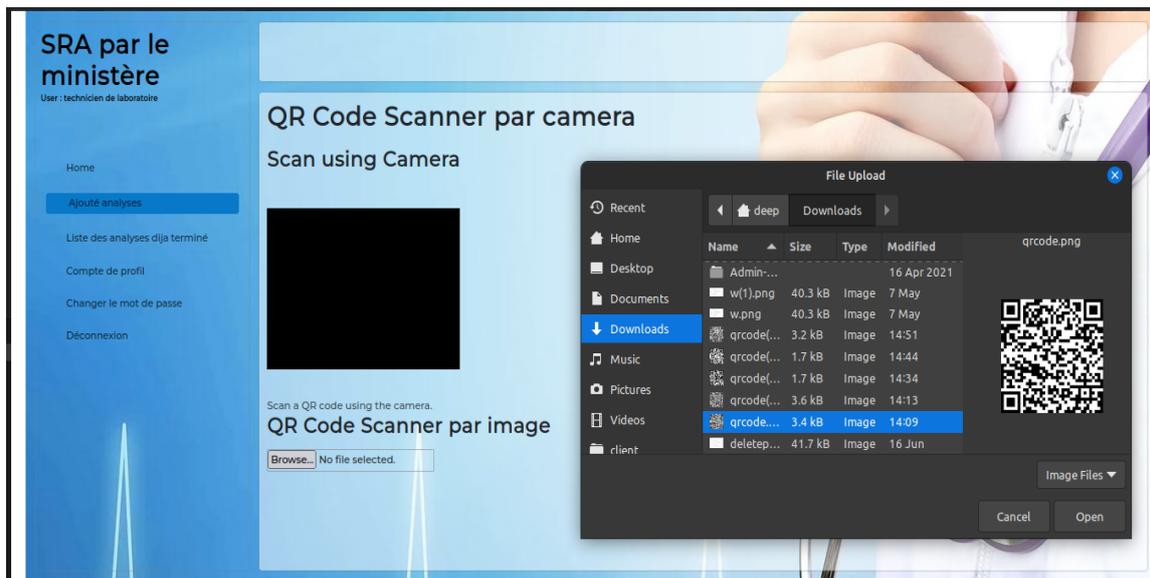


FIGURE 4.38 – scanner le qr-code

Ajouter les infomatients des analyses

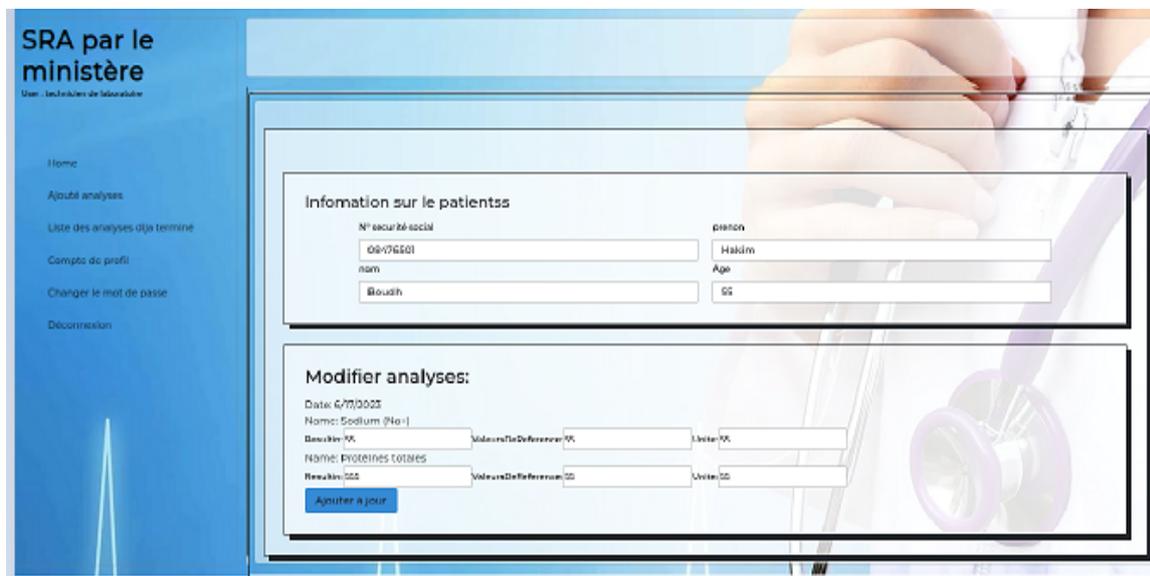


FIGURE 4.39 – Ajouter des analyses

## Espace pour scanner le qr-code pour ajouter les radiographies des patients

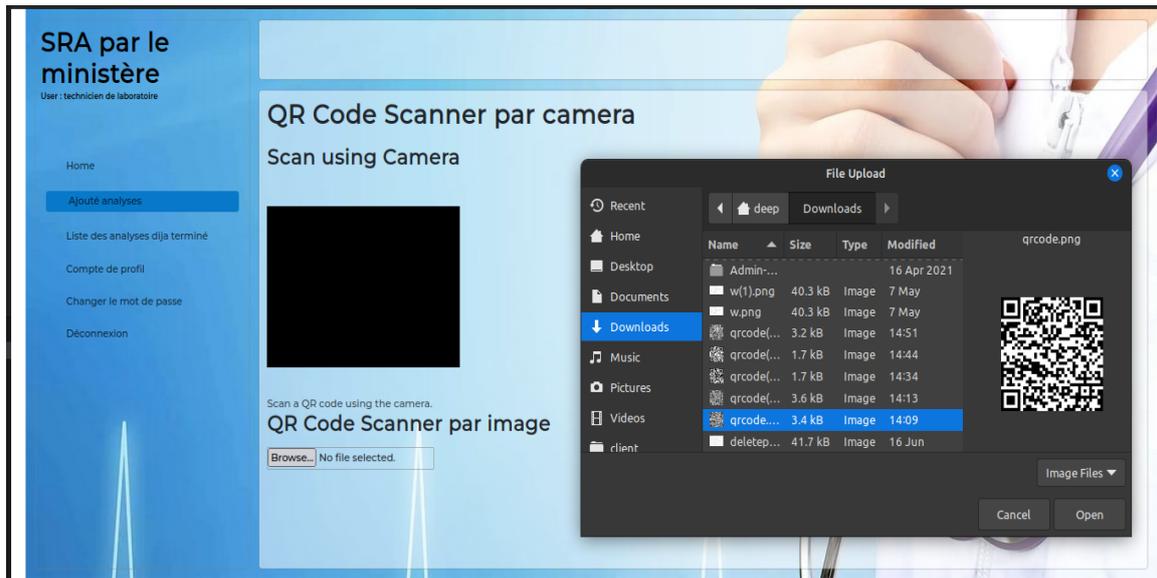


FIGURE 4.40 – Scanner le qr-code

## Uploader des radiographies



FIGURE 4.41 – Uploader des radiographies

## Espace Deep learning

### Extrait de l'exécution deep learning des radiographies

```

machinelearning.py > ...
43
44 #La première ligne définit la taille d'image attendu pour le modèle. Ici, chaque image doit avoir une taille de 224x224 pixels et 3 canaux de couleur
45 image_shape = (224, 224, 3)
46
47
48 #La deuxième ligne définit la taille des lots (ou mini-lots) d'images qui seront utilisés pour l'entraînement du modèle. Ici, chaque lot contient 32 im
49 batch_size = 32
50
51 #La troisième ligne définit le nombre d'époques (ou itérations) d'entraînement que le modèle effectuera sur toutes les données d'entraînement.
52 epochs = 10
53
54 train_datagen = ImageDataGenerator(rescale=1./255)
55 train_dataset = train_datagen.flow_from_directory(train_dir, target_size=image_shape[:2], batch_size=batch_size, class_mode='categorical')
56
57 val_datagen = ImageDataGenerator(rescale=1./255)
58 val_dataset = val_datagen.flow_from_directory(val_dir, target_size=image_shape[:2], batch_size=batch_size, class_mode='categorical')
59
60 model = tf.keras.models.Sequential([
61     tf.keras.layers.experimental.preprocessing.Rescaling(1./255, input_shape=image_shape),
62     tf.keras.layers.Conv2D(filters=32, kernel_size=3, activation='relu', input_shape=image_shape),
63     tf.keras.layers.MaxPooling2D(pool_size=(2, 2)),
64     tf.keras.layers.Flatten(),
65     tf.keras.layers.Dense(units=2, activation='softmax')
66 ])
67
68 model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
69 history = model.fit(train_dataset, epochs=epochs, validation_data=val_dataset)
70
71 print("Training complete.")
72
73 model = tf.keras.models.load_model('/home/deep/')

```

```

deep@deep-Lenovo-G50-80:~/xray-app$ ./xray-app.py /usr/bin/env /bin/python3 /home/deep/.vscode/extensions/ms-python.python-2023.8.0/pythonFiles/lib/python/debugpy/adapter
./xray-app.py/launcher:49111 -- /home/deep/xray-app/machinelearning.py
2023-08-17 15:08:17.151075948: I tensorflow/compiler/xla/platform/cpu_feature_guard.cc:193] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library
to enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
2023-08-17 15:08:01.100932: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libcudart.so.11.0':
dlopen: libcudart.so.11.0: cannot open shared object file: No such file or directory
2023-08-17 15:08:01.162921: I tensorflow/compiler/xla/stream_executor/cuda/cudart_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on
your machine.
2023-08-17 15:08:08.827991: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libwinifor.so.7': dl
error: libwinifor.so.7: cannot open shared object file: No such file or directory
2023-08-17 15:08:08.827991: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libwinifor-plugin.so
.7': dlopen: libwinifor-plugin.so.7: cannot open shared object file: No such file or directory
2023-08-17 15:08:08.827991: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libwinifor-plugin.so
.7': dlopen: libwinifor-plugin.so.7: cannot open shared object file: No such file or directory
2023-08-17 15:08:08.827991: W tensorflow/compiler/xla/stream_executor/cuda/cuda_driver.cc:265] failed call to cuInit: UNKNOWN ERROR (303)
TF Runtime warning: Cannot dlopen some TensorRT libraries. If you would like
to use Nvidia GPU with TensorRT, please make sure the missing libraries mentioned above are installed properly.
Found 690 images belonging to 2 classes.
2023-08-17 15:08:19.439515: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libcuda.so.1': dlerr
or: libcuda.so.1: cannot open shared object file: No such file or directory
2023-08-17 15:08:19.439515: W tensorflow/compiler/xla/stream_executor/cuda/cuda_driver.cc:265] failed call to cuInit: UNKNOWN ERROR (303)
2023-08-17 15:08:19.439723: I tensorflow/compiler/xla/stream_executor/cuda/cuda_diagnostics.cc:156] kernel driver does not appear to be running on this host
(deep-Lenovo-G50-80) /proc/driver/nvidia/version does not exist
2023-08-17 15:08:19.452768: I tensorflow/core/platform/cpu_feature_guard.cc:193] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library
(foundNN) to use the following CPU instructions in performance-critical operations: AVX2 FMA
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
2023-08-17 15:08:20.236268: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 19267584 exceeds 10% of free system memory.
Epoch 1/10
2023-08-17 15:08:22.752918: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 19267584 exceeds 10% of free system memory.
2023-08-17 15:08:22.883486: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 201867264 exceeds 10% of free system memory.
2023-08-17 15:08:23.151192: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 19267584 exceeds 10% of free system memory.
2023-08-17 15:08:23.648733: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 59466816 exceeds 10% of free system memory.
259/277 [-----] . ETA: 13s - loss: 0.6566 - accuracy: 0.6006

```

FIGURE 4.42 – Lancer l'exécution pour créer le modèle

```

machinelearning.py > ...
43
44 #La première ligne définit la taille d'image attendu pour le modèle. Ici, chaque image doit avoir une taille de 224x224 pixels et 3 canaux de couleur
45 image_shape = (224, 224, 3)
46
47
48 #La deuxième ligne définit la taille des lots (ou mini-lots) d'images qui seront utilisés pour l'entraînement du modèle. Ici, chaque lot contient 32 im
49 batch_size = 32
50
51 #La troisième ligne définit le nombre d'époques (ou itérations) d'entraînement que le modèle effectuera sur toutes les données d'entraînement.
52 epochs = 10
53
54 train_datagen = ImageDataGenerator(rescale=1./255)
55 train_dataset = train_datagen.flow_from_directory(train_dir, target_size=image_shape[:2], batch_size=batch_size, class_mode='categorical')
56
57 val_datagen = ImageDataGenerator(rescale=1./255)
58 val_dataset = val_datagen.flow_from_directory(val_dir, target_size=image_shape[:2], batch_size=batch_size, class_mode='categorical')
59
60 model = tf.keras.models.Sequential([
61     tf.keras.layers.experimental.preprocessing.Rescaling(1./255, input_shape=image_shape),
62     tf.keras.layers.Conv2D(filters=32, kernel_size=3, activation='relu', input_shape=image_shape),
63     tf.keras.layers.MaxPooling2D(pool_size=(2, 2)),
64     tf.keras.layers.Flatten(),
65     tf.keras.layers.Dense(units=2, activation='softmax')
66 ])
67
68 model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
69 history = model.fit(train_dataset, epochs=epochs, validation_data=val_dataset)
70
71 print("Training complete.")
72
73 model = tf.keras.models.load_model('/home/deep/')

```

```

deep@deep-Lenovo-G50-80:~/xray-app$ ./xray-app.py /usr/bin/env /bin/python3 /home/deep/.vscode/extensions/ms-python.python-2023.8.0/pythonFiles/lib/python/debugpy/adapter
./xray-app.py/launcher:49111 -- /home/deep/xray-app/machinelearning.py
2023-08-17 15:08:17.151075948: I tensorflow/compiler/xla/platform/cpu_feature_guard.cc:193] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library
to enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
2023-08-17 15:08:01.100932: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libcudart.so.11.0':
dlopen: libcudart.so.11.0: cannot open shared object file: No such file or directory
2023-08-17 15:08:01.162921: I tensorflow/compiler/xla/stream_executor/cuda/cudart_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on
your machine.
2023-08-17 15:08:08.827991: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libwinifor.so.7': dl
error: libwinifor.so.7: cannot open shared object file: No such file or directory
2023-08-17 15:08:08.827991: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libwinifor-plugin.so
.7': dlopen: libwinifor-plugin.so.7: cannot open shared object file: No such file or directory
2023-08-17 15:08:08.827991: W tensorflow/compiler/xla/stream_executor/cuda/cuda_driver.cc:265] failed call to cuInit: UNKNOWN ERROR (303)
TF Runtime warning: Cannot dlopen some TensorRT libraries. If you would like
to use Nvidia GPU with TensorRT, please make sure the missing libraries mentioned above are installed properly.
Found 690 images belonging to 2 classes.
2023-08-17 15:08:19.439515: W tensorflow/compiler/xla/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'libcuda.so.1': dlerr
or: libcuda.so.1: cannot open shared object file: No such file or directory
2023-08-17 15:08:19.439515: W tensorflow/compiler/xla/stream_executor/cuda/cuda_driver.cc:265] failed call to cuInit: UNKNOWN ERROR (303)
2023-08-17 15:08:19.439723: I tensorflow/compiler/xla/stream_executor/cuda/cuda_diagnostics.cc:156] kernel driver does not appear to be running on this host
(deep-Lenovo-G50-80) /proc/driver/nvidia/version does not exist
2023-08-17 15:08:19.452768: I tensorflow/core/platform/cpu_feature_guard.cc:193] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library
(foundNN) to use the following CPU instructions in performance-critical operations: AVX2 FMA
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
2023-08-17 15:08:20.236268: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 19267584 exceeds 10% of free system memory.
Epoch 1/10
2023-08-17 15:08:22.752918: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 19267584 exceeds 10% of free system memory.
2023-08-17 15:08:22.883486: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 201867264 exceeds 10% of free system memory.
2023-08-17 15:08:23.151192: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 19267584 exceeds 10% of free system memory.
2023-08-17 15:08:23.648733: W tensorflow/tsl/framework/cpu_allocator_impl.cc:82] Allocation of 59466816 exceeds 10% of free system memory.
259/277 [-----] . ETA: 13s - loss: 0.6566 - accuracy: 0.6006

```

FIGURE 4.43 – Fin d'exécution

## Interface de deep learning des radiographies



FIGURE 4.44 – la première étape

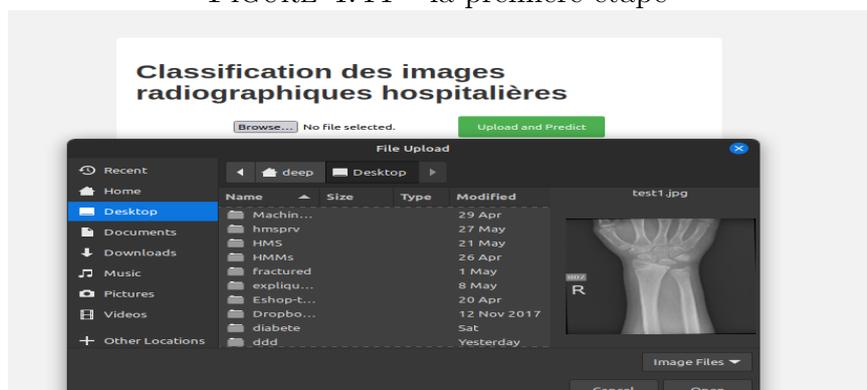


FIGURE 4.45 – la deuxième étape



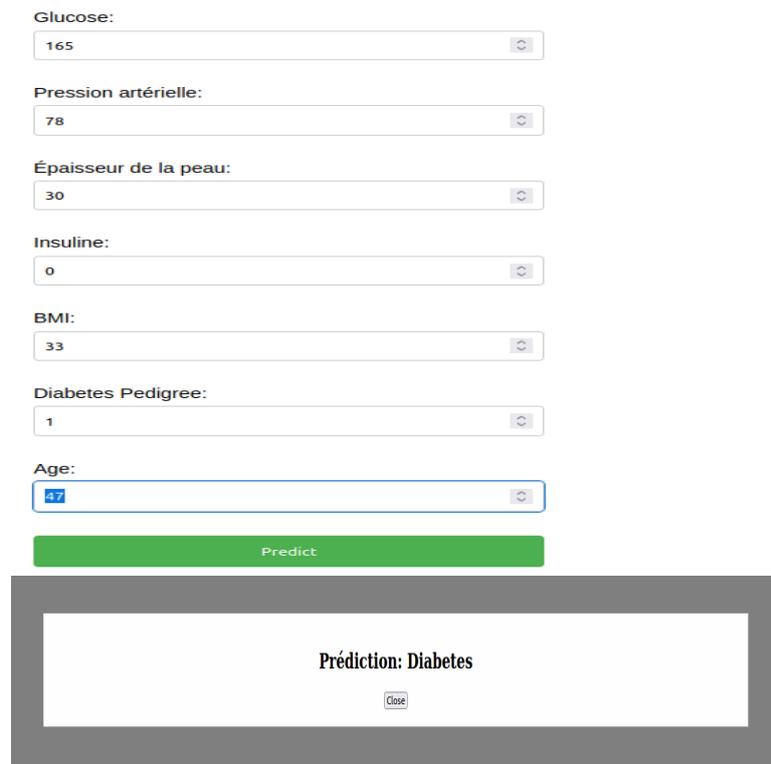
FIGURE 4.46 – la troisième étape

## Extrait de l'exécution deep learning des analyses Diabétiques

```

43
44 # La première ligne définit la taille d'image attendue pour le modèle. Ici, chaque image doit avoir une taille de 224x224 pixels et 3 canaux de couleur
45 image_shape = (224, 224, 3)
46
47 # La deuxième ligne définit la taille des lots (ou mini-lots) d'images qui seront utilisés pour l'entraînement du modèle. Ici, chaque lot contient 32 im
48 batch_size = 32
49
50 # La troisième ligne définit le nombre d'époques (ou itérations) d'entraînement que le modèle effectuera sur toutes les données d'entraînement.
51 epochs = 10
52
53 train_datagen = ImageDataGenerator(rescale=1./255)
54 train_datagen = train_datagen.flow_from_directory(train_dir, target_size=image_shape[:2], batch_size=batch_size, class_mode='categorical')
55
56 val_datagen = ImageDataGenerator(rescale=1./255)
57 val_datagen = val_datagen.flow_from_directory(val_dir, target_size=image_shape[:2], batch_size=batch_size, class_mode='categorical')
58
59 model = tf.keras.models.Sequential([
60     tf.keras.layers.experimental.preprocessing.Rescaling(1./255, input_shape=image_shape),
61     tf.keras.layers.Conv2D(filters=32, kernel_size=3, activation='relu', input_shape=image_shape),
62     tf.keras.layers.MaxPooling2D(pool_size=(2, 2)),
63     tf.keras.layers.Dense(units=2, activation='softmax')
64 ])
65
66 model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
67
68 history = model.fit(train_datagen, epochs=epochs, validation_data=val_datagen)
69
70 print("Training complete.")
71
72 model = tf.keras.models.load_model('/home/deep/')
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2
```

## Les étapes des tests diabétique avec DEEP LEARNING



Glucose: 165

Pression artérielle: 78

Épaisseur de la peau: 30

Insuline: 0

BMI: 33

Diabetes Pedigree: 1

Age: 47

Predict

Prédiction: Diabetes

Close

FIGURE 4.49 – Les étapes des tests Diabétique avec DEEP LEARNING

Dans cette partie les médecin et les les patients ils ont l'accès de remplir le formulaire et voir le résultat (diabétique ou non).

## 4.4 Conclusion

Dans ce dernier chapitre, nous avons mis en avant tout ce qui est langages, framework, outils et logiciels qui ont contribué à l'élaboration de notre application ainsi que ces principales fonctionnalités en exposant quelques interfaces accompagnées de leurs descriptions afin d'avoir une vue générale de notre application.

# Conclusion générale

En conclusion, ce mémoire a exploré l'utilisation des solutions Big Data dans le domaine de la santé en Algérie. Les résultats indiquent que l'adoption de ces technologies offre des avantages potentiels tels que l'amélioration de la gestion des données médicales, la prise de décision éclairée et l'amélioration des soins de santé.

Cependant, certaines considérations doivent être prises en compte pour une mise en œuvre réussie. La confidentialité des données médicales est un aspect crucial qui nécessite des mesures de sécurité adéquates pour protéger les informations sensibles des patients. De plus, la formation du personnel de santé à l'utilisation des solutions Big Data est essentielle pour en tirer pleinement parti.

En adoptant les solutions Big Data, l'Algérie peut améliorer l'efficacité des soins de santé, optimiser la gestion des ressources et améliorer la qualité des services médicaux. Cependant, cela nécessitera une planification stratégique, des investissements appropriés dans l'infrastructure technologique et la collaboration entre les professionnels de la santé et les experts en technologie.

En fin de compte, l'utilisation des solutions Big Data dans le domaine de la santé en Algérie peut contribuer à moderniser le système de santé, à améliorer les résultats des patients et à répondre aux défis croissants auxquels le secteur est confronté. Il est essentiel d'encourager la sensibilisation, la recherche et l'adoption progressive de ces technologies pour bénéficier pleinement de leurs avantages potentiels.

# Bibliographie

- [1] Grady Booch, James Rumbaugh, and Ivar Jacobson. The unified modeling language user guide. *Addison-Wesley Professional*, 1998.
- [2] Kenneth N Cukier and Viktor Mayer-Schönberger. Big data is better data. *Wall Street Journal*, 26 :C1, 2013.
- [3] Kenneth N Cukier and Viktor Mayer-Schönberger. Big data is better data. *Wall Street Journal*, 26 :C1, 2013.
- [4] Jean Dupont and Marie Martin. Les défis du big data. *Journal of Big Data*, 5(1) :1–8, 2018.
- [5] Jean Dupont and Marie Martin. Les avantages du big data. *Journal de la technologie*, 7(1) :12–20, 2021.
- [6] Marie Dupont. Big data et la santé. *Revue Médicale de France*, 73(2) :45–50, 2017.
- [7] Martin Kleppmann. Data-intensive applications with the kappa architecture. In *Proceedings of the 11th ACM SIGPLAN-SIGOPS international conference on Distributed systems*, pages 1–1. ACM, 2015.
- [8] Jean-Baptiste Lagrange. *Big Data : définitions et architectures*. Dunod, 2017.
- [9] Nathan Marz and James Warren. Big data : Principles and best practices of scalable realtime data systems. In *Manning Publications Co.*, 2015.
- [10] Sourav Mukherjee and Shilpa Bhatia. Big data architecture : A complete guide. *Journal of Big Data*, 2(1) :1–29, 2015.

- [11] Nicolas Rodondi, Arnaud Chiolero, Ana-Maria Mocanu, Fabrice Lillo-Le Louët, Olivier Grosгурin, Pascal Bovet, and Fred Paccaud. Le big data en santé : enjeux et perspectives. *Revue médicale suisse*, 14(625-626) :1476–1481, 2018.
- [12] Nicolas Rodondi, Arnaud Chiolero, Ana-Maria Mocanu, Fabrice Lillo-Le Louët, Olivier Grosгурin, Pascal Bovet, and Fred Paccaud. Le big data en santé : enjeux et perspectives. *Revue médicale suisse*, 14(625-626) :1476–1481, 2018.

# Annexe 1

**LABORATOIRE D'ANALYSES MEDICALES.**  
Dr BERKAT  
MEDECIN BIOLOGISTE  
Tel : +213 (0)659 78 15 21 - Email : berkatalab@gmail.com

QR code | Edité le : 23/04/2023

NID : 2023423-001580  
Patient (e) : AYACHE SARA  
Né (e) le : 23-4-1999 Age : 24 ans  
Prélèvement du : 23/04/2023 08:53:00

**Compte Rendu d'Analyses Médicales**

*Prélèvement non effectué au laboratoire*  
**Biochimie sanguine**

	Résultat	Unité	Valeurs Usuelles	Antécédent
CALCIUM SANGUIN	86.00	mg/L	86.00 - 103.00	
PROTIDES TOTAUX	71.0	g/L	66.0 - 83.0	
ALBUMINEMIE	40.2	g/L	35.0 - 52.0	
VITAMINE D2/D3 : 25 HYDROXY VITAMINE D	<u>&lt;8.1</u>	ng/mL	9.10 - 48.50	

**Hormonologie**

	Résultat	Unité	Valeurs Usuelles	Antécédent
TSHus	2.689	mUI/mL	0.350 - 5.100	
PTH	39.64	pg/L	12.00 - 88.00	

Laboratoire d'Analyses Médicales  
**Dr N. BERKAT**  
Médecin Biologiste  
N°6655/TZO

FIGURE 50 – Liste des analyses

## Annexe 2

**Etablissement Public Hospitalier d'Ain Bessem**  
**Laboratoire central de biologie – unité de biochimie**

Nom: Ayach Prénoms: Samir N°  
Date: Service:

Paramètres sanguins	Résultats	Valeurs de référence	unités
Glucose	/	0,7 – 1,1	g/l
Urée		0,15–0,50	g/l
Créatinine		H: 8–13 F: 5–12	mg/l
Acide urique		H: 35 – 72 F : 26 – 60	mg/l
Protéines totales		60 – 80	g/l
Albumine		35 – 50	g/l
Sodium (Na <sup>+</sup> )		135 -145	mEq/l
potassium (K <sup>+</sup> )	3,5 – 5	mEq/l	
Chlore (Cl <sup>-</sup> )	95 – 105	mEq/l	
Calcium	110	A: 85–105	mg/l
Magnésium		18 – 25	mg/l
Phosphore		25– 48	mg/l
Bilirubine totale	/	≤10	mg/l
Bilirubine directe		≤2	mg/l
Bilirubine indirecte		≤ 8	mg/l
ASAT		H: <50 F: <35	UI/l
ALAT		H: <50 F: <35	UI/l
GGT		H: 5-85 F: 5-55	UI/l
PAL		<125	UI/l
LDH	135–225	UI/l	
Lipase	< 60	UI/l	
Amylase			
Cholestérol		1,3 – 2	g/l
HDL cholestérol		H : > 0,45 F : >0,55	g/l
LDL cholestérol		<1,6	g/l
Triglycérides		0,5 – 1,5	g/l
CRP		< 6	mg/l

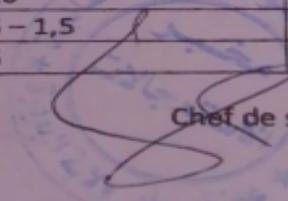
  
 Chef de service

FIGURE 51 – Liste des analyses

## ملخص

يتركز مشروع درجة الماجستير هذا على إنشاء تطبيق ويب يستخدم البيانات الضخمة والتعلم العميق لمعالجة التحديات في إدارة المستشفيات ومعالجة المرضى وتحليل الصور الشعاعية. من خلال استغلال التنقيب في البيانات والتحليلات والتعلم الآلي، يقوم التطبيق بإدارة البيانات المستشفى بكفاءة مع دمج قواعد البيانات الموزعة لتحقيق القدرة على التوسع. واجهات سهلة الاستخدام تمكن من التواصل المتسق بين طاقم المستشفى والأطباء والمرضى. يعمل التطبيق على مراقبة البيانات في الوقت الحقيقي والتنبؤ بالأمراض والتقارير التلقائية لتحسين الكفاءة. تحليل التعلم العميق يعزز التشخيصات الدقيقة والكشف المبكر عن الأمراض وتوصيات العلاج الشخصية. يهدف التطبيق إلى تعزيز إدارة المستشفيات ونتائج المرضى واتخاذ القرارات القائمة على البيانات، مما يقلل من التكاليف والأعباء الإدارية.

## Abstract

This master's degree project focuses on creating a web application that utilizes big data and deep learning to address challenges in hospital management, patient handling, and radiographic image analysis. By leveraging data mining, analytics, and machine learning, the application efficiently manages hospital data while integrating distributed databases for scalability. User-friendly interfaces enable seamless communication among hospital staff, doctors, and patients. Real-time data monitoring, disease forecasting, and automated reporting improve efficiency. Deep learning analysis enhances accurate diagnoses, early disease detection, and personalized treatment recommendations. The application aims to enhance hospital management, patient outcomes, and data-driven decision-making, reducing costs and administrative burdens.

**Key words :** internet, web, ...

# Résumé

Ce projet de Master se concentre sur la création d'une application Web qui utilise le big data et l'apprentissage approfondi pour relever les défis de la gestion hospitalière, de la prise en charge des patients et de l'analyse d'images radiographiques. En exploitant l'exploration de données, les analyses et l'apprentissage automatique, l'application gère efficacement les données hospitalières tout en intégrant des bases de données distribuées pour assurer la scalabilité. Des interfaces conviviales permettent une communication fluide entre le personnel hospitalier, les médecins et les patients. La surveillance des données en temps réel, les prévisions des maladies et les rapports automatisés améliorent l'efficacité. L'analyse de l'apprentissage approfondi améliore les diagnostics précis, la détection précoce des maladies et les recommandations de traitement personnalisées. L'application vise à améliorer la gestion hospitalière, les résultats des patients et la prise de décision basée sur les données, tout en réduisant les coûts et les charges administratives.

**Mots clés** : internet, web, ...