



# République du Sénégal

Un Peuple – Un but – Une foi

-----  
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR, DE LA RECHERCHE ET DE  
L'INNOVATION

UNIVERSITE ALIOUNE DIOP



N° .....

UFR Sciences Appliquées et Technologies de l'Information et de la  
Communication (SATIC)

Département Technologie de l'Information et de la Communication (TIC)

**Systemes et Réseaux (SR)**

**MEMOIRE DE FIN DE CYCLE**

**Présenté par :**

Serigne fallou DIA

Pour l'obtention du diplôme de

**Master en Systemes et Réseaux**

**SUJET : Mise en place d'un système de téléconsultation basé  
sur l'IA et l'IoT**

**Encadrant (s) : Pr Abdou khadre DIOP**

**Année académique 2022-2024**

## Dédicaces

À ma mère, dont l'amour inconditionnel et la sagesse m'ont guidée tout au long de ce parcours. À mon père, qui m'a inculqué la rigueur et la persévérance. À mes sœurs Mariama, Ami, Fama et à Dior . À tous mes frères, votre présence chaleureuse est mon plus beau soutien. Et à mes camarades de promo, avec qui j'ai partagé des moments inoubliables, ce mémoire est un hommage à notre amitié sincère.

## *Remerciement*

*Je tiens tout d'abord à exprimer ma profonde gratitude à Dieu pour m'avoir accordé la force et la persévérance nécessaires à l'accomplissement de ce travail à qui mon amour est sans limites.*

*Mes remerciements vont ensuite à mes parents, pour leur soutien inconditionnel et leurs sacrifices durant tous mes années scolaire depuis 2004.*

*Je suis également très reconnaissant envers le Pr **Abdou Khadre Diop** à qui j'ai été confiés durant toutes mes années universitaire pour ses précieux conseils et son encadrement.*

*Enfin, je souhaite remercier chaleureusement tous mes collègues pour leur aide et leur encouragement tout au long de ce projet.*

## Liste des figures

Figure 1: machine learning system.....	10
Figure 2: ML dans l'industrie.....	11
Figure 3: consultation à l'hôpital .....	11
Figure 4: apprentissage supervisé .....	13
Figure 5: apprentissage non supervisé .....	13
Figure 6: apprentissage par renforcement .....	14
Figure 7: réseau de neurones .....	15
Figure 8 application de Machine learning.....	16
Figure 9: intelligence artificielle .....	4
Figure 10: dates d'evolution de l'IA .....	5
Figure 11: comparaison IA fort IA faible.....	6
Figure 12: robot médical .....	8
Figure 13: la collecte de données .....	17
Figure 14: bot de collecte par IoT .....	18
Figure 15: collecte par enquête .....	20
Figure 16: data processing.....	21
Figure 17: bibliotheque .....	28
Figure 18: symptomes .....	29
Figure 19: maladie.....	29
Figure 20: pretraitement .....	30
Figure 21: visualisation .....	30
Figure 22: pretraitement .....	
Figure 23: linkage des symptômes .....	31
Figure 25: correlation des symptômes .....	31
Figure 26: correlation des donnees test.....	33
Figure 27: random forest .....	34
Figure 28: métrique .....	35
Figure 29: accuracy .....	36
Figure 30: naive bayes.....	37
Figure 31: accuracy .....	37
Figure 32: accuracy .....	38
Figure 33: architecture .....	41

Figure 34:flask.....	42
Figure 35: golang .....	43
Figure 36:interface graphique .....	45
Figure 37: application web .....	46
Figure 38: bouton web.....	46
Figure 39: prompt de l'application web.....	46

## Liste des tableaux

Tableau 1: liste des symptômes associés.....	24
Tableau 2: liste des maladies repertoriées .....	26

# 1 Table des matières

Dédicaces .....	ii
<i>Remerciement</i> .....	iii
Liste des figures .....	iv
Liste des tableaux .....	vi
1 Introduction Générale .....	i
1.1 Contexte et motivation.....	i
1.2 Problématique .....	i
1.3 Objectifs généraux et spécifiques de l'étude.....	2
1.4 Méthodologie Générale .....	2
Chapitre 1 : L'intelligence Artificielle .....	3
1.1 Introduction à l'intelligence artificielle.....	3
1.2 Historique de l'intelligence artificielle.....	4
1.3 Types d'intelligence artificielle .....	5
1.3.1 IA faible (IA étroite) .....	5
1.3.2 IA forte (IA générale).....	6
1.3.3 Apprentissage profond (Deep Learning) .....	6
1.4 Les approches en intelligence artificielle .....	7
1.4.1 Approche symbolique .....	7
1.4.2 Approche connexionniste .....	7
1.5 Applications de l'intelligence artificielle .....	7
1.6 Éthique et enjeux de l'intelligence artificielle.....	8
1.7 Conclusion .....	9
Chapitre 2 : Généralité sur le Machine Learning .....	10
2.2 Historique du Machine Learning .....	12
2.3 Types d'apprentissage automatique .....	12
2.3.1 Apprentissage supervisé.....	12
2.3.2 Apprentissage non supervisé :.....	13
2.3.3 Apprentissage par renforcement.....	13
2.4 Algorithme de Machine Learning.....	14
2.5 Applications du Machine Learning .....	15
2.6 Conclusion .....	16
Chapitre 3 : Collecte de données, les maladies répertoriés et certains de leurs symptômes ....	17
3.1 Introduction à la collecte de données .....	17

3.2	Méthodologie de la collecte de données .....	17
3.2.1	Sources de données .....	17
3.2.2	Validation des données.....	20
3.3	Prétraitement des données .....	20
3.3.1	Nettoyage des données .....	20
3.3.2	Transformation des données.....	21
3.4	Sélection des maladies et symptômes.....	21
3.4.1	Critères de sélection .....	21
3.4.2	Liste des symptômes étudiés .....	22
Chapitre 4 : Conception du système de prédiction de maladie .....		28
4.1	Introduction .....	28
4.2	Étapes de la modélisation .....	28
4.2.1	Prétraitement des données .....	28
4.3	Mise en place de la prédiction avec les algorithmes de prédiction.....	33
4.3.1	Algorithme des arbres de décision .....	34
4.3.2	Algorithme Random Forest.....	35
4.3.3	Algorithme Naive Bayes .....	36
4.3.4	Algorithme K-Nearest Neighbors (KNN) .....	38
4.4	Conclusion .....	39
Chapitre 5 : Mise en place d'une plateforme pour permettre l'interaction utilisateur .....		40
5.1	Introduction .....	40
5.2	Architecture globale du système.....	40
5.3	Technologies utilisées.....	41
5.4	Architecture du backend :.....	41
5.4.1	Flask .....	42
5.4.2	Golang .....	42
5.4.3	Communication entre les composants :.....	43
5.5	Architecture du frontend.....	44
5.5.1	HTML .....	44
5.5.2	CSS.....	44
5.5.3	JavaScript natif .....	44
5.6	Résultat final .....	45
5.7	Conclusion .....	47
6	Conclusion Générale .....	48
Références : .....		49





# 1 Introduction Générale

## 1.1 Contexte et motivation

La prédiction des maladies à partir des symptômes rapportés par les patients représente une avancée significative dans le domaine de la santé. Avec l'essor des technologies numériques et de l'intelligence artificielle, il est désormais possible de recueillir des données en temps réel et d'utiliser des algorithmes sophistiqués pour analyser ces informations. Cela permet non seulement d'améliorer la précision des diagnostics, mais aussi de rendre les soins de santé plus accessibles et personnalisés. En exploitant les données fournies par les utilisateurs, il devient possible de détecter des patterns et des corrélations qui échappent parfois à l'œil humain, ouvrant ainsi la voie à des interventions précoces et plus efficaces.

### **Importance de la prédiction des maladies**

La prédiction des maladies à partir des symptômes présente plusieurs avantages. Elle peut contribuer à réduire le fardeau des maladies en permettant une détection précoce, améliorer l'efficacité des traitements en facilitant des diagnostics plus précis, et rendre les services de santé plus accessibles en permettant aux utilisateurs de recevoir des avis médicaux sans avoir à se déplacer. De plus, cette approche peut jouer un rôle crucial dans la gestion des maladies chroniques et dans la prévention des épidémies. En somme, la prédiction des maladies via l'analyse des symptômes est une innovation prometteuse pour le futur de la santé publique.

## 1.2 Problématique

La problématique centrale de cette étude réside dans le développement d'un modèle de prédiction des maladies qui soit à la fois précis et accessible aux utilisateurs non spécialisés. Plus spécifiquement, comment peut-on concevoir un système capable d'interpréter les symptômes fournis par les utilisateurs pour prédire les maladies potentielles avec une fiabilité suffisante pour être utile dans un contexte clinique ou personnel ? Cette question soulève plusieurs défis, notamment en termes de collecte et de traitement des données, de conception d'algorithmes de prédiction, et d'intégration de ces outils dans des systèmes utilisables par le grand public.

### 1.3 Objectifs généraux et spécifiques de l'étude

Dans le but d'apporter des solutions à l'ensemble des problématiques soulignés en haut dans ce mémoire nous avons comme objectifs :

- ❖ **Objectif général :** Développer un modèle de prédiction des maladies basées sur les symptômes fournis par les utilisateurs grâce au machine learning et de l'intelligence artificielle.
- ❖ **Objectifs de la Liste :**
  - ❖ Identifier et sélectionner les données pertinentes pour la prédiction des maladies.
  - ❖ Concevoir et implémenter des algorithmes de machine learning capables d'analyser ces données de manière efficace.
  - ❖ Évaluer la performance des modèles prédictifs développés à l'aide de métriques appropriées.
  - ❖ Proposer une interface utilisateur intuitive permettant la saisie des symptômes et la visualisation des résultats de prédiction.

### 1.4 Méthodologie Générale

#### Aperçu de l'approche adoptée pour atteindre les objectifs

Pour atteindre les objectifs définis, notre étude sera articulée autour de quatre chapitres :

- ❖ **Le chapitre 1 :** Porte sur l'intelligence artificielle, les solutions existantes dans le domaine de la médecine.
- ❖ **Le chapitre 2 :** fait la généralité sur le machine Learning, la collecte et l'exploitation des données.
- ❖ **Le chapitre 3 :** articule sur la collecte de données, les maladies répertoriés, les symptômes et leurs manifestations.
- ❖ **Le chapitre 4 :** étale la modélisation et conception du système de prédiction de maladie.
- ❖ **Le chapitre 5 :** porte sur la mise en place d'une plateforme pour permettre l'interaction utilisateur.

# Chapitre 1 : L'intelligence Artificielle

## 1.1 Introduction à l'intelligence artificielle

L'intelligence artificielle (IA) est un domaine de l'informatique qui vise à créer des systèmes capables de simuler des processus cognitifs humains tels que l'apprentissage, le raisonnement, la perception et la prise de décision. L'IA a connu une croissance exponentielle au cours des dernières décennies, bouleversant des secteurs tels que la santé, la finance, et l'industrie. (Stuart Russell, 2010)

Née des rêves d'imiter l'intelligence humaine, il a connu une évolution fulgurante depuis les premiers travaux de Turing dans les années 1950. Après avoir traversé des périodes d'enthousiasme démesuré suivies de désillusions, l'IA connaît aujourd'hui un essor sans précédent grâce aux progrès de l'apprentissage automatique et de la puissance de calcul.

Les systèmes d'IA, qui s'appuient sur des algorithmes complexes capables d'apprendre à partir de données, sont désormais présents dans de nombreux aspects de notre vie quotidienne. De la reconnaissance vocale des assistants virtuels aux recommandations personnalisées des plateformes de streaming, en passant par les diagnostics médicaux assistés par ordinateur, l'IA transforme profondément notre société.

Cependant, cette révolution technologique soulève de nombreuses questions. Les biais algorithmiques, la perte d'emplois, la sécurité des systèmes d'IA et les questions éthiques liées à l'autonomie des machines sont autant de défis auxquels la société doit faire face. Il est donc essentiel de développer une réflexion approfondie sur les implications de l'IA et de mettre en place des garde-fous pour garantir un développement responsable de cette technologie.

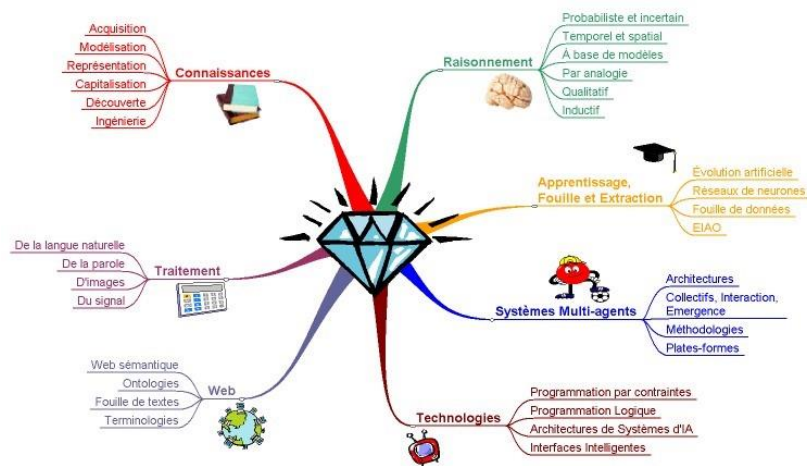


Figure 1: intelligence artificielle<sup>1</sup>

## 1.2 Historique de l'intelligence artificielle

L'idée de machines capables de penser remonte à l'antiquité, mais c'est au 20<sup>e</sup> siècle que l'IA a véritablement pris forme. Le terme "intelligence artificielle" fut officiellement introduit par John McCarthy. Depuis lors, l'IA a évolué avec le développement des réseaux de neurones, des systèmes experts et plus récemment du deep learning. (Nilsson, 1998)

**Les prémices** de l'IA moderne peuvent être situées dans les années 1940, avec les travaux d'Alan Turing sur la machine de Turing et les premiers modèles de calcul. En 1956, la conférence de Dartmouth marque un tournant décisif : c'est lors de cet événement que le terme "intelligence artificielle" est forgé par John McCarthy, et que les fondations théoriques de cette nouvelle discipline sont posées.

Les **années 1960 et 1970** sont marquées par un grand optimisme et de nombreux projets ambitieux, tels que les systèmes experts et les programmes de traduction automatique. Cependant, les limites technologiques de l'époque et la complexité de l'intelligence humaine ont rapidement mis en évidence les difficultés de la tâche. Cette période est parfois appelée le "premier hiver de l'IA".

**Les années 1980** voient l'émergence des systèmes experts, des programmes capables de raisonner et de prendre des décisions dans des domaines spécifiques. Ces systèmes ont connu

un certain succès dans l'industrie, mais leur fragilité et leur difficulté de maintenance ont limité leur diffusion.

**Le renouveau** de l'IA dans les années 2010 est principalement dû aux progrès dans le domaine de l'apprentissage profond, qui repose sur l'utilisation de réseaux de neurones artificiels à plusieurs couches. Ces réseaux sont capables d'apprendre de grandes quantités de données et d'accomplir des tâches complexes, telles que la reconnaissance d'images, la traduction automatique ou la génération de texte.

**En résumé**, l'histoire de l'IA est marquée par des avancées spectaculaires, mais aussi par des périodes de doute et de remise en question. Les défis restent nombreux, mais les perspectives d'avenir sont prometteuses.

## LES DATES CLES DE L'INTELLIGENCE ARTIFICIELLE

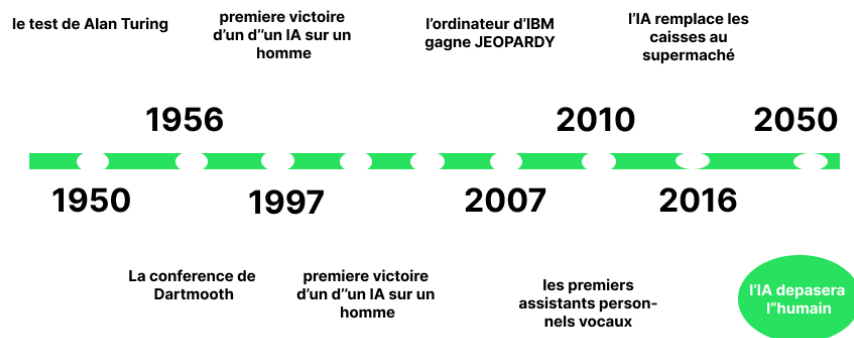


Figure 2: dates d'évolution de l'IA

### 1.3 Types d'intelligence artificielle

L'IA peut être divisée en plusieurs sous-types, en fonction de ses capacités et de ses applications

#### 1.3.1 IA faible (IA étroite)

L'IA faible, ou étroite, est conçue pour effectuer une tâche spécifique, souvent mieux que les humains, mais elle est limitée à un domaine très précis. (Bostrom N. (., 2014) Elle n'a pas la capacité de généraliser ou de "penser" en dehors de sa programmation initiale.

Exemple :

- ❖ Les assistants vocaux comme Siri ou Google Assistant.

- ❖ Les systèmes de recommandation de Netflix ou Amazon.

### 1.3.2 IA forte (IA générale)

L'IA forte, ou générale, est l'objectif ultime de nombreux chercheurs dans le domaine de l'intelligence artificielle. Elle vise à créer des machines capables de comprendre, apprendre et appliquer leur intelligence dans n'importe quel domaine, comme le ferait un être humain. Cependant, ce type d'IA reste largement théorique à ce jour.

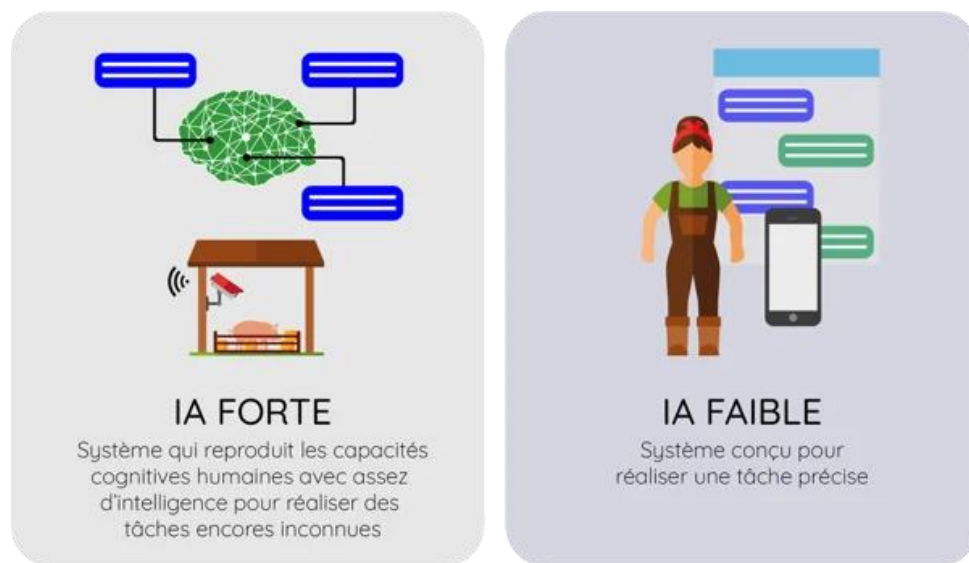


Figure 3: comparaison IA forte IA faible

### 1.3.3 Apprentissage profond (Deep Learning)

Le deep learning est une méthode d'apprentissage automatique qui repose sur des réseaux de neurones profonds. (Ian Goodfellow, 2016) Il permet aux machines de reconnaître des motifs complexes et d'apprendre à partir de grandes quantités de données. Le deep learning a permis des avancées significatives dans des domaines tels que la vision par ordinateur, la reconnaissance vocale, et le traitement du langage naturel.

Exemple :

- ❖ Les voitures autonomes utilisant des réseaux de neurones pour analyser leur environnement.
- ❖ La reconnaissance faciale dans les téléphones intelligents.

## 1.4 Les approches en intelligence artificielle

### 1.4.1 Approche symbolique

L'approche symbolique de l'intelligence artificielle repose sur des règles formelles et une représentation explicite des connaissances. Elle est utilisée dans les systèmes experts et dans la programmation logique.

Exemple :

- ❖ Systèmes de diagnostic médical basés sur des règles.

### 1.4.2 Approche connexionniste

L'approche connexionniste, largement utilisée dans le Deep Learning, se base sur la modélisation des processus cérébraux sous forme de réseaux de neurones artificiels. Ces réseaux sont formés par des algorithmes d'apprentissage qui ajustent les connexions entre les neurones pour accomplir une tâche donnée.

Exemple :

- ❖ Réseaux de neurones profonds utilisés pour la reconnaissance d'image ou la traduction automatique.

## 1.5 Applications de l'intelligence artificielle

L'IA est déjà omniprésente dans notre quotidien, et ses applications se multiplient :

- ❖ Santé : Aide au diagnostic médical, analyse d'images radiologiques, découverte de médicaments.
- ❖ Transport : Véhicules autonomes, optimisation des itinéraires de vol et de conduite.
- ❖ Commerce : Recommandations de produits, analyse des comportements d'achat.
- ❖ Finance : Trading automatisé, détection de fraudes. (Stuart Russell, 2010)





Figure 4: robot médical

## 1.6 Éthique et enjeux de l'intelligence artificielle

L'IA soulève également de nombreuses questions éthiques. Les risques potentiels incluent la perte d'emplois due à l'automatisation, les biais dans les algorithmes, et la menace d'une IA non contrôlée. (Bostrom N. , 2014) L'émergence de ces technologies impose une réflexion sérieuse sur la manière de les réguler et de les utiliser de manière éthique et responsable.

Jusqu'à présent, nous nous sommes concentrés sur les aspects techniques et sociétaux de l'IA. Il est également essentiel de considérer les implications philosophiques et existentielles de cette technologie :

- ❖ **La nature de l'intelligence :** Qu'est-ce que l'intelligence ? L'IA peut-elle véritablement "penser" ou est-ce une simulation ? Cette question soulève des débats sur la conscience, la nature de l'esprit et la place de l'homme dans l'univers.
- ❖ **Les relations humaines :** L'IA risque-t-elle de modifier nos relations interpersonnelles ? Comment allons-nous interagir avec des machines de plus en plus autonomes et intelligentes ?
- ❖ **Le sens de la vie :** Si les machines peuvent accomplir de plus en plus de tâches, quel sera le rôle de l'homme dans la société ? Quel sens donner à notre existence dans un monde où les machines pourraient nous surpasser dans de nombreux domaines ?

**Des questions clés à explorer**

- ❖ **Responsabilité** : Qui est responsable des actions d'une IA ? Le développeur, l'utilisateur, ou l'IA elle-même ?
- ❖ **Droit** : Comment adapter le cadre juridique existant à l'émergence de l'IA ? Faut-il accorder des droits aux robots ?
- ❖ **Équité** : Comment garantir que les bénéfices de l'IA soient répartis équitablement et que personne ne soit laissé pour compte ?
- ❖ **Sécurité** : Comment protéger les systèmes d'IA contre les cyberattaques et les utilisations malveillantes ?

### Une approche interdisciplinaire

Pour appréhender pleinement les enjeux éthiques de l'IA, il est nécessaire de mobiliser les connaissances de différentes disciplines :

- ❖ **Philosophie** : Pour réfléchir sur les questions fondamentales liées à la nature de l'intelligence, de la conscience et de l'existence.
- ❖ **Sociologie** : Pour analyser les impacts sociaux de l'IA et les transformations qu'elle engendre.
- ❖ **Droit** : Pour élaborer un cadre juridique adapté à l'IA.
- ❖ **Éthique** : Pour définir les principes éthiques qui doivent guider le développement et l'utilisation de l'IA.
- ❖ **Sciences cognitives** : Pour mieux comprendre le fonctionnement du cerveau humain et les mécanismes de l'intelligence.

## 1.7 Conclusion

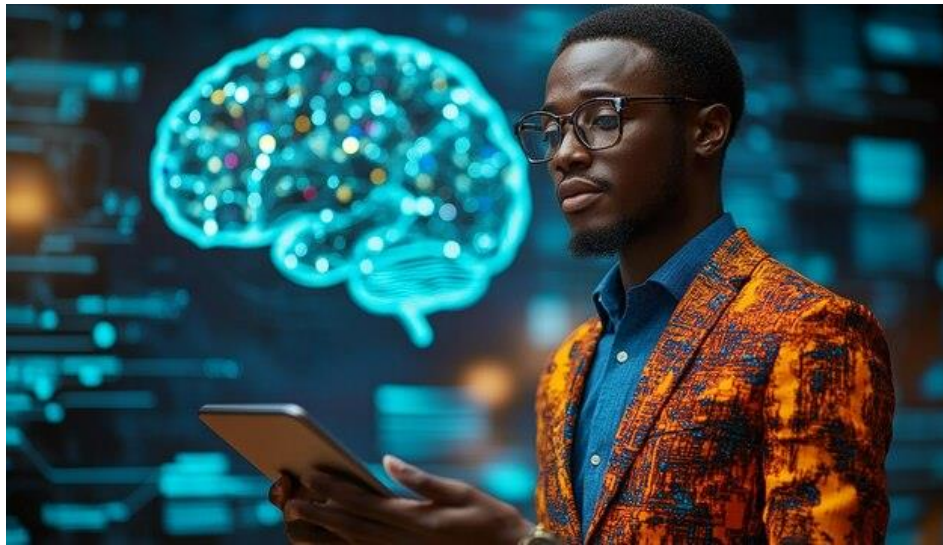
L'intelligence artificielle est un domaine complexe et en pleine expansion, avec un potentiel énorme pour transformer notre société. Bien que certaines applications telles que l'IA faible soient déjà bien intégrées dans notre quotidien, le développement d'une IA forte reste un défi pour les chercheurs. L'avenir de l'IA dépendra non seulement des avancées technologiques, mais aussi de la manière dont nous choisissons de gérer ses implications éthiques.

## Chapitre 2 : Généralité sur le Machine Learning

### 2.1 Introduction au Machine Learning

Le Machine Learning, ou apprentissage automatique, est un domaine en plein essor de l'intelligence artificielle (IA) qui permet aux machines d'apprendre à partir de données, sans être explicitement programmées pour chaque tâche. (Stuart Russell, 2010) Il se base sur des algorithmes capables d'identifier des motifs, de faire des prédictions et d'améliorer leur performance au fil du temps à partir des expériences accumulées.

Le machine Learning est aujourd'hui utilisé dans une multitude de secteurs, allant des recommandations personnalisées sur les plateformes de streaming aux véhicules autonomes, en passant par les diagnostics médicaux.



*Figure 5: machine learning system<sup>2</sup>*

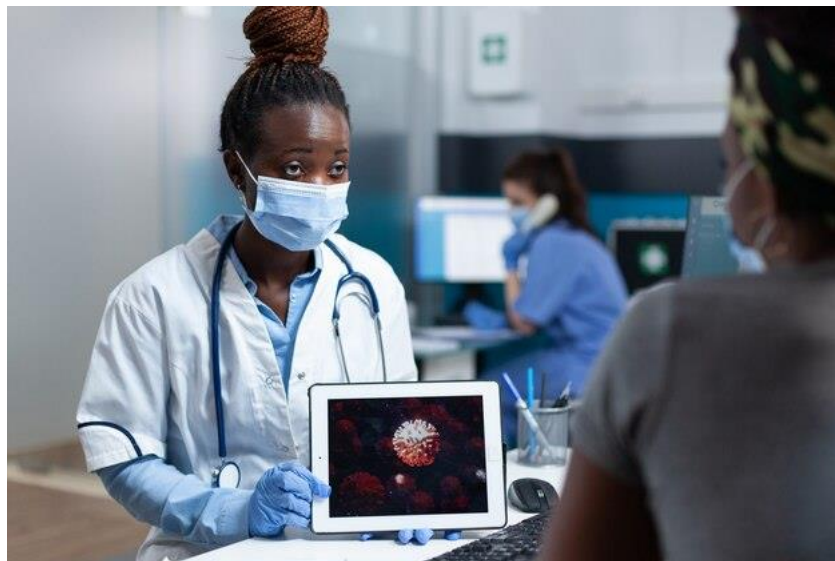
**Exemple d'un système de Machine Learning :**

Logistique et fabrication : Les robots d'apprentissage automatique peuvent être utilisés pour automatiser les tâches de manutention et d'emballage dans les entrepôts.



*Figure 6: MI dans l'industrie*

❖ L'autosurveillance de la croissance des enfants dans le domaine de la santé :



*Figure 7: consultation à l'hôpital<sup>3</sup>*

Aujourd'hui avec le développement du numérique, nous pouvons en citer encore beaucoup mais ces deux exemples suffiront pour l'instant.

## 2.2 Historique du Machine Learning

L'idée que les machines puissent "apprendre" n'est pas nouvelle. Elle trouve ses racines dans les travaux de pionniers de l'intelligence artificielle tels qu'Alan Turing et Arthur Samuel. Dans les années 1950, Turing proposa un test pour déterminer si une machine pouvait se comporter de manière intelligente. C'est Arthur Samuel, en 1959, qui utilisa pour la première fois le terme de "Machine Learning" lorsqu'il développa un programme capable de jouer aux échecs et d'améliorer son niveau de jeu avec l'expérience. (Stuart Russell, 2010).

## 2.3 Types d'apprentissage automatique

Le Machine Learning peut être catégorisé en plusieurs types d'apprentissage en fonction de la nature des données disponibles et de l'objectif poursuivi. (Bishop, 2006).

### 2.3.1 *Apprentissage supervisé.*

L'apprentissage supervisé est le type d'apprentissage le plus courant. Dans ce cadre, l'algorithme apprend à partir d'un ensemble de données étiquetées, où chaque donnée d'entrée est associée à une réponse correcte. L'objectif de l'algorithme est de découvrir une fonction qui relie les entrées aux sorties. Cela permet ensuite de prédire la réponse correcte pour de nouvelles données non observées.

#### **Exemples d'applications :**

- ❖ La reconnaissance d'images, où chaque image est associée à une étiquette (par exemple : chat, chien, etc.).
- ❖ La prédiction de valeurs boursières.

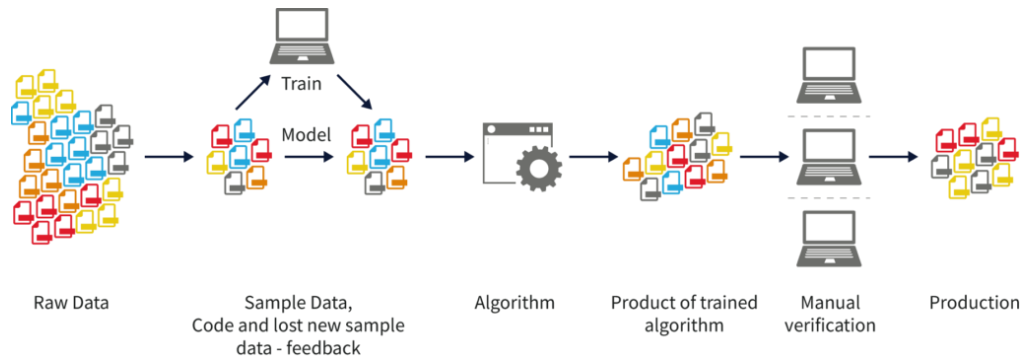


Figure 8: *apprentissage supervisé*<sup>4</sup>

### 2.3.2 *Apprentissage non supervisé :*

Contrairement à l'apprentissage supervisé, l'apprentissage non supervisé travaille avec des données non étiquetées. L'algorithme essaie de découvrir des motifs ou des structures sous-jacentes dans les données. L'apprentissage non supervisé est souvent utilisé pour le regroupement (clustering) ou la réduction de dimensionnalité.

Exemples d'applications :

- ❖ Le regroupement de clients ayant des comportements d'achat similaires.
- ❖ La détection d'anomalies dans les transactions bancaires.

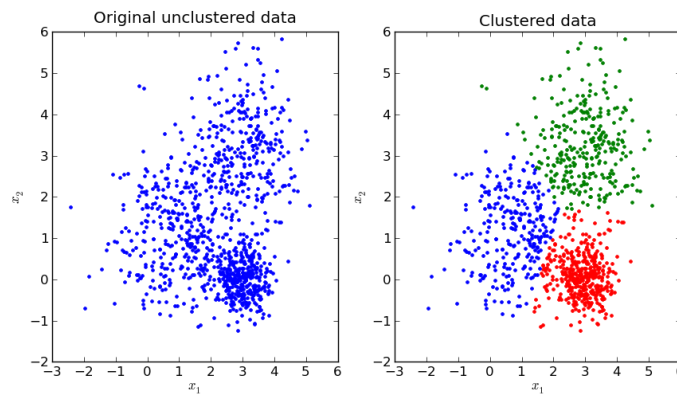


Figure 9: *apprentissage non supervisé*<sup>5</sup>

### 2.3.3 *Apprentissage par renforcement*

L'apprentissage par renforcement est une méthode où un agent apprend à interagir avec un environnement en recevant des récompenses ou des punitions en fonction de ses actions. L'agent cherche à maximiser la récompense totale au fil du temps en apprenant une politique optimale.

Exemples d'applications :

- ❖ Les jeux vidéo où une IA apprend à jouer et à s'améliorer avec le temps.
- ❖ La conduite autonome, où le véhicule apprend à naviguer de manière optimale dans un environnement complexe.

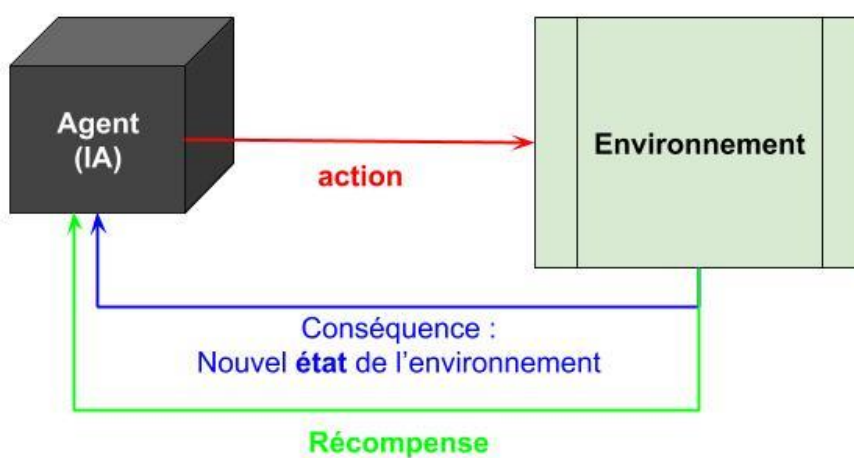


Figure 10: apprentissage par renforcement

## 2.4 Algorithme de Machine Learning

En termes d'algorithme de machine learning on peut citer trois catégories :

Les algorithmes de classification, les algorithmes de régression et les algorithmes de Deep Learning

### Réseaux de neurones ou Deep Learning

Les réseaux de neurones artificiels imitent le fonctionnement du cerveau humain. Ils sont composés de couches de neurones connectés entre eux et peuvent apprendre des représentations complexes à partir des données. Le deep learning, une sous-catégorie du machine learning, utilise des réseaux de neurones profonds pour résoudre des tâches particulièrement complexes comme la reconnaissance vocale ou la vision par ordinateur. (Stuart Russell, 2010).



## Exemple

Reconnaissance d'images dans des applications comme Google Photos ou les filtres faciaux sur Snapchat.

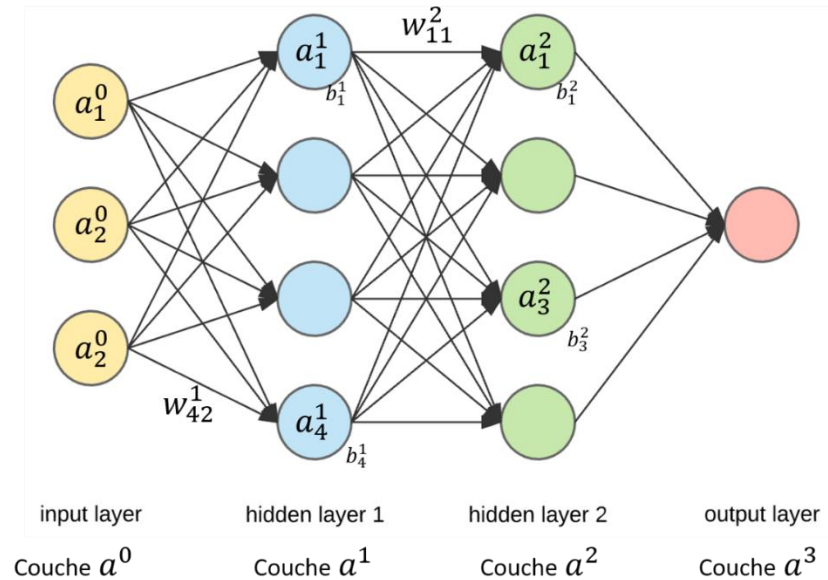


Figure 11: réseau de neurones<sup>6</sup>

## 2.5 Applications du Machine Learning

Le machine Learning est une technologie polyvalente qui trouve des applications dans de nombreux secteurs :

Santé : Diagnostic médical automatisé, identification des anomalies dans les radiographies, prédiction comme dans le cadre de ce mémoire.

- ❖ Finance : Détection de fraudes, prévision des tendances de marché.
- ❖ Commerce en ligne : Recommandations personnalisées, segmentation de la clientèle.
- ❖ Transport : Véhicules autonomes, optimisation des itinéraires.
- ❖ Industrie : Maintenance prédictive, automatisation des processus.





*Figure 12 application de Machine learning*

## **2.6 Conclusion**

Le machine learning, en tant que sous-domaine de l'intelligence artificielle, a révolutionné de nombreuses industries en permettant aux machines d'apprendre à partir de données et d'améliorer leur performance avec le temps. Que ce soit à travers l'apprentissage supervisé, non supervisé ou par renforcement, il existe une large variété d'approches et d'algorithmes capables de répondre à des problématiques spécifiques. Cette technologie est vouée à jouer un rôle toujours plus central dans l'avenir de l'automatisation et de l'intelligence artificielle.

## Chapitre 3 : Collecte de données, les maladies répertoriés et certains de leurs symptômes

### 3.1 Introduction à la collecte de données

La collecte de données est une étape cruciale dans toute étude scientifique. Elle consiste à rassembler des informations pertinentes qui permettront d'analyser et de comprendre les phénomènes étudiés. Dans le cadre de cette étude, les données collectées concernent des maladies spécifiques et leurs symptômes associés, avec pour objectif de mieux comprendre la nature de ces pathologies, leurs facteurs de risque, et leur répartition dans la population étudiée.



Figure 13: la collecte de données

La qualité et la précision des données sont essentielles, car elles influencent directement les résultats de l'étude. Ainsi, une collecte rigoureuse permet d'assurer des conclusions fiables et de poser des bases solides pour d'éventuelles interventions ou recommandations médicales.

### 3.2 Méthodologie de la collecte de données

La méthodologie de collecte des données utilisée dans cette étude repose sur plusieurs étapes clés, visant à garantir l'exhaustivité et la fiabilité des informations recueillies. Les principales étapes comprennent :

#### 3.2.1 Sources de données

La collecte de données est une étape cruciale dans tout projet de recherche, notamment dans le domaine de la santé. Deux méthodes principales se distinguent : les enquêtes et l'Internet des

objets (IoT). Chacune présente des avantages et des inconvénients spécifiques, et le choix de la méthode dépendra des objectifs de la recherche, des ressources disponibles et du contexte.

#### ❖ Collecte de données par IoT dans le domaine de la santé

Lors de mon cursus, arrivé en Master 1 j'ai ambitionné de créer par moi-même un mini robot qui au-delà d'être utile au gens par sa surveillance par reconnaissance faciale permet la prise de constantes chez un malade. En effet lors d'un projet de téléconsultation je pu développer en intégralité ce bot qui est équipé de suffisamment de capteurs permettant de recueillir entre autres chez un patient la température du corps, le rythme cardiaque... et ces informations sont transmises via une plateforme web afin de le stocker pour permettre l'interaction avec un médecin à distance.



Figure 14: bot de collecte par IoT

#### Avantages :

- ❖ **Données en temps réel:** Les capteurs IoT permettent de collecter des données physiologiques (rythme cardiaque, pression artérielle, etc.) en continu et en temps réel, offrant une vision plus précise de l'état de santé des individus.
- ❖ **Objectivité :** Les données collectées par les capteurs sont généralement plus objectives que les données auto-rapportées dans les enquêtes.
- ❖ **Détection précoce :** Les données IoT peuvent permettre de détecter des anomalies ou des tendances précoces, facilitant ainsi la prise en charge médicale.

#### Inconvénients :

- ❖ **Coût** : La mise en place d'un système IoT peut être coûteuse, notamment en termes d'achat des capteurs et de développement de l'infrastructure.
- ❖ **Complexité technique** : La gestion et l'analyse des données IoT peuvent nécessiter des compétences techniques spécifiques.
- ❖ **Collecte de données par enquête**

Les données utilisées pour ce projet proviennent principalement d'une étude de l'université de Columbia réalisée à l'hôpital presbytérien de New York en 2004 : [Disease \(columbia.edu\)](http://Disease(columbia.edu))

À ce là j'ai ajouté quelques enquêtes publiées dans kaggle. Elles illustrent l'associations maladie-symptôme générées par une méthode automatisée basée sur les informations contenues dans les résumés de sortie textuels des patients du New York Presbytérien Hospital admis en 2004. La première colonne indique la maladie, la seconde le nombre de résumés de sortie contenant une mention positive et actuelle de la maladie, et le symptôme associé. Les associations pour les 150 maladies les plus fréquentes basées sur ces notes ont été calculées et les symptômes sont classés en fonction de la force de l'association. La méthode utilise le système de traitement du langage naturel MEDLEE pour obtenir les codes UMLS des maladies et des symptômes à partir des notes ; ensuite, des méthodes statistiques basées sur les fréquences et les cooccurrences ont été utilisées pour obtenir les associations. Une description plus détaillée de la méthode automatisée est disponible dans Wang X, Chused A, Elhadad N, Friedman C, Markatou M. Automated knowledge acquisition from clinical reports. AMIA Annu Symp Proc. 2008. p. 783-7. PMCID : PMC2656103.<sup>7</sup>

---

<sup>7</sup> [Disease \(columbia.edu\)](http://Disease(columbia.edu))



Figure 15: collecte par enquête

### 3.2.2 Validation des données

Une fois les données collectées, elles doivent être validées afin de s'assurer de leur fiabilité. Ce processus inclut :

- Vérification des doublons : Les données sont vérifiées afin d'éviter les répétitions ou les informations incohérentes.
- Vérification de la qualité : Les informations sont comparées à des standards établis pour s'assurer de leur exactitude.

Dans notre cas les données sont déjà validé et publié dans l'Oms et dans quelques sites internet.

## 3.3 Prétraitement des données

Avant l'analyse, les données doivent être prétraitées pour garantir leur bon format et leur qualité. Le prétraitement des données inclut plusieurs étapes :

### 3.3.1 Nettoyage des données

Le nettoyage des données consiste à éliminer ou corriger les erreurs et incohérences présentes dans les ensembles de données. Cela inclut :

- ❖ **Suppression des valeurs manquantes** : Les enregistrements incomplets ou erronés sont soit complétés à partir de sources fiables, soit supprimés s'ils ne peuvent pas être corrigés.
- ❖ **Uniformisation des données** : Les données collectées sous différents formats (ex. : dates, unités de mesure) sont harmonisées pour permettre leur analyse.

### 3.3.2 Transformation des données

Après le nettoyage, certaines données peuvent nécessiter une transformation avant d'être analysées :

- ❖ **Normalisation** : Les données quantitatives sont souvent normalisées pour garantir qu'elles sont comparables entre elles.
- ❖ **Encodage des variables** : Les variables catégorielles, comme les types de symptômes ou de maladies, sont encodées sous forme numérique pour être utilisables par les algorithmes d'analyse.



Figure 16: data processing

## 3.4 Sélection des maladies et symptômes

### 3.4.1 Critères de sélection

Les maladies étudiées dans le cadre de cette recherche ont été sélectionnées en fonction de plusieurs critères :

- ❖ Prévalence : Seules les maladies les plus fréquentes dans la population cible ont été sélectionnées.
- ❖ Impact sur la santé : Les maladies sélectionnées sont celles qui présentent un risque significatif pour la santé publique.
- ❖ Disponibilité des données : Seules les maladies pour lesquelles des données fiables et complètes sont disponibles ont été incluses dans l'étude.

### 3.4.2 Liste des symptômes étudiés

Les symptômes étudiés dans cette recherche couvrent une large gamme de signes cliniques, allant de troubles légers à des affections plus graves. Ces symptômes ont été sélectionnés en raison de leur prévalence et de leur lien avec les maladies analysées dans cette étude. Voici une sélection des principaux symptômes étudiés :

Symptôme	Description
back_pain	Mal de dos
constipation	Constipation
abdominal_pain	Douleur abdominale
diarrhoea	Diarrhée
mild_fever	Fièvre légère
yellow_urine	Urine jaune
yellowing_of_eyes	Jaunissement des yeux (ictère)
acute_liver_failure	Insuffisance hépatique aiguë
fluid_overload	S surcharge hydrique
swelling_of_stomach	Gonflement de l'estomac
swollen_lymph_nodes	Ganglions lymphatiques enflés
malaise	Malaise général
blurred_and_distorted_vision	Vision floue et déformée
phlegm	Flegme
throat_irritation	Irritation de la gorge
redness_of_eyes	Rougeur des yeux
sinus_pressure	Pression dans les sinus
runny_nose	Écoulement nasal
congestion	Congestion nasale
chest_pain	Douleur thoracique
weakness_in_limbs	Faiblesse dans les membres



<b>fast_heart_rate</b>	Rythme cardiaque rapide
<b>pain_during_bowel_movements</b>	Douleur pendant la selle
<b>pain_in_anal_region</b>	Douleur dans la région anale
<b>bloody_stool</b>	Selles sanglantes
<b>irritation_in_anus</b>	Irritation anale
<b>neck_pain</b>	Douleur au cou
<b>dizziness</b>	Étourdissements
<b>cramps</b>	Crampes
<b>bruising</b>	Ecchymoses (bleus)
<b>obesity</b>	Obésité
<b>swollen_legs</b>	Jambes enflées
<b>swollen_blood_vessels</b>	Vaisseaux sanguins dilatés
<b>puffy_face_and_eyes</b>	Visage et yeux bouffis
<b>enlarged_thyroid</b>	Thyroïde hypertrophiée
<b>brittle_nails</b>	Ongles cassants
<b>swollen_extremities</b>	Extrémités enflées
<b>excessive_hunger</b>	Faim excessive
<b>extra_marital_contacts</b>	Relations sexuelles extraconjugales
<b>drying_and_tingling_lips</b>	Lèvres sèches et picotantes
<b>slurred_speech</b>	Parole trouble
<b>knee_pain</b>	Douleur au genou
<b>hip_joint_pain</b>	Douleur à la hanche
<b>muscle_weakness</b>	Faiblesse musculaire
<b>stiff_neck</b>	Torticolis
<b>swelling_joints</b>	Articulations enflées
<b>movement_stiffness</b>	Raideur des mouvements
<b>spinning_movements</b>	Sensations de rotation
<b>loss_of_balance</b>	Perte d'équilibre
<b>unsteadiness</b>	Instabilité
<b>weakness_of_one_body_side</b>	Faiblesse d'un côté du corps
<b>loss_of_smell</b>	Perte d'odorat
<b>bladder_discomfort</b>	Gêne de la vessie
<b>foul_smell_of_urine</b>	Urine malodorante
<b>continuous_feel_of_urine</b>	Envie permanente d'uriner
<b>passage_of_gases</b>	Émission de gaz
<b>internal_itching</b>	Démangeaisons internes
<b>toxic_look_(typhos)</b>	Facies typhique (fièvre typhoïde)
<b>depression</b>	Dépression
<b>irritability</b>	Irritabilité



<b>muscle_pain</b>	Douleurs musculaires
<b>altered_sensorium</b>	Confusion mentale
<b>red_spots_over_body</b>	Taches rouges sur le corps
<b>belly_pain</b>	Mal au ventre
<b>abnormal_menstruation</b>	Règles anormales
<b>dischromic_patches</b>	Taches cutanées décolorées
<b>watering_from_eyes</b>	Larmolement
<b>increased_appetite</b>	Augmentation de l'appétit
<b>polyuria</b>	Mictions fréquentes
<b>family_history</b>	Antécédents familiaux
<b>muroid_sputum</b>	Crachats mucoïdes
<b>rusty_sputum</b>	Crachats rouillés
<b>lack_of_concentration</b>	Troubles de la concentration
<b>visual_disturbances</b>	Troubles visuels
<b>receiving_blood_transfusion</b>	Transfusion sanguine
<b>receiving_unsterile_injections</b>	Injections non stériles
<b>coma</b>	Coma

Tableau 1: liste des symptômes associés

### 3.1.1 Sélection des maladies

Les maladies sélectionnées pour cette étude ont été choisies en fonction de plusieurs critères, tels que leur prévalence, leur impact sur la santé publique et la disponibilité des données. Elles couvrent un large éventail de pathologies, allant des maladies infectieuses aux troubles métaboliques. Voici une présentation détaillée des principales maladies étudiées :

<b>Maladie (Anglais)</b>	<b>Maladie (Français)</b>	<b>Description</b>
<b>Fungal infection</b>	Infection mycosique	Infection causée par un champignon.
<b>Allergy</b>	Allergie	Réaction excessive du système immunitaire à une substance généralement inoffensive.
<b>GERD</b>	Reflux gastro-œsophagien (RGO)	Remontée acide de l'estomac vers l'œsophage.
<b>Chronic cholestasis</b>	Cholestase chronique	Affection empêchant l'écoulement normal de la bile du foie.
<b>Drug Reaction</b>	Effet indésirable d'un médicament	Réaction indésirable à un médicament.
<b>Peptic ulcer disease</b>	Ulcère gastro-duodéal	Plaies se développant sur la paroi de l'estomac ou du duodénum.

<b>AIDS</b>	Sida (Syndrome d'immunodéficience acquise)	Stade le plus avancé de l'infection par le VIH.
<b>Diabetes</b>	Diabète	Affection chronique caractérisée par un taux élevé de sucre dans le sang.
<b>Gastroenteritis</b>	Gastro-entérite	Inflammation de l'estomac et des intestins.
<b>Bronchial Asthma</b>	Asthme bronchique	Maladie inflammatoire chronique des voies respiratoires.
<b>Hypertension</b>	Hypertension artérielle	Pression artérielle anormalement élevée.
<b>Migraine</b>	Migraine	Mal de tête intense pouvant s'accompagner de nausées, de vomissements et d'une sensibilité à la lumière et au son.
<b>Cervical spondylosis</b>	Spondylose cervicale	Arthrose touchant les os et les articulations du cou.
<b>Paralysis (brain hemorrhage)</b>	Paralysie (hémorragie cérébrale)	Perte de la fonction musculaire due à un saignement dans le cerveau.
<b>Jaundice</b>	Ictère	Coloration jaune de la peau et du blanc des yeux.
<b>Malaria</b>	Paludisme	Maladie infectieuse transmise par les moustiques.
<b>Chicken pox</b>	Varicelle	Maladie infantile contagieuse caractérisée par une éruption cutanée vésiculeuse.
<b>Dengue</b>	Dengue	Maladie virale transmise par les moustiques, pouvant être grave.
<b>Typhoid</b>	Fièvre typhoïde	Maladie infectieuse due à la bactérie Salmonella Typhi.
<b>Hepatitis A</b>	Hépatite A	Inflammation du foie due au virus de l'hépatite A.
<b>Hepatitis B</b>	Hépatite B	Inflammation du foie due au virus de l'hépatite B.
<b>Hepatitis C</b>	Hépatite C	Inflammation du foie due au virus de l'hépatite C.
<b>Hepatitis D</b>	Hépatite D	Inflammation du foie nécessitant la présence de l'hépatite B pour se développer.
<b>Hepatitis E</b>	Hépatite E	Inflammation du foie due au virus de l'hépatite E.
<b>Alcoholic hepatitis</b>	Hépatite alcoolique	Inflammation du foie causée par une consommation excessive d'alcool.

<b>Tuberculosis</b>	Tuberculose	Maladie infectieuse contagieuse touchant généralement les poumons.
<b>Common Cold</b>	Rhume	Infection virale des voies respiratoires supérieures.
<b>Pneumonia</b>	Pneumonie	Infection des poumons.
<b>Dimorphic hemorrhoids(piles)</b>	Hémorroïdes	Veines dilatées dans le rectum ou l'anus.
<b>Heart attack</b>	Crise cardiaque	Mort brutale d'une partie du muscle cardiaque due à un arrêt de la circulation sanguine.
<b>Varicose veins</b>	Varices	Veines dilatées et tortueuses visibles sous la peau.
<b>Hypothyroidism</b>	Hypothyroïdie	Fonctionnement insuffisant de la glande thyroïde.
<b>Hyperthyroidism</b>	Hyperthyroïdie	Fonctionnement excessif de la glande thyroïde.
<b>Hypoglycemia</b>	Hypoglycémie	Taux de sucre dans le sang anormalement bas.
<b>Osteoarthritis</b>	Arthrose	Forme d'arthrite touchant le cartilage des articulations.
<b>Arthritis</b>	Arthrite	Inflammation douloureuse d'une ou plusieurs articulations.
<b>(vertigo) Paroxysmal Positional Vertigo</b>	Vertige paroxystique positionnel bénin (VPPB)	Sensation brève de vertige déclenchée par certains mouvements de la tête.
<b>Acne</b>	Acné	Affection

Tableau 2: liste des maladies repertoriées

## Conclusion

Dans ce chapitre, nous avons présenté les différentes étapes de la collecte des données, ainsi que les maladies et symptômes étudiés. L'analyse des symptômes, tels que les douleurs abdominales, les écoulements nasaux ou les ganglions lymphatiques enflés, permet de comprendre la diversité des affections traitées. Les maladies sélectionnées couvrent un large spectre, allant des infections virales telles que la **fièvre typhoïde** et la **tuberculose**, aux affections métaboliques comme le **diabète** et l'**hypothyroïdie**. Cette collecte rigoureuse de données servira de base pour l'application des algorithmes de machine learning dans les chapitres suivants, afin de détecter et prédire les maladies sur la base des symptômes.

L'intégration de ces données dans un modèle informatique repose sur la précision et la diversité des informations collectées, renforçant ainsi la robustesse des prédictions. L'utilisation des

technologies d'intelligence artificielle dans l'analyse médicale nous permet de développer des solutions innovantes pour le diagnostic précoce, tout en améliorant la prise en charge des patients.

# Chapitre 4 : Conception du système de prédiction de maladie

## 4.1 Introduction

La modélisation et la conception du système de prédiction de maladies consistent à transformer les données de symptômes et maladies en informations exploitables par un algorithme de machine learning. L'objectif est de créer un modèle prédictif capable d'identifier les maladies à partir d'un ensemble de symptômes. Ce chapitre présente les différentes étapes de cette modélisation, les choix technologiques effectués, ainsi que les aspects techniques liés à la conception de ce système.

## 4.2 Étapes de la modélisation

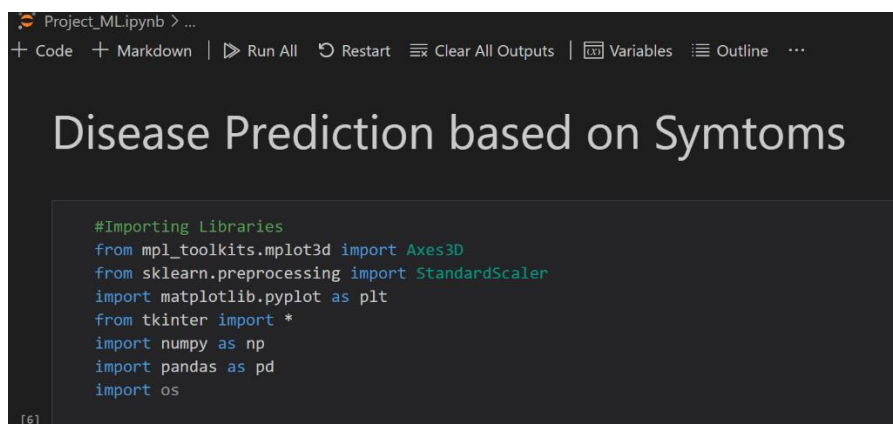
La modélisation d'un système de prédiction de maladies s'effectue en plusieurs étapes, depuis la collecte des données jusqu'à l'entraînement des algorithmes. Voici les principales étapes :

### 4.2.1 Prétraitement des données

Les données collectées doivent être nettoyées et préparées avant d'être utilisées par un modèle de machine learning. Le prétraitement inclut la gestion des données manquantes, la normalisation des valeurs, et la transformation des variables catégorielles en format numérique.

- ❖ **Technologies utilisés** : Vscode, Google Colabs , numpy, Pandas et python comme langage de programmation.

Tout d'abord nous allons afficher les données grâce à des bibliothèque de sklearn.



```
Project_ML.ipynb > ...
+ Code + Markdown | ▶ Run All ↺ Restart ≡ Clear All Outputs | 📄 Variables 📖 Outline ...

Disease Prediction based on Symtoms

#Importing Libraries
from mpl_toolkits.mplot3d import Axes3D
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
from tkinter import *
import numpy as np
import pandas as pd
import os

[6]
```

Figure 17: bibliotheque

Ensuite nous allons mettre les symptômes et les maladies respectueusement dans une liste1 et liste2.

```
#List of the symptoms is listed here in list 11.

l1=['back_pain','constipation','abdominal_pain','diarrhoea','mild_fever','yellow_urine',
'yellowing_of_eyes','acute_liver_failure','fluid_overload','swelling_of_stomach',
'swelled_lymph_nodes','malaise','blurred_and_distorted_vision','phlegm','throat_irritation',
'redness_of_eyes','sinus_pressure','runny_nose','congestion','chest_pain','weakness_in_limbs',
'fast_heart_rate','pain_during_bowel_movements','pain_in_anal_region','bloody_stool',
'irritation_in_anus','neck_pain','dizziness','cramps','bruising','obesity','swollen_legs',
'swollen_blood_vessels','puffy_face_and_eyes','enlarged_thyroid','brittle_nails',
'swollen_extremeties','excessive_hunger','extra_marital_contacts','drying_and_tingling_lips',
'slurred_speech','knee_pain','hip_joint_pain','muscle_weakness','stiff_neck','swelling_joints',
'movement_stiffness','spinning_movements','loss_of_balance','unsteadiness',
'weakness_of_one_body_side','loss_of_smell','bladder_discomfort','foul_smell_of_urine',
'continuous_feel_of_urine','passage_of_gases','internal_itching','toxic_look_(typhos)',
'depression','irritability','muscle_pain','altered_sensorium','red_spots_over_body','belly_pain',
'abnormal_menstruation','dischromic_patches','watering_from_eyes','increased_appetite','polyuria',
'rusty_sputum','lack_of_concentration','visual_disturbances','receiving_blood_transfusion',
'receiving_unsterile_injections','coma','stomach_bleeding','distention_of_abdomen',
'history_of_alcohol_consumption','fluid_overload','blood_in_sputum','prominent_veins_on_calf',
'palpitations','painful_walking','pus_filled_pimples','blackheads','scurrying','skin_peeling',
'silver_like_dusting','small_dents_in_nails','inflammatory_nails','blister','red_sore_around_nose',
'yellow_crust_ooze']

len(l1)

95
```

Figure 18: symptômes

```
#List of Diseases is listed in list disease.

maladie=['Fungal infection', 'Allergy', 'GERD', 'Chronic cholestasis',
'Drug Reaction', 'Peptic ulcer disease', 'AIDS', 'Diabetes ',
'Gastroenteritis', 'Bronchial Asthma', 'Hypertension ', 'Migraine',
'Cervical spondylosis', 'Paralysis (brain hemorrhage)', 'Jaundice',
'Malaria', 'Chicken pox', 'Dengue', 'Typhoid', 'hepatitis A',
'Hepatitis B', 'Hepatitis C', 'Hepatitis D', 'Hepatitis E',
'Alcoholic hepatitis', 'Tuberculosis', 'Common Cold', 'Pneumonia',
'Dimorphic hemmorhoids(piles)', 'Heart attack', 'Varicose veins',
'Hypothyroidism', 'Hyperthyroidism', 'Hypoglycemia',
'Osteoarthritis', 'Arthritis',
'(vertigo) Parosymal Positional Vertigo', 'Acne',
'Urinary tract infection', 'Psoriasis', 'Impetigo']

# disease = [df['prognosis'].unique()]
print(len(maladie))

41
```

Figure 19: maladie

Nous pouvons donc voir que nous avons au total **95** symptômes pour **41** maladies à prédire. Après cette étape il faut maintenant charger les données dans un formats adéquat c'est ce que j'ai fait au préalable et j'ai choisi donc le format csv.

```
#Reading the training .csv file
df=pd.read_csv("Dataset/training.csv")
DF= pd.read_csv('Dataset/training.csv', index_col='prognosis')
#Replace the values in the imported file by pandas by the inbuilt function replace in pandas.

df.replace({'prognosis':{'Fungal infection':0,'Allergy':1,'GERD':2,'Chronic cholestasis':3,'Drug Reaction':4,
'Peptic ulcer disease':5,'AIDS':6,'Diabetes ':7,'Gastroenteritis':8,'Bronchial Asthma':9,'Hypertension ':10,
'Migraine':11,'Cervical spondylosis':12,
'Paralysis (brain hemorrhage)':13,'Jaundice':14,'Malaria':15,'Chicken pox':16,'Dengue':17,'Typhoid':18,'hepatitis A':19,
'Hepatitis B':20,'Hepatitis C':21,'Hepatitis D':22,'Hepatitis E':23,'Alcoholic hepatitis':24,'Tuberculosis':25,
'Common Cold':26,'Pneumonia':27,'Dimorphic hemmorhoids(piles)':28,'Heart attack':29,'Varicose veins':30,'Hypothyroidism':31,
'Hyperthyroidism':32,'Hypoglycemia':33,'Osteoarthritis':34,'Arthritis':35,
'(vertigo) Paroymsal  Positional Vertigo':36,'Acne':37,'Urinary tract infection':38,'Psoriasis':39,
'Impetigo':40}},inplace=True)
#df.head()
DF.head(100)
```

Figure 20: pretraitement

A présent nous avons chargé les données et on s'est assuré que les de remplacer dans la forme de base les index du fichier par la liste des maladies que j'ai répertorié. Et le résultat à la sortie fût :

	itching	skin_rash	nodal_skin_eruptions	continuous_sneezing	shivering	chills	joint_pain	stomach_pain	acidity	ulcers_on_t
prognosis										
Fungal infection	1	1	1	0	0	0	0	0	0	0
Fungal infection	0	1	1	0	0	0	0	0	0	0
Fungal infection	1	0	1	0	0	0	0	0	0	0
Fungal infection	1	1	0	0	0	0	0	0	0	0
Fungal infection	1	1	1	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...
Bronchial Asthma	0	0	0	0	0	0	0	0	0	0
Bronchial Asthma	0	0	0	0	0	0	0	0	0	0
Bronchial Asthma	0	0	0	0	0	0	0	0	0	0
Bronchial Asthma	0	0	0	0	0	0	0	0	0	0
Bronchial Asthma	0	0	0	0	0	0	0	0	0	0

Figure 21: visualisation

A présent que cette étape est passé maintenant on crée une fonction qui me permettra d'afficher les données sous format graphique afin de permettre une interprétation plus humaine sur la relation entre maladie et symptômes.

Il donne :

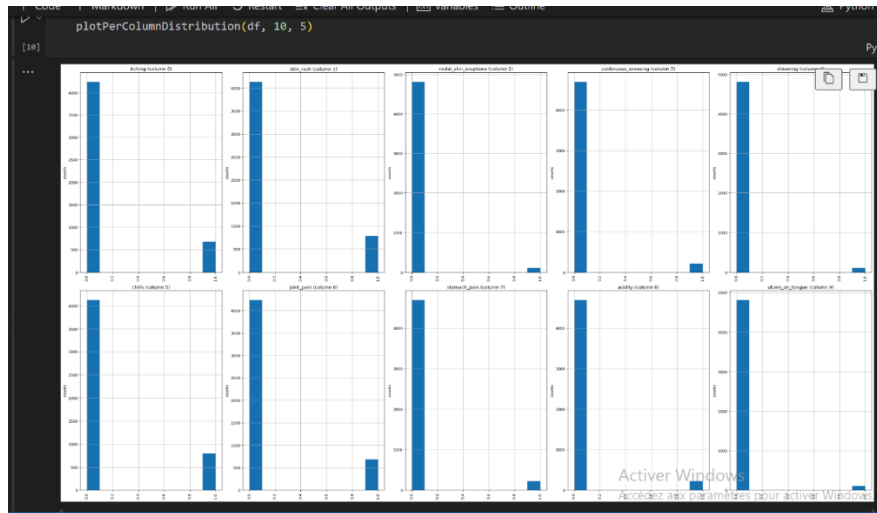


Figure 22: linkage des symptômes

Et aussi, une fonction qui permet de tracer les matrices de dispersions afin de d'analyser l'importance d'un symptôme par rapport à une maladie et d'avoir ainsi le coefficient de dispersion qui n'aura que 3 valeurs d'interprétation :

- ❖ Une valeur positive : signifie que pour une maladie donnée le symptôme en question est obligatoire
- ❖ Une valeur nul : signifie que le symptôme peut être présent ou pas
- ❖ Et une valeur négative : pour dire que le symptôme n'est pas lié à la maladie.

Et le résultat fût :

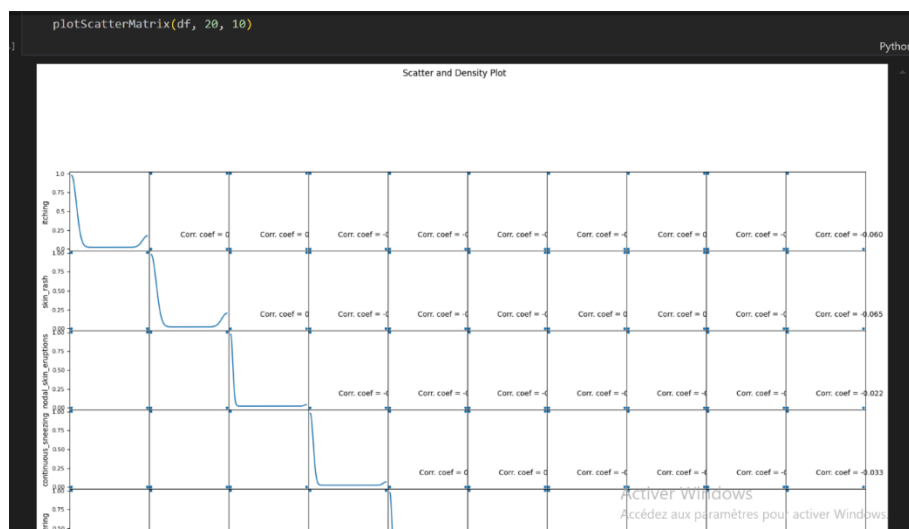


Figure 23: corrélation des symptômes



Dans la modélisation des systèmes de prédiction de maladies, l'un des outils analytiques essentiels est le coefficient de corrélation, qui mesure la force et la direction de la relation linéaire entre deux variables. Pour prédire une maladie à partir de symptômes, l'analyse de la corrélation est utile pour comprendre l'influence des différents symptômes sur l'apparition de maladies spécifiques.

Le **coefficient de corrélation** est une mesure statistique qui quantifie la relation linéaire entre deux variables. Sa valeur varie de -1 à +1 :

- ❖ **+1** indique une corrélation parfaitement positive : lorsque la valeur d'une variable augmente, l'autre augmente également de manière proportionnelle.
- ❖ **0** indique aucune corrélation : il n'y a pas de relation linéaire entre les variables.
- ❖ **-1** indique une corrélation parfaitement négative : lorsque la valeur d'une variable augmente, l'autre diminue proportionnellement.

Dans le cadre d'un système de prédiction de maladies, les variables que nous analysons sont les symptômes et les maladies. Par exemple, nous pouvons calculer la corrélation entre l'apparition de symptômes comme la **douleur abdominale** et des maladies telles que la **gastrite** ou l'**ulcère gastro-duodéal**.

Dans un ensemble de données sur les symptômes et les maladies, le coefficient de corrélation peut être utilisé pour identifier les symptômes qui sont fortement associés à certaines maladies. Cela peut être très utile pour :

- ❖ **Sélection des caractéristiques** : Identifier les symptômes qui ont un impact significatif sur la prédiction d'une maladie donnée.
- ❖ **Simplification du modèle** : En éliminant les symptômes qui n'ont pas de corrélation significative avec les maladies (corrélation proche de 0), on peut réduire la dimensionnalité du problème et ainsi améliorer l'efficacité et la précision du modèle prédictif.

Par exemple, une corrélation élevée entre des symptômes comme la **fièvre** et la **jaunisse** peut indiquer une association avec des maladies comme l'**hépatite** ou la **fièvre typhoïde**.

Il est également important de visualiser ces corrélations pour comprendre plus facilement la relation entre les variables. Une matrice de dispersion combinée avec des graphiques de densité, comme dans la fonction (`plotScatterMatrix`), permet de visualiser les relations entre plusieurs

symptômes en même temps. Chaque sous-graphique montre une paire de symptômes, tandis que la densité diagonale montre la distribution de chaque symptôme. De plus, l'annotation avec les coefficients de corrélation facilite l'interprétation des relations entre les symptômes.

Maintenant notre trun est prêt et il faut faire le meme travail pour les données tests aussi si les deux sont séparé en amont avant toute traitement et c'est notre cas ici.

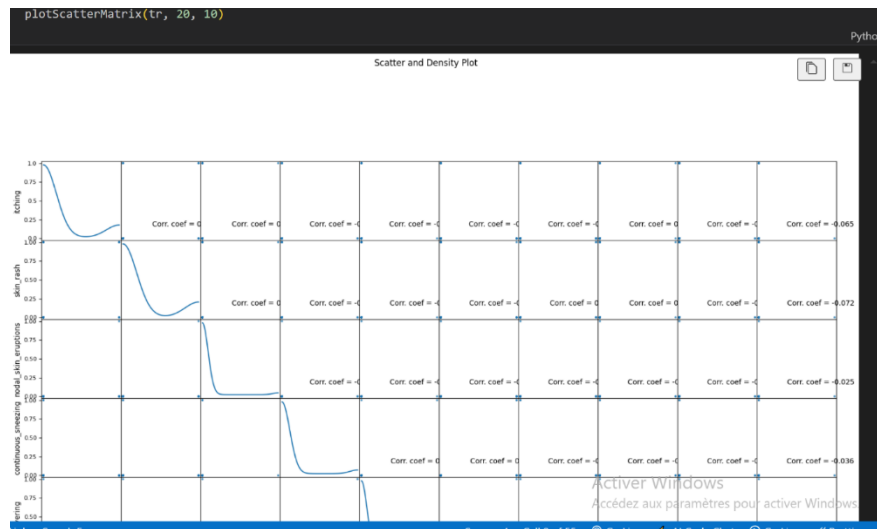


Figure 24: correlation des donnes test

On constate que les diagonales sont différentes de celles des données Trun.

### 4.3 Mise en place de la prédiction avec les algorithmes de prédiction

La mise en œuvre d'un système de prédiction de maladies repose sur l'utilisation d'algorithmes de machine learning capables d'exploiter les données cliniques. Ces algorithmes apprennent des relations entre les symptômes et les maladies à partir de données historiques, pour ensuite effectuer des prédictions sur de nouveaux cas. Plusieurs algorithmes peuvent être utilisés dans ce cadre, chacun ayant ses avantages et ses inconvénients en fonction des caractéristiques des données disponibles.

Dans le cadre de notre système de prédiction des maladies à partir de symptômes, nous avons utilisé quatre algorithmes de machine learning populaires : **l'algorithme des arbres de décision**, **les forêts aléatoires**, **K-Nearest Neighbors (KNN)** et **Naive Bayes**. Chacun de ces algorithmes a été choisi en raison de sa capacité à traiter efficacement des données médicales, où les relations entre symptômes et maladies peuvent être complexes et non linéaires.

### 4.3.1 Algorithme des arbres de décision

L'algorithme des arbres de décision est une méthode d'apprentissage supervisé qui se base sur la division réursive des données selon les valeurs des attributs. Dans le contexte de la prédiction des maladies, un arbre de décision évalue les symptômes présentés par un patient et, à chaque nœud de l'arbre, pose une question (comme "Le patient a-t-il de la fièvre ?") pour classer le patient dans une catégorie de maladie. (Quinlan, 1986)

- ❖ Avantages : Cet algorithme est facile à interpréter, car il produit des modèles visuellement compréhensibles. Cela est particulièrement utile dans le secteur médical, où il est essentiel que les médecins puissent comprendre comment le modèle prend ses décisions.
- ❖ Inconvénients : L'algorithme des arbres de décision peut être sujet à l'overfitting (surapprentissage) lorsqu'il est appliqué à de grands ensembles de données, surtout s'il n'y a pas d'élagage (pruning) pour limiter la profondeur de l'arbre.

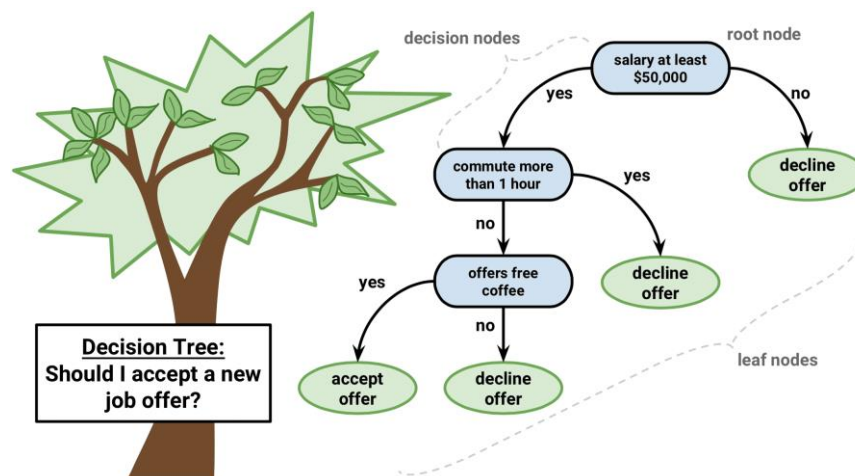


Figure 25: random forest

#### Utilisation :

ceci appliqué sur nos données nous donne une prédiction avec pour **97%** avec pour chaque client qui teste le système la possibilité de lui tracer un graphique de dispersion sur chaque modèle montrant sur quoi la précision se base t'elle avec le coefficient de dispersion.

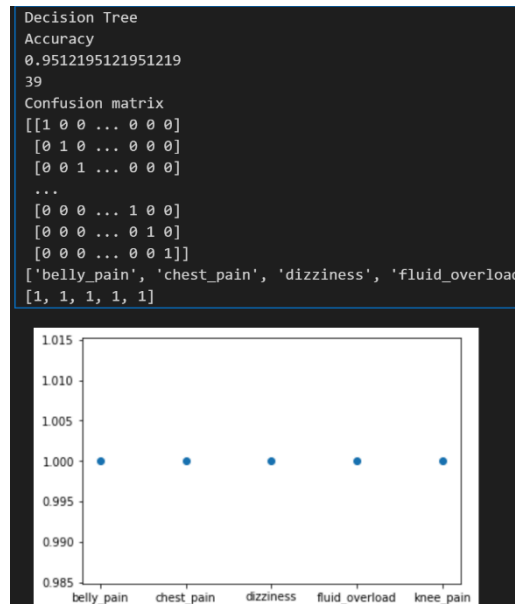


Figure 26: métrique

#### 4.3.2 Algorithme Random Forest

Les forêts aléatoires (Random Forest) sont une extension des arbres de décision, où plusieurs arbres de décision sont construits sur des sous-échantillons de données et leurs prédictions sont combinées pour obtenir un résultat final plus robuste. En combinant les résultats de plusieurs arbres, cet algorithme réduit les risques d'overfitting. (Breiman, 2001)

- ❖ Avantages : Random Forest est plus robuste que l'algorithme des arbres de décision seul. Il fournit de bonnes performances, même avec des ensembles de données complexes et bruités. De plus, il peut gérer un grand nombre de symptômes avec des interactions complexes.
- ❖ Inconvénients : Bien que Random Forest offre de bonnes performances, le modèle final est beaucoup plus difficile à interpréter que les arbres de décision individuels. Il est aussi plus lent à entraîner.

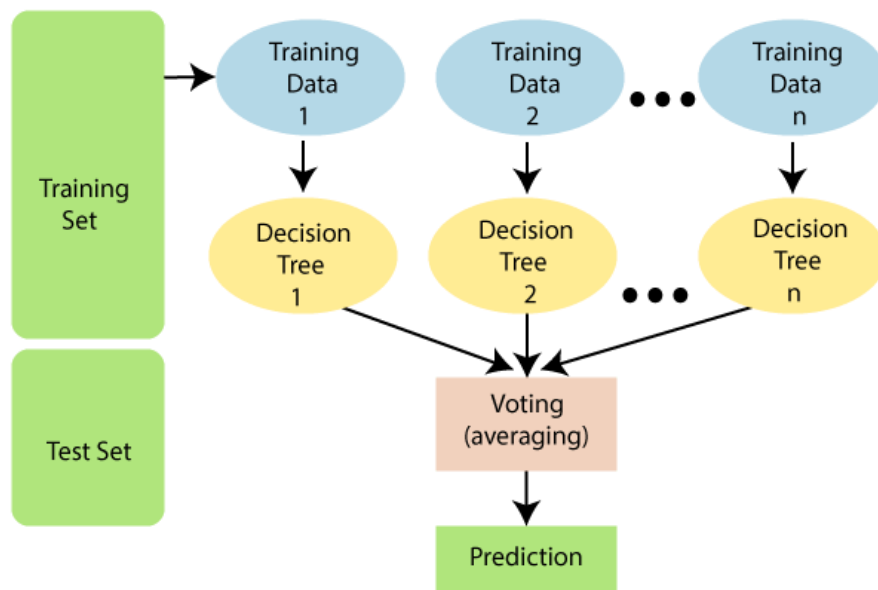


Figure 27 : Algorithme de random forest

### Utilisation :

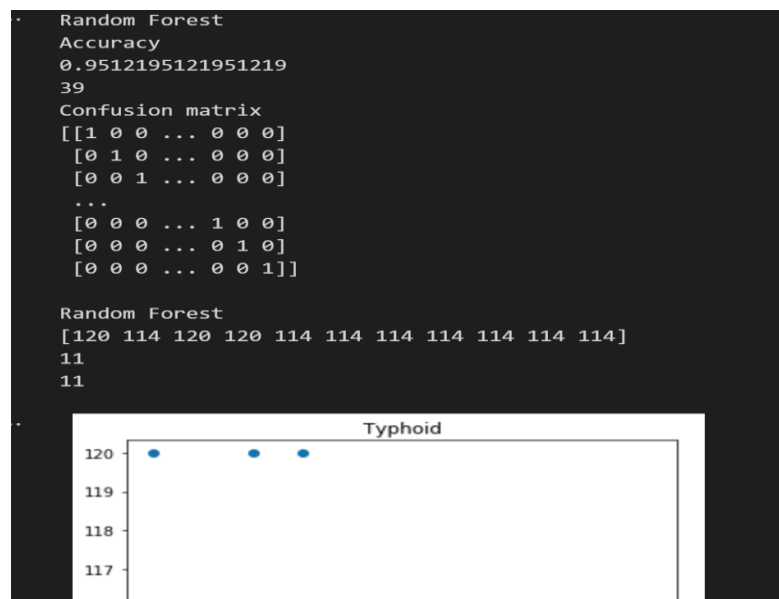


Figure 28: accuracy

### 4.3.3 Algorithme Naive Bayes

L'algorithme Naive Bayes est basé sur la théorie bayésienne et repose sur une hypothèse forte d'indépendance entre les symptômes. Malgré cette hypothèse simplificatrice, Naive Bayes est connu pour donner de bons résultats dans des contextes où les données médicales sont modérément corrélées.

- ❖ Avantages : Naive Bayes est rapide à entraîner et efficace même avec des données de grande dimension. Il est particulièrement adapté aux problèmes où la compréhension des probabilités sous-jacentes est importante.
- ❖ Inconvénients : L'hypothèse d'indépendance des variables est souvent irréaliste dans un contexte médical, car de nombreux symptômes sont interconnectés. Cela peut parfois réduire la précision des prédictions, surtout en présence de données fortement corrélées. (Zhang, 2004)

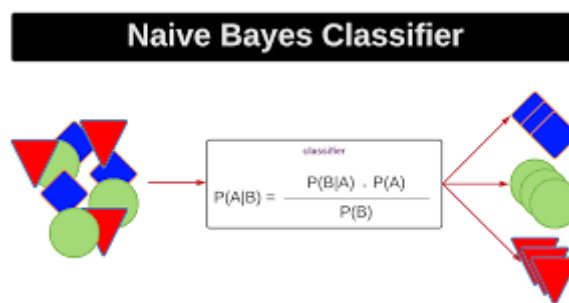


Figure 29: naive bayes

8

## Utilisation :

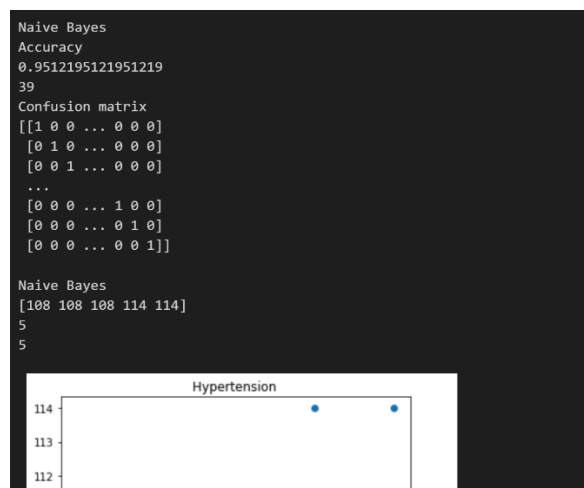


Figure 30: accuracy

#### 4.3.4 Algorithme *K-Nearest Neighbors* (KNN)

L'algorithme K-Nearest Neighbors est une approche de classification basée sur la similarité entre les cas. Pour prédire une maladie, cet algorithme compare les symptômes du patient avec ceux des patients précédemment répertoriés et classe le patient dans la catégorie de la maladie qui correspond à ses voisins les plus proches dans l'ensemble de données.

- ❖ Avantages : KNN est simple à mettre en œuvre et fonctionne bien lorsque les classes (maladies) sont bien séparées dans l'espace des données. Il est non paramétrique, ce qui signifie qu'il ne fait aucune hypothèse sur la distribution des données.
- ❖ Inconvénients : KNN peut être coûteux en termes de calcul, car il nécessite la comparaison de chaque nouveau cas avec l'ensemble des données d'entraînement. De plus, il est sensible aux valeurs aberrantes et à l'échelle des données (une normalisation est souvent nécessaire).

##### Utilisation :



Figure 31: accuracy

## 4.4 Conclusion

Dans ce chapitre, nous avons détaillé les différentes étapes de la modélisation et de la mise en œuvre d'un système de prédiction de maladies basé sur des symptômes. Nous avons commencé par la collecte et la préparation des données, qui constitue une étape essentielle pour garantir la qualité et la pertinence des prédictions. Ensuite, nous avons exploré plusieurs algorithmes de machine learning, notamment l'algorithme de l'arbre de décision, la forêt aléatoire, le K-Nearest Neighbors (KNN) et le Naive Bayes. Chacun de ces algorithmes a été sélectionné en fonction de ses performances et de ses capacités à traiter des données médicales complexes.



# Chapitre 5 : Mise en place d'une plateforme pour permettre l'interaction utilisateur

## 5.1 Introduction

L'objectif final de notre étude est de développer une plateforme interactive permettant aux utilisateurs, qu'ils soient des patients ou des professionnels de santé, d'interagir facilement avec le système de prédiction de maladies. Cette plateforme servira de pont entre les algorithmes de machine learning développés et l'utilisateur final, offrant une interface intuitive et accessible pour le diagnostic médical assisté par l'intelligence artificielle.

## 5.2 Architecture globale du système

L'architecture du système global repose sur un modèle client-serveur, avec une interaction fluide entre l'utilisateur, les algorithmes de prédiction, et les bases de données. Voici les principales composantes de l'architecture :

- ❖ **Frontend (Client)** : L'interface utilisateur est accessible via un navigateur web ou une application mobile. C'est ici que l'utilisateur saisit ses symptômes, reçoit les prédictions et interagit avec la plateforme.
- ❖ **Backend (Serveur)** : Le serveur héberge l'application, exécute les algorithmes de machine learning et gère les bases de données. Il reçoit les requêtes du client, les traite, et renvoie les résultats.
- ❖ **Base de données** : Une base de données centralisée stocke les informations médicales, les symptômes, les prédictions passées et les autres données utilisateur. Elle est essentielle pour la persistance des données et leur récupération rapide.
- ❖ **IA**: un modèle exporté et enregistré avec pickle.
- ❖ **API** : développement de mon propre API avec golang et Flask

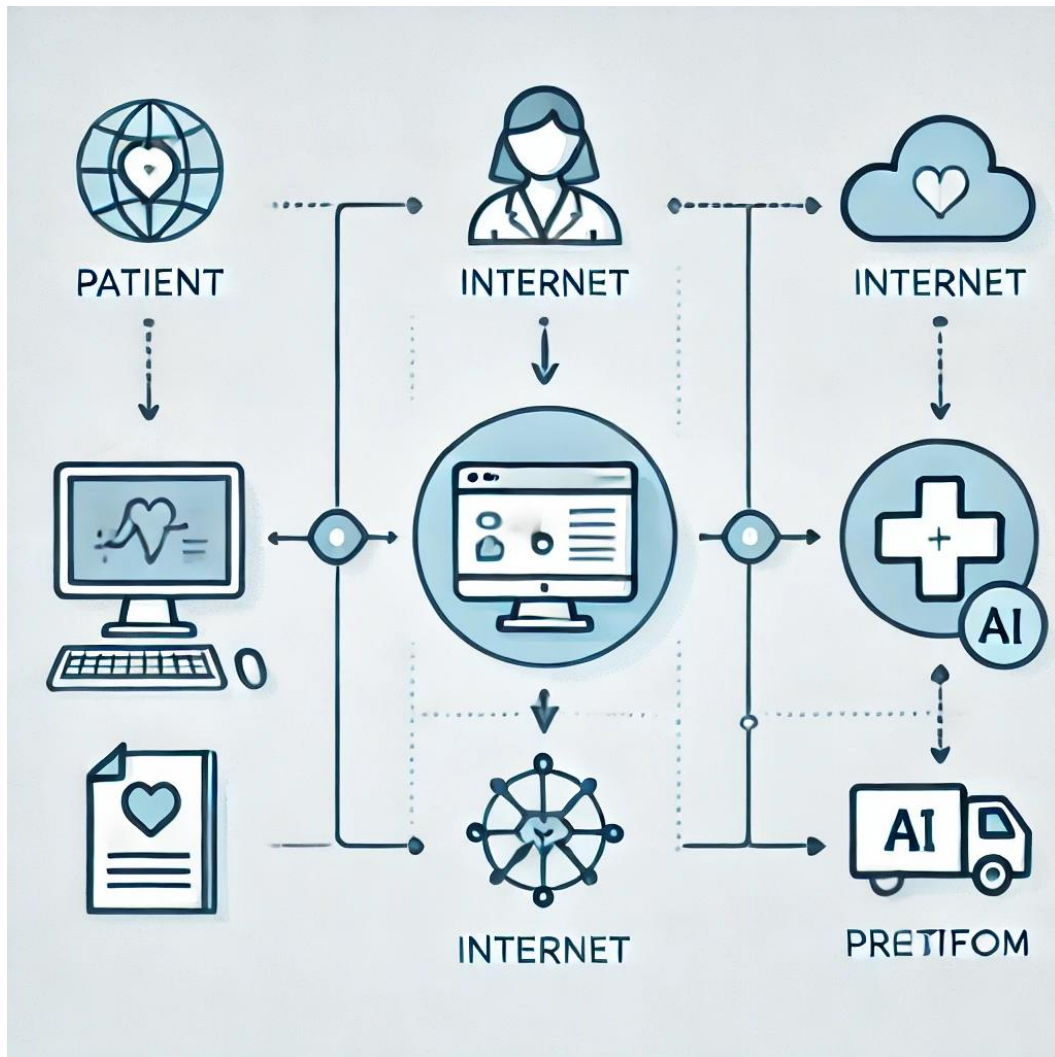


Figure 32: séquence de fonctionnement du système

### 5.3 Technologies utilisées

Pour une approche beaucoup plus moderne nous avons donc choisi d'utiliser une architecture backend, frontend comme dans le cadre du web2.0.

### 5.4 Architecture du backend :

Pour la conception de l'architecture backend de la plateforme de prédiction de maladies, deux technologies principales ont été utilisées :

- ❖ **Flask** : Un micro-framework web en Python, utilisé ici pour créer des endpoints permettant d'intégrer et de charger le modèle d'intelligence artificielle (IA).
- ❖ **Golang (Go)** : Un langage de programmation rapide et performant, utilisé pour développer l'application web qui interagit avec l'utilisateur.

### 5.4.1 Flask

Flask est utilisé pour exposer un **endpoint API**. Cet endpoint est chargé de recevoir les requêtes des utilisateurs (symptômes envoyés), d'invoquer le modèle d'IA pour faire les prédictions, et de renvoyer les résultats au frontend de l'application. Flask est particulièrement adapté pour ce type de tâche car il est léger, flexible et bien intégré avec Python, facilitant ainsi la manipulation du modèle d'IA.

- ❖ **Chargement du modèle** : Le modèle d'IA, entraîné sur des données de maladies et symptômes, est intégré dans le backend Flask. Lorsqu'une requête est envoyée, Flask appelle le modèle pour prédire la maladie la plus probable en fonction des symptômes fournis.
- ❖ **API REST** : L'endpoint développé en Flask fonctionne comme une API REST qui reçoit les données en format JSON, les traite, puis retourne les résultats sous forme de prédiction dans une structure JSON simple à interpréter pour le frontend.



Figure 33:flask

### 5.4.2 Golang

Golang est utilisé pour gérer la partie **web application** du système. Grâce à ses performances et à sa capacité à gérer de nombreuses connexions simultanées et des potentiels erreurs, Golang permet de créer une interface utilisateur interactive et réactive.

- ❖ **Interaction avec Flask** : L'application web en Go interagit avec l'API Flask via des requêtes HTTP, en envoyant les informations des utilisateurs (symptômes) et en recevant les prédictions de l'IA en retour.
- ❖ **Gestion des utilisateurs** : L'application web, développée en Go, gère également les sessions utilisateur, la collecte des informations via des formulaires web, et l'affichage des résultats prédictifs.

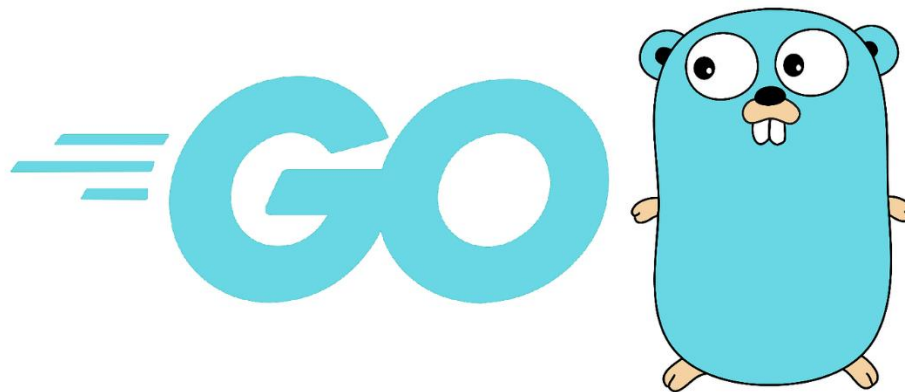


Figure 34: golang

#### 5.4.3 Communication entre les composants :

L'architecture repose sur la communication entre Flask (chargé de l'IA) et Go (pour l'application web). Ces deux technologies interagissent principalement via des appels HTTP.

- ❖ Le frontend web (développé avec Go) récupère les inputs des utilisateurs et les envoie à l'API Flask.
- ❖ Flask traite ces données, appelle le modèle IA, et renvoie la prédiction au frontend.
- ❖ Le frontend Go affiche ensuite les résultats à l'utilisateur, offrant une interface simple pour la consultation des maladies potentielles.

#### Avantages de cette architecture :

- ❖ **Modularité** : La séparation entre le modèle IA (Flask) et l'application web (Go) permet de développer, maintenir et améliorer chaque composant indépendamment.
- ❖ **Performance** : Go est réputé pour sa rapidité, ce qui en fait un choix idéal pour une application web nécessitant une réponse rapide et capable de gérer plusieurs utilisateurs simultanément.
- ❖ **Flexibilité** : L'utilisation de Flask pour l'IA permet de tirer parti des bibliothèques Python pour le machine learning, tout en conservant la légèreté nécessaire pour déployer le modèle rapidement.

## 5.5 Architecture du frontend.

Pour la partie frontend de la plateforme, les technologies suivantes ont été utilisées :

- ❖ **HTML** : Pour structurer le contenu de la page web.
- ❖ **CSS** : Pour styliser l'interface utilisateur, assurer une mise en page attractive et responsive.
- ❖ **JavaScript (JS)** : Utilisé de manière native pour rendre l'interface interactive et gérer les interactions avec le backend.

### 5.5.1 HTML

L'interface utilisateur a été développée en **HTML** afin de structurer les différentes sections de la page, notamment le formulaire où les utilisateurs peuvent entrer leurs symptômes. Les éléments HTML incluent des champs de texte, des boutons et des sections pour afficher les résultats de la prédiction.

- ❖ **Formulaire de saisie** : Un formulaire HTML permet aux utilisateurs de renseigner leurs symptômes, qui sont ensuite envoyés au backend pour analyse.
- ❖ **Affichage des résultats** : Après la prédiction de la maladie par l'IA, les résultats sont affichés directement sur la page via des éléments HTML bien définis, comme des div ou des tableaux.

### 5.5.2 CSS

**CSS** a été utilisé pour styliser les différentes composantes de la plateforme afin d'offrir une interface utilisateur moderne, claire et facile à naviguer.

- ❖ **Responsivité** : Grâce à des techniques comme les **media queries**, l'interface est responsive, s'adaptant automatiquement aux différentes tailles d'écran (smartphones, tablettes, ordinateurs).
- ❖ **Mise en page et esthétique** : Des grilles, des animations légères et des effets de survol permettent une expérience utilisateur agréable et fluide.

### 5.5.3 JavaScript natif

L'usage de **JavaScript natif** dans cette architecture a pour but de rendre l'application interactive et dynamique. Aucune bibliothèque ou framework (comme React ou Angular) n'a été utilisé, privilégiant un code simple et léger.

- ❖ **Envoi des données au backend** : JS est utilisé pour capturer les informations saisies par l'utilisateur dans le formulaire, et les envoyer au backend Flask via des requêtes **AJAX** ou des **fetch API**. Cette approche permet de rendre l'application plus fluide sans avoir besoin de recharger toute la page à chaque fois qu'une prédiction est effectuée.
- ❖ **Traitement des réponses** : Une fois que le backend Flask renvoie la prédiction, JavaScript récupère cette réponse et met à jour dynamiquement l'interface utilisateur pour afficher les résultats sans rechargement de la page.

## 5.6 Résultat final

Le système se présente d'abord sous deux pages : une page d'accueil, et une page de test de l'IA.

La page d'accueil se présente ainsi, tout d'abord nous avons fait un descriptif et mis des boutons de navigation.

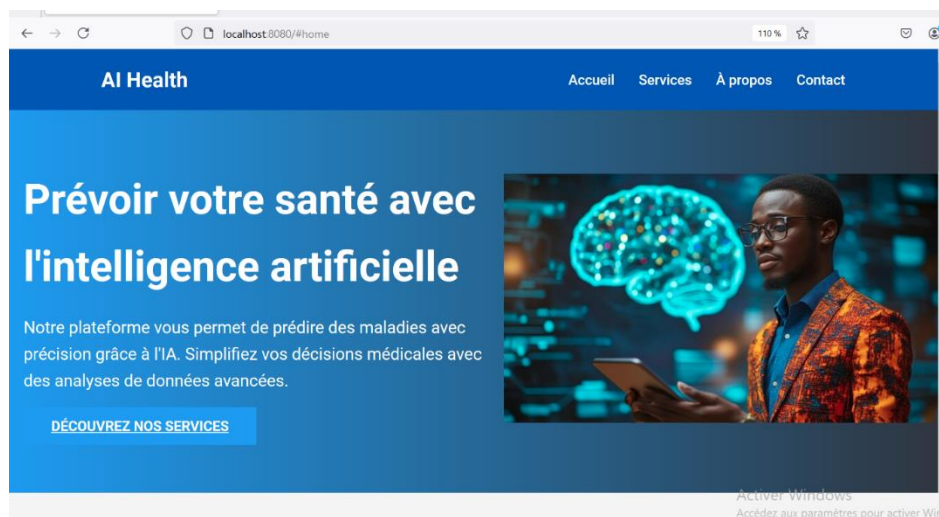


Figure 35: interface graphique



Figure 36: application web



Figure 37: bouton web

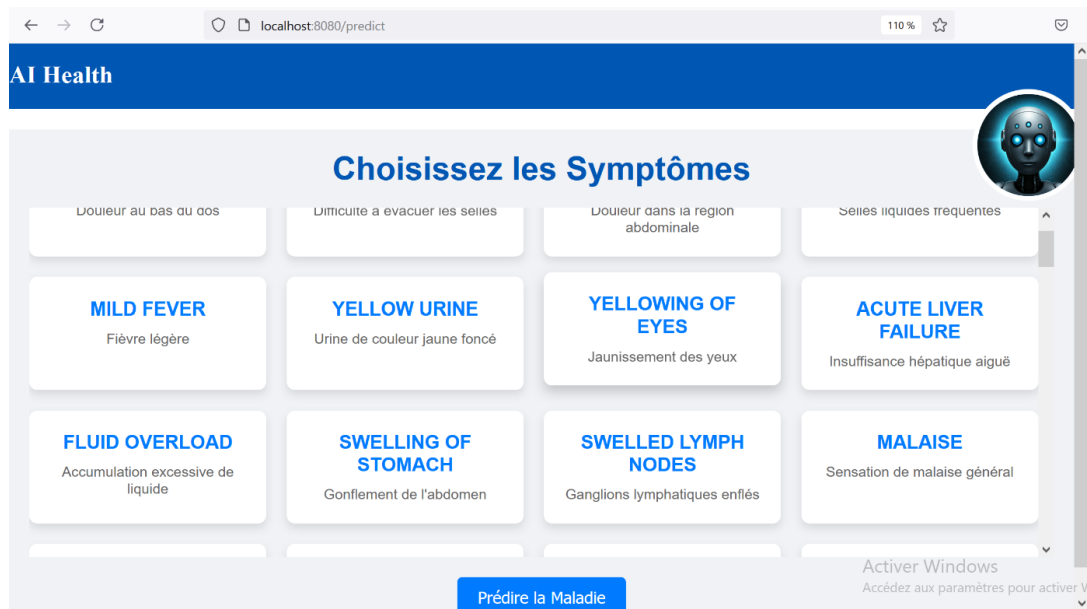


Figure 38: prompt de l'application web



## 5.7 Conclusion

La mise en place de cette plateforme de prédiction des maladies constitue une avancée significative dans l'amélioration de l'accès aux soins et la prise en charge préventive. En intégrant plusieurs technologies, nous avons pu créer un système robuste qui relie de manière fluide les utilisateurs à l'IA pour la prédiction des maladies.

L'architecture backend repose sur **Flask** pour la gestion des API d'intelligence artificielle, tandis que **Golang** assure une gestion performante de l'application web côté serveur. D'un point de vue frontend, l'utilisation combinée de **HTML**, **CSS** et **JavaScript** permet de créer une interface utilisateur réactive et intuitive, facilitant l'interaction des patients avec la plateforme. Cette structure garantit une rapidité de traitement, une évolutivité, et une fiabilité dans les prédictions de maladies.

L'implémentation de cette architecture apporte non seulement une solution technique efficace, mais elle pose aussi les bases d'une plateforme évolutive, capable de s'adapter à de nouveaux besoins et de nouvelles technologies à l'avenir. La plateforme, en étant flexible et modulaire, permet de répondre à des cas d'usage variés tout en offrant une expérience utilisateur optimisée.



## 6 Conclusion Générale

Ce travail de recherche et de développement a permis de concevoir et de mettre en œuvre une plateforme de prédiction de maladies basée sur l'intelligence artificielle. Nous avons parcouru plusieurs étapes, allant de la compréhension des concepts théoriques de l'IA, à la collecte et l'analyse des données médicales, jusqu'à la modélisation et l'intégration des algorithmes prédictifs. Le projet a démontré la pertinence et l'efficacité des modèles d'apprentissage automatique, tels que le **Decision Tree**, **Random Forest**, **K-Nearest Neighbors**, et **Naive Bayes** dans le domaine de la prédiction des maladies à partir de symptômes fournis.

Le développement technique, en utilisant **Flask** pour la gestion des API et **Golang** pour la création de l'application web backend, a permis de construire un système robuste et évolutif. Le **frontend**, développé avec **HTML**, **CSS** et **JavaScript**, offre une interface utilisateur intuitive, facilitant l'interaction des patients avec la plateforme.

L'objectif initial, qui était de fournir un outil accessible et performant capable d'assister les patients dans l'identification précoce de maladies potentielles, a été largement atteint. Les résultats obtenus montrent que l'utilisation des algorithmes de machine learning peut améliorer significativement la détection précoce de maladies et contribuer à une meilleure prise en charge médicale.

Cependant, il reste des perspectives d'amélioration. L'intégration de nouvelles sources de données, telles que des données en temps réel provenant de dispositifs IoT, ou l'ajout de fonctionnalités de suivi médical personnalisé, pourraient enrichir la plateforme et accroître son efficacité. En outre, une attention particulière devra être portée à l'aspect éthique et à la protection des données des utilisateurs.

Ainsi, cette plateforme ouvre la voie à de futures recherches et développements dans le domaine de la e-santé, contribuant à l'amélioration des services de santé accessibles à distance et à la prévention médicale assistée par l'intelligence artificielle.

## Références :

- [1].Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York, NY: Springer.
- [2].Bostrom, N. (. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- [3].Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- [4].Breiman, L. (2001). *Machine Learning*.
- [5].Ian Goodfellow, Y. B. (2016). *Deep Learning*. Cambridge, MA: MIT Press.
- [6].Nilsson, N. J. (1998). *Artificial Intelligence: A New Synthesis*. San Francisco, CA: Morgan Kaufmann.
- [7].Quinlan, J. R. (1986). *Induction of Decision Trees*.
- [8].Stuart Russell, P. N. (2010). *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall.
- [10].Zhang, H. (2004). Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society.
- [11]. [https://fr.freepik.com/images-ia-premium/presentation-innovante-intelligence-artificielle-hologramme-cerebral-numerique\\_342195563.htm#fromView=search&page=1&position=12&uuid=fab3c302-145c-48f5-8a8f-3720baf63ba](https://fr.freepik.com/images-ia-premium/presentation-innovante-intelligence-artificielle-hologramme-cerebral-numerique_342195563.htm#fromView=search&page=1&position=12&uuid=fab3c302-145c-48f5-8a8f-3720baf63ba)
- [12]. [https://fr.freepik.com/photos-premium/medecin-medecin-afro-americaain-tenant-tablette-illustration-covid\\_19412179.htm#fromView=search&page=1&position=7&uuid=25082ff5-c344-4298-9e33-32b989a55e30](https://fr.freepik.com/photos-premium/medecin-medecin-afro-americaain-tenant-tablette-illustration-covid_19412179.htm#fromView=search&page=1&position=7&uuid=25082ff5-c344-4298-9e33-32b989a55e30)
- <sup>1</sup> <https://gt2.ariis.fr/les-algorithmes-dexploitation/lapprentissage-supervise/>
- <sup>1</sup> <https://fr.linedata.com/quest-ce-que-lapprentissage-non-supervise>
- <sup>1</sup> <https://penseeartificielle.fr/tout-pour-bien-debuter-en-deep-learning-5/principe-apprentissage-par-renforcement/>
- <sup>1</sup> <https://www.aspexit.com/reseau-de-neurones-on-va-essayer-de-demystifier-un-peu-tout-ca-1/>
- <sup>1</sup>
- [14].[https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.oezratty.net%2Fwordpress%2F2016%2Favancees-intelligence-artificielle-6%2F&psig=AOvVaw2T2yU\\_VvfeHFRgQLO5oKXY&ust=1728955168640000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCLjXl5jajIkDFQAAAAAdAAAAABAE](https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.oezratty.net%2Fwordpress%2F2016%2Favancees-intelligence-artificielle-6%2F&psig=AOvVaw2T2yU_VvfeHFRgQLO5oKXY&ust=1728955168640000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCLjXl5jajIkDFQAAAAAdAAAAABAE)
- <sup>1</sup> [Disease \(columbia.edu\)](#)
- [15].<https://www.google.com/url?sa=i&url=https%3A%2F%2Fdataaspirant.com%2Fhow-decision-tree-algorithm-works%2F&psig=AOvVaw1ZuuZPiW8HG6rBloEYtpIb&ust=1729133148614000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCKD0pZzxkYkDFQAAAAAdAAAAABAJ>

- [16].<https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.javatpoint.com%2Fmachine-learning-random-forest-algorithm&psig=AOvVaw1zCA6z9HCpvDo9qSgsSrFS&ust=1729133316650000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCNCRwuHxkYkDFQAAAAAdAAAAABAE>
- [17]. <https://www.google.com/url?sa=i&url=https%3A%2F%2Fmlarchive.com%2Fmachine-learning%2Fthe-ultimate-guide-to-naive-bayes%2F&psig=AOvVaw3ODK9W2w4aozAP2lo0ugmw&ust=1729136279975000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCLDws7T9kYkDFQAAAAAdAAAAABAE>
- [18].<https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.monocubed.com%2Fblog%2Fflask-vs-django%2F&psig=AOvVaw36F-dJM2dcagz4bKrrgtPh&ust=1729735842056000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCMiotai2o4kDFQAAAAAdAAAAABAE>
- [19].<https://www.google.com/url?sa=i&url=https%3A%2F%2Fmedium.com%2Fnerd-for-tech%2Flearn-golang-in-one-blog-fdd568e6f631&psig=AOvVaw3ujmYC4gFMJD6PuTdL02d&ust=1729736134679000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCODL2bC3o4kDFQAAAAAdAAAAABAE>
- [20]<sup>1</sup>[https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.oezratty.net%2Fwordpress%2F2016%2Favancees-intelligence-artificielle-6%2F&psig=AOvVaw2T2yU\\_VvfeHFRgQLO5oKXY&ust=1728955168640000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCLjX15jaiIkDFQAAAAAdAAAAABAE](https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.oezratty.net%2Fwordpress%2F2016%2Favancees-intelligence-artificielle-6%2F&psig=AOvVaw2T2yU_VvfeHFRgQLO5oKXY&ust=1728955168640000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCLjX15jaiIkDFQAAAAAdAAAAABAE)
- [21]fig3: [https://fr.freepik.com/photos-premium/medecin-medecin-afro-americaain-tenant-tablette-illustration-covid\\_19412179.htm#fromView=search&page=1&position=7&uuid=25082ff5-c344-4298-9e33-32b989a55e300](https://fr.freepik.com/photos-premium/medecin-medecin-afro-americaain-tenant-tablette-illustration-covid_19412179.htm#fromView=search&page=1&position=7&uuid=25082ff5-c344-4298-9e33-32b989a55e300)
- [22] fig 4 <https://gt2.ariis.fr/les-algorithmes-dexploitation/lapprentissage-supervise/>
- [23]fig 5 : [https://fr.freepik.com/images-ia-premium/presentation-innovante-intelligence-artificielle-hologramme-cerebral-numerique\\_342195563.htm#fromView=search&page=1&position=12&uuid=fab3c302-145c-48f5-8a8f-3720baf63ba](https://fr.freepik.com/images-ia-premium/presentation-innovante-intelligence-artificielle-hologramme-cerebral-numerique_342195563.htm#fromView=search&page=1&position=12&uuid=fab3c302-145c-48f5-8a8f-3720baf63ba)
- [24] Fig 9: <sup>1</sup> <https://fr.linedata.com/quest-ce-que-lapprentissage-non-supervisee>
- [25]fig10 :<https://penseeartificielle.fr/tout-pour-bien-debuter-en-deep-learning-5/principe-apprentissage-par-renforcement/>
- [25 ] fig 11: <sup>1</sup> <https://www.aspexit.com/reseau-de-neurones-on-va-essayer-de-demystifier-un-peu-tout-ca-1/>
- [27] fig17 : <sup>1</sup> <https://www.google.com/url?sa=i&url=https%3A%2F%2Fdataaspirant.com%2Fhow-decision-tree-algorithm-works%2F&psig=AOvVaw1ZuuZPiW8HG6rBloEYtpIb&ust=1729133148614000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCKD0pZzxkYkDFQAAAAAdAAAAABAJ>

[28] fig 26: <sup>1</sup> <https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.javatpoint.com%2Fmachine-learning-random-forest-algorithm&psig=AOvVaw1zCA6z9HCpvDo9qSgsSrFS&ust=1729133316650000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCNCRwuHxkYkDFQAAAAAdAAAAABAE>

[29] fig 34: <sup>1</sup>

<https://www.google.com/url?sa=i&url=https%3A%2F%2Fmedium.com%2Fnerd-for-tech%2Flearn-golang-in-one-blog-fdd568e6f631&psig=AOvVaw3ujmYC4gFMJDu6PuTdL02d&ust=1729736134679000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCODL2bC3o4kDFQAAAAAdAAAAABAE>

[30] fig 30: <https://www.google.com/url?sa=i&url=https%3A%2F%2Fmlarchive.com%2Fmachine-learning%2Fthe-ultimate-guide-to-naive-bayes%2F&psig=AOvVaw3ODK9W2w4aozAP2lo0ugmw&ust=1729136279975000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCLDws7T9kYkDFQAAAAAdAAAAABAE>

[31] fig 33: <sup>1</sup>

<https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.monocubed.com%2Fblog%2Fflask-vs-django%2F&psig=AOvVaw36F-dJM2dcagz4bKrrgtPh&ust=1729735842056000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCMiotai2o4kDFQAAAAAdAAAAABAE>