**Université Cheikh Anta Diop**

**Ecole Supérieure Polytechnique**

Département Génie Informatique

Master 2 en Intelligence Artificielle et Big Data (IABD)

**Veille technologique**

*2023*

*-*

*2024*



Réalisé

par :

**Mohamed ADEJA**

**Ali Mahomet Taher**

**E**

**l ghazali Moctar**

**Enseignant :**

**Pr. Doudou FALL**

**Projet de programmation : IA, Big Data et sécurité**

**:**

*MediAsiist*

*:*

Assistant Intelligent pour la prédiction du Diabète

[Table des matières **Partie 1 – Charte** 3](#_Toc10588)

[**Vision:** 3](#_Toc10589)

[**Langage d'implémentation :** 3](#_Toc10590)

[**Choix de la base de données :** 3](#_Toc10591)

[**Dépôt de contrôle de version :** 4](#_Toc10592)

[**Membres de l'équipe et coordonnées :** 4](#_Toc10593)

[**Réunion de groupe hebdomadaire :** 4](#_Toc10594)

[**Principales caractéristiques :** 4](#_Toc10595)

[**Description narrative :** 5](#_Toc10596)

[**Partie 2 – Document sur les exigences** 6](#_Toc10597)

[**Objectif du système** 6](#_Toc10598)

[**Exigences fonctionnelles :** 6](#_Toc10599)

[**Analyse de menaces** 8](#_Toc10600)

[**Objectifs de sécurité** 10](#_Toc10601)

[**Partie 3 et Partie 4 –Document d'Assurance Qualité pour le Projet** 10](#_Toc10602)

[**Approche d'Assurance Qualité** 10](#_Toc10603)

[**Tâches Accomplies** 11](#_Toc10604)

[**Outil d'analyse statique : Flake8** 11](#_Toc10605)

[**Code finale du projet :** 12](#_Toc10606)

**Partie 1 – Charte**

## Vision :

Il s’agit d’un assistant intelligent conçu pour les médecins, visant à améliorer les diagnostics et les soins pour leurs patients. Il fournit des évaluations médicales précises et rapides, basées sur les informations fournies par les médecins, leur permettant de prendre des décisions éclairées et de fournir des traitements adaptés.

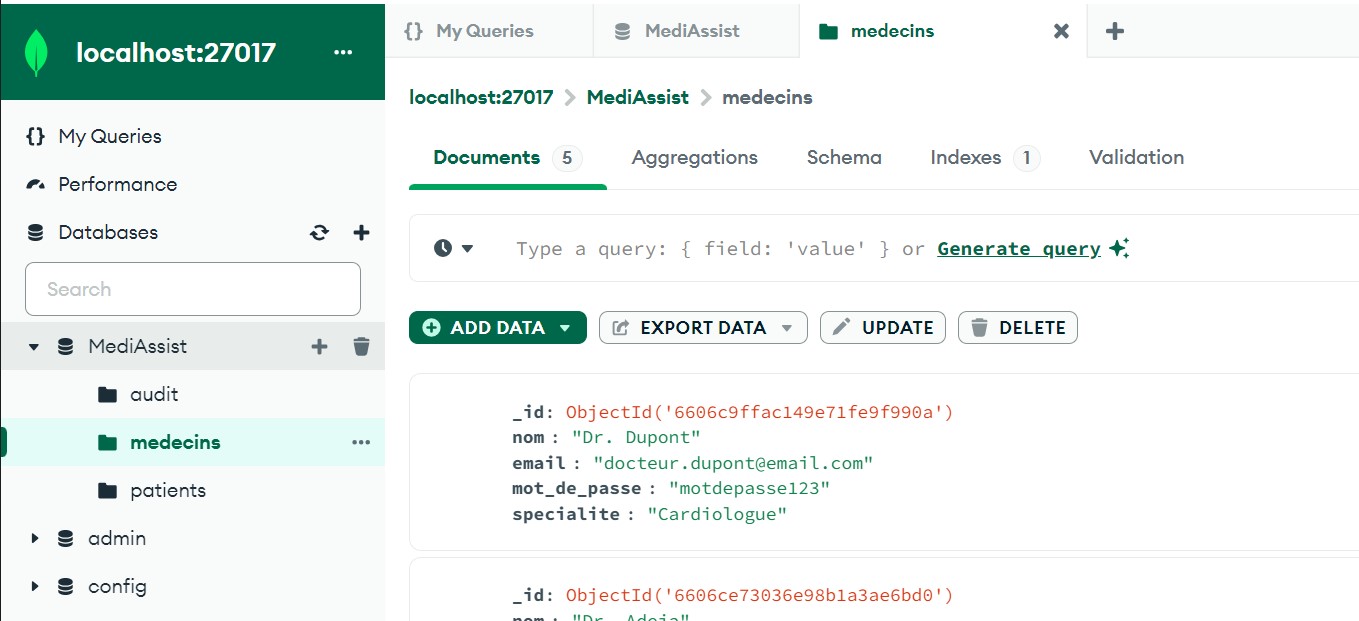
## Langage d'implémentation :

Nous proposons d'utiliser le langage Python pour l'implémentation du projet . Nous avons une solide expérience dans le développement d'applications en utilisant Python, ainsi que dans l'utilisation de bibliothèques populaires d'apprentissage automatique telles que TensorFlow, scikit-learn et Keras.

## Choix de la base de données :

Nous avons choisi MongoDB comme base de données pour notre application, et nous allons nous connecter à cette base de données en utilisant PyMongo qui est une bibliothèque Python qui permet d'interagir avec MongoDB. Elle fournit une API simple et intuitive pour effectuer des opérations de lecture, d'écriture et de manipulation des données dans la base de données.

Voici notre base de données avec ses différentes collections :



## Dépôt de contrôle de version :

Nous utiliserons GitHub comme plateforme de contrôle de version pour le développement du logiciel. Vous pouvez accéder au dépôt du projet à l'adresse suivante : <https://github.com/mtam360/projet_systeme_logiciel.git>

## Membres de l'équipe et coordonnées :

* Ali Mahamat Taher : alimahamattaher@esp.sn
* El Ghazali Moctar : elghazaliahmedmoctar@esp.sn
* Mohamed ADEJA : mohamedadeja@esp.sn

## Réunion de groupe hebdomadaire :

Nous prévoyons de nous réunir en tant qu'équipe chaque semaine, en dehors des cours, pour discuter de l'état d'avancement du projet. La réunion aura lieu tous les mardis et les jeudis à 14h00 en ligne ou en présentiel. Les membres de l'équipe sont tenus d'assister à la réunion et de partager leurs mises à jour, leurs défis et leurs idées.

## Principales caractéristiques :

* **Collecte de données médicales :** notre logiciel permettra aux médecins de saisir les données médicales et les symptômes de leurs patients de manière structurée et conviviale.
* **Diagnostic médical avancé :** Grâce à des techniques d'apprentissage automatique, l'assistant fournit des évaluations médicales précises basées sur les informations fournies par les médecins.
* **Recommandations de traitement :** notre assistant propose des recommandations de traitement basées sur les meilleures pratiques médicales et les données disponibles.
* **Interface utilisateur intuitive :** L'assistant est doté d'une interface utilisateur conviviale, permettant aux médecins de naviguer facilement et d'interagir avec le système.
* **Confidentialité et sécurité des données :** Les données médicales des patients sont traitées de manière confidentielle et sécurisée, conformément aux réglementations en matière de protection des données.
* **Support multiplateforme :** le logiciel est accessible sur différentes plateformes, telles que les ordinateurs de bureau et les appareils mobiles.

**Éléments de sécurité essentiels :**

* Chiffrement des données : Les données médicales sensibles sont chiffrées en transit et au repos pour garantir leur confidentialité.
* Contrôle d'accès : Un système de contrôle d'accès est mis en place pour permettre uniquement aux médecins autorisés d'accéder aux données médicales sensibles.
* Gestion des identités : Des mécanismes d'authentification et d'autorisation robustes sont utilisés pour vérifier l'identité des médecins et contrôler leur accès au système.
* Suivi des journaux et audit : Un système de suivi des journaux est implémenté pour enregistrer les activités du système, détecter les tentatives d'accès non autorisées et faciliter les audits de sécurité.

## Description narrative :

Notre projet consiste à mettre en place un assistant intelligent conçu spécifiquement pour les médecins, leur offrant un moyen efficace d'améliorer les diagnostics et les soins pour leurs patients. L'assistant permet aux médecins de saisir les données médicales et les symptômes de manière structurée, facilitant ainsi l'analyse et la recherche d'informations pertinentes. Grâce à des techniques d'apprentissage automatique avancées, notre logiciel permettra de fournir des évaluations médicales précises basées sur les informations fournies par les médecins. Ces évaluations aident les médecins à prendre des décisions éclairées et à fournir des traitements adaptés à leurs patients.

L'interface utilisateur intuitive permet aux médecins d'interagir facilement avec l'assistant et d'accéder aux résultats de manière claire et compréhensible. Une fois qu'un médecin a créé son compte, il a la possibilité de créer un nouveau patient à tout moment. Il peut commencer par remplir le dossier médical du patient, bien que toutes les informations ne soient pas obligatoires dans le formulaire. De plus, il est toujours possible de modifier ces informations ultérieurement. Ensuite, le médecin peut engager une discussion avec le modèle d'intelligence artificielle qui a été spécifiquement entraîné sur des données de santé. Au cours de cette conversation, le médecin peut discuter des symptômes du patient et évoquer différents scénarios possibles afin de parvenir à un diagnostic précis. Le modèle peut également recommander au médecin de poser des questions complémentaires au patient, ce qui pourrait améliorer la prise de décision.

La confidentialité et la sécurité des données sont des priorités essentielles raison pour laquelle toutes les données médicales sensibles sont chiffrées en transit et au repos pour garantir leur confidentialité. Un système de contrôle d'accès est mis en place pour que seuls les médecins autorisés puissent accéder aux données sensibles. Des mécanismes d'authentification et d'autorisation robustes sont utilisés pour vérifier l'identité des médecins et contrôler leur accès au système. De plus, un système de suivi des journaux est implémenté pour enregistrer les activités du système, détecter les tentatives d'accès non autorisées et faciliter les audits de sécurité.

Il s’agit donc d’une solution multiplateforme, accessible sur différents appareils tels que les ordinateurs de bureau et les appareils mobiles. Il vise à améliorer les diagnostics et les soins médicaux en fournissant aux médecins des outils avancés pour prendre des décisions éclairées et fournir des traitements adaptés à leurs patients.

# Partie 2 – Document sur les exigences

Ce document sur les exigences vise à définir les fonctionnalités et les contraintes du système d'assistant intelligent pour les médecins. Il servira de référence pour le développement du logiciel et assurera une compréhension claire des attentes du projet.

## Objectif du système

Notre système a pour objectif de mettre en place un assistant médical intelligent spécialement conçu pour les médecins, afin d'améliorer les diagnostics et les soins prodigués à leurs patients. Grâce à une interface intuitive et conviviale, les médecins peuvent interagir facilement avec l'assistant et accéder aux résultats de manière claire et compréhensible.

L'assistant permet aux médecins de saisir les données médicales et les symptômes de manière structurée, facilitant ainsi l'analyse et la recherche d'informations pertinentes. En utilisant des techniques d'apprentissage automatique avancées, notre logiciel fournit des évaluations médicales précises basées sur les informations fournies par les médecins. Ces évaluations aident les médecins à prendre des décisions éclairées et à fournir des traitements adaptés à leurs patients.

## Exigences fonctionnelles :

**2.1 Enregistrement des médecins:**

Type d'utilisateur : Médecins

Actifs : Informations d'identification appropriées

Importance : Must have

User story : En tant que médecin, je veux pouvoir créer un compte sécurisé en fournissant des informations d'identification appropriées.

**2.2 Gestion des patients :**

Type d'utilisateur : Médecins

Actifs : Informations médicales des patients, dossiers des patients

Importance : Must have

User story : En tant que médecin, je veux pouvoir ajouter de nouveaux patients à ma liste, enregistrer leurs informations médicales de base et accéder facilement à leurs dossiers.

**2.3 Saisie structurée des données médicales :**

Type d'utilisateur : Médecins

Actifs : Données médicales des patients (antécédents médicaux, symptômes, traitements précédents, etc.)

Importance : Should have

User story : En tant que médecin, je veux pouvoir saisir les données médicales des patients de manière structurée, en utilisant des champs appropriés pour les informations telles que les antécédents médicaux, les symptômes, les traitements précédents, etc.

**1.4 Évaluations médicales précises :**

Type d'utilisateur : Médecins

Actifs : Techniques d'apprentissage automatique avancées, informations fournies par les médecins

Importance : should have

User story : En tant que médecin, je veux pouvoir engager une conversation interactive avec l'assistant IA, discuter des symptômes du patient et explorer différents scénarios pour parvenir à un diagnostic précis. Les évaluations fournis par l’assistant doivent être claires, compréhensibles et utiles dans mon processus de diagnostic et de traitement.

**2.5 Administration du système :**

Type d'utilisateur : Administrateur du système

Actifs : Comptes utilisateur, autorisations d'accès, données médicales

Importance : Must have

User story : En tant qu'administrateur du système, je veux pouvoir gérer les comptes utilisateur et les autorisations d'accès et pouvoir, de plus, effectuer des sauvegardes régulières des données médicales pour assurer la sécurité et la récupération en cas de problème.

## Analyse de menaces

**Menaces malveillantes potentielles :**

1. Accès non autorisé aux données sensibles : Des attaquants pourraient tenter de contourner les mécanismes de contrôle d'accès pour accéder aux informations médicales sensibles des patients. Cela pourrait compromettre la confidentialité et l'intégrité des données.
2. Vol ou divulgation de données sensibles : Des attaquants pourraient chercher à voler ou à divulguer les données médicales sensibles des patients, ce qui pourrait entraîner des conséquences néfastes pour la vie privée et la réputation des patients.
3. Manipulation des données : Les attaquants pourraient tenter de manipuler les données médicales des patients afin de fausser les résultats ou d'induire en erreur les médecins lors de la prise de décisions médicales.
4. Attaques par déni de service : Des attaquants pourraient chercher à perturber le fonctionnement du système en lançant des attaques par déni de service, ce qui pourrait entraîner une interruption du service et une perte de productivité pour les médecins.

**Attaquants potentiels :**

1. Pirates informatiques : Des individus ou des groupes de pirates informatiques pourraient chercher à accéder illégalement aux données médicales sensibles dans le but de les vendre ou de les exploiter à des fins lucratives.
2. Professionnels de la santé malveillants : Des médecins ou d'autres professionnels de la santé ayant un accès autorisé au système pourraient abuser de leurs privilèges et accéder à des informations confidentielles pour des motifs personnels ou malveillants.
3. Concurrents : Des concurrents ou des entreprises concurrentes pourraient tenter de compromettre votre système dans le but de voler des informations médicales sensibles ou de perturber vos opérations.
4. Organisations criminelles : Des organisations criminelles pourraient être intéressées par les données médicales sensibles en raison de leur valeur sur le marché noir, où elles pourraient être utilisées à des fins frauduleuses ou pour des activités illégales.

**Vulnérabilités potentielles :**

1. Vulnérabilités logicielles : Des vulnérabilités dans les logiciels utilisés pour le système, y compris les bibliothèques tierces et les frameworks, pourraient être exploitées par des attaquants pour compromettre la sécurité du système.
2. Erreurs de configuration : Des erreurs de configuration dans les paramètres de sécurité du système pourraient permettre à des attaquants d'accéder à des fonctionnalités ou à des données sensibles de manière non autorisée.
3. Défauts de conception : Des défauts de conception dans l'architecture du système pourraient créer des points faibles qui pourraient être exploités par des attaquants.
4. Ingénierie sociale : Les attaquants pourraient utiliser des techniques d'ingénierie sociale pour manipuler les utilisateurs autorisés et obtenir des informations confidentielles ou des accès non autorisés.

**Mesures de sécurité recommandées :**

1. Authentification et autorisation robustes : Mettez en place des mécanismes d'authentification solides pour vérifier l'identité des médecins et contrôler leur accès au système. Utilisez des méthodes d'autorisation appropriées pour limiter l'accès aux données sensibles en fonction des privilèges des utilisateurs.
2. Chiffrement des données : Appliquez des techniques de chiffrement pour protéger les données médicales sensibles en transit et au repos, en veillant à ce que les clés de chiffrement soient bien protégées.
3. Surveillance des journaux d'activité : Mettez en place un système de surveillance des journaux d'activité pour détecter toute activité suspecte ou non autorisée.
4. Tests de sécurité réguliers : Effectuez des tests de sécurité réguliers pour identifier les vulnérabilités du système et appliquer les correctifs nécessaires.
5. Sensibilisation à la sécurité : Fournissez une formation et une sensibilisation régulières aux médecins et au personnel sur les meilleures pratiques de sécurité, y compris la gestion des mots depasse, la protection des informations sensibles et la détection des attaques potentielles.

## Objectifs de sécurité

**Confidentialité des données :** Toutes les données médicales sensibles doivent être chiffrées lors de leur transit et de leur stockage afin de garantir leur confidentialité.

**Contrôle d'accès :** Un système de contrôle d'accès robuste doit être mis en place pour permettre uniquement aux médecins autorisés d'accéder aux données médicales sensibles.

**Authentification et autorisation :** Des mécanismes d'authentification solides doivent être utilisés pour vérifier l'identité des médecins et contrôler leur accès au système. Des autorisations appropriées doivent être définies pour limiter l'accès aux fonctionnalités spécifiques en fonction du rôle et des privilèges du médecin.

# Partie 3 et Partie 4 –Document d'Assurance Qualité pour le Projet

## Approche d'Assurance Qualité

Afin de garantir la sécurité, la fiabilité et la transparence de notre système, nous avons mis en œuvre une approche d'assurance qualité rigoureuse, couvrant l'ensemble du cycle de développement, de la gestion de projet aux tests en passant par la sécurité du code.

**Gestion de projet agile pour une collaboration efficace:**

* Adoption d'une méthodologie agile avec des sprints réguliers et des revues de code fréquentes.
* Collaboration transparente au sein de l'équipe pour une progression fluide et une prise de décision rapide.

**Tests complets pour un système robuste:**

* Mise en place d'une stratégie de test complète comprenant des tests unitaires, d'intégration, fonctionnels et de sécurité.
* Garantie de la robustesse et de la fiabilité du système grâce à des tests rigoureux.

**Analyse statique pour un code de qualité supérieure:**

* Utilisation d'outils d'analyse statique pour évaluer la qualité du code source.
* Détection des erreurs de programmation et identification des vulnérabilités potentielles dès le départ.

**Sécurité des données au cœur de nos priorités :**

* Implémentation de mécanismes de cryptage robustes pour protéger la confidentialité des données des électeurs.
* Mise en place de contrôles d'accès stricts pour limiter l'accès aux données sensibles.
* Protection des données contre les intrusions et les cybermenaces grâce à des mesures de sécurité adéquates.

## Tâches Accomplies

À ce jour, les principales tâches accomplies dans le cadre du projet sont les suivantes :

* **Enregistrement des médecins**

Un médecin peut créer un compte sécurisé en fournissant des informations d'identification appropriées au niveau de la base de données.

* **Gestion des patients :**

Un médecin a la possibilité d’ajouter de nouveaux patients à sa liste, enregistrer leurs informations médicales de base et accéder facilement à leurs dossiers.

* **Évaluations médicales précises :**

Le médecin peut grâce au modèle d’intelligence artificiel évaluer la situation sanitaire du patient et d'identifier les risques potentiels pour la santé des patients plus tôt et de prendre des décisions médicales plus éclairées.

## Outil d'analyse statique : Flake8

Dans le cadre de notre travail, nous avons effectué une analyse approfondie des outils d'analyse statique disponibles pour améliorer la qualité et la fiabilité de notre code. Après avoir examiné différentes options, nous avons décidé d'adopter **Flake8** comme outil principal d'analyse statique.

**Flake8** est un outil d'analyse statique de code pour le langage de programmation Python. Il est largement utilisé dans la communauté du développement Python pour améliorer la qualité, la lisibilité et la cohérence du code. Flake8 combine plusieurs outils d'analyse de code, notamment PyFlakes, pycodestyle (anciennement connu sous le nom de pep8) et McCabe, pour fournir une analyse complète du code source.

Voici un aperçu des fonctionnalités offertes par Flake8 :

1. **Vérification de conformité aux normes de codage :** Flake8 vérifie si le code respecte les normes de codage recommandées par PEP 8, qui est la convention de codage officielle pour Python. Il signale les erreurs de style, les indentations incorrectes, les longueurs de ligne dépassant les limites, les espaces inutiles, etc.
2. **Détection des erreurs potentielles :** Flake8 détecte les erreurs courantes dans le code telles que les variables non utilisées, les imports inutiles, les assignations incorrectes, les expressions redondantes, les blocs de code inaccessibles, etc. Cela permet de repérer les problèmes susceptibles de causer des bugs ou des comportements indésirables.
3. **Calcul de la complexité cyclomatique :** Flake8 utilise l'outil McCabe pour calculer la complexité cyclomatique du code. La complexité cyclomatique mesure la complexité structurelle d'une fonction ou d'un module et fournit des indications sur sa maintenabilité. Un code avec une complexité cyclomatique élevée peut être difficile à comprendre et à maintenir.
4. **Personnalisation des règles** : Flake8 offre la possibilité de personnaliser les règles d'analyse en utilisant des fichiers de configuration. Cela permet d'adapter les vérifications spécifiques aux besoins du projet ou de l'organisation.

Division des données en X et y ---

target\_name='Outcome'

y= df\_new[target\_name]

X=df\_new.drop(target\_name,axis=1)

Y prédictions données - données d'entraînement

suppression de la colonne Résultat et conservation de toutes les autres colonnes sous la forme X

X.head()

**Algorithmes de classification :**

**Support Vector Machine (SVM) :**

**Le SVM est un algorithme d'apprentissage supervisé utilisé pour la classification et la régression. Il cherche à trouver l’hyperplan optimal séparant des classes avec la plus grande marge. Lorsque les données ne sont pas linéaires, une fonction de noyau permet de projeter les données dans un espace de dimension supérieure. J'ai choisi cet algorithme pour mon projet Assistant Intelligent pour le Diagnostic Médical, car il est efficace pour classer des données complexes et pertinent dans le cadre de la détection de pathologies.**

**Enregistrement du modèle**

**Pour pouvoir l'importer et enchainer directement avec son utilisation que de le réentrainer encore.**

**import os**

**os.makedirs('./saved\_models', exist\_ok=True)**

**import pickle**

**pickle.dump(model, open('./saved\_models/model\_01.pkl', 'wb'))**

**# Ajuster (fit) le modèle sur les données d'entraînement**

**model.fit(X\_train, y\_train)**

**# Sauvegarder le modèle**

**import joblib**

**joblib.dump(model, "saved\_models/model\_01.pkl")**

**['saved\_models/model\_01.pkl']**

# Code finale du projet :

<https://github.com/mtam360/projet_systeme_logiciel.git>