

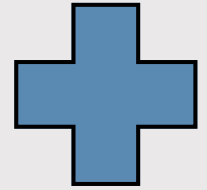
PROJET



WLAB

Créé par
Wild's Anatomy

2024



SOMMAIRE



01. L'EQUIPE

02. LES OBJECTIFS

03. LES RESSOURCES

04. L'ORGANISATION

05. LES RECHERCHES

06. LES ALGORITHMES

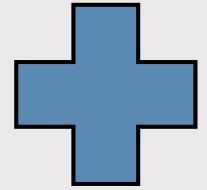
07. LANCEMENT APPLICATION

08. AXES D'AMELIORATION

09. LIENS DES TRAVAUX



L'EQUIPE



COMPOSEE DE 5 DATA
ANALYST

Luana, Victoria, Laura, Linda
et Pauline

Notre crédo :
**"Votre problème, notre
solution !"**





LES OBJECTIFS

Dans le cadre de recherches sur certaines maladies chroniques, l'Institut Pasteur et l'INRA nous sollicitent.

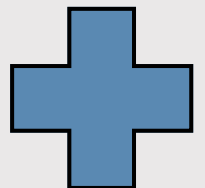
Ils souhaitent des algorithmes capables de prédire une maladie ainsi que leur intégration dans un applicatif qui sera utilisé dans le milieu médical.

1

PREDICTIONS DE MALADIES

2

APPLICATIF FACILE D'UTILISATION





LES RESSOURCES

5 JEUX DE DONNEES MEDICALES ANONYMES
5 MALADIES CHRONIQUES A ETUDIER



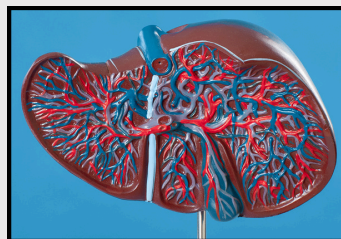
COEUR



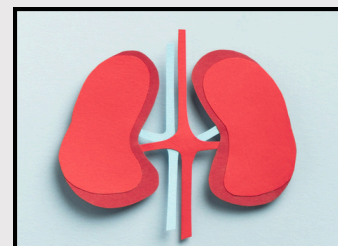
CANCER DU SEIN



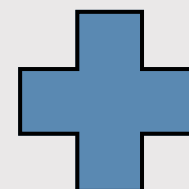
DIABETE



FOIE



REINS





L'ORGANISATION

1

2

3

4

5

Plannification

- Brainstorming
- Méthode SCRUM
- Création d'un AirTable d'équipe
- Création d'un Drive commun
- Partage des bases de données
- Planning par sprint hebdomadaire

Méthodologie

- SPRINT = 1 semaine
- Création d'un référentiel commun
- Création d'un workflow commun
- Nettoyage individuel des données
- Partage des difficultés
- Solution en équipe

Technologies

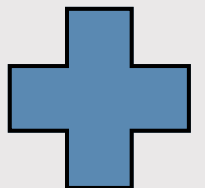
- AirTable
- Drive/Discord
- Google scholar
- Pandas (Python)
- NumPy
- Matplotlib
- Seaborn
- Scikit-learn
- Joblib
- Django
- Streamlit

Activités

- Lecture des données
- Lectures scientifiques
- Compréhension des variables scientifiques
- Nettoyage et traitements
- Analyses descriptives
- Classification et choix des métriques
- Encodage
- Machine learning
- Création applicatif et intégration

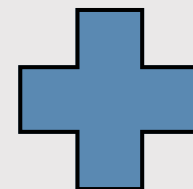
Ajustements

- Vérifications et contrôles
- Harmonisation des modèles
- Derniers ajustements
- Création de la présentation

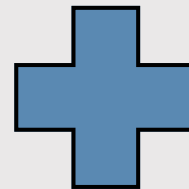




**NOUS FAISONS
TOUJOURS DE
NOTRE MIEUX
POUR NOTRE
CLIENT ET NOUS
AVONS
ENTREPRIS DES
RECHERCHES
POUR
COMPRENDRE
LES DONNEES
MEDICALES**



LES RECHERCHES



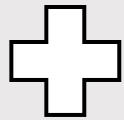
Recherches
sur les sites
dédiés aux
scientifiques

- Création d'un référentiel commun
- Workflow commun

Croisement
des
informations
entre les
maladies



Recherche des
meilleurs
algorithmes adaptés
au milieu médical



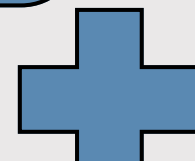
LES ALGORITHMES

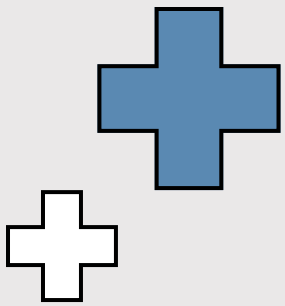
Après avoir effectué nos recherches, nettoyé et préparé les jeux de données : lancement des modèles de prédictions

Chaque jeu de données dispose d'une colonne avec une cible binaire correspondant au malade/non malade.

Cette colonne essentielle, nous a permis de tester nos modèles et de l'affiner.

Utilisation de techniques telles que la Cross Validation pour trouver le meilleur modèle





**MODELE :
RANDOM FOREST**

Encoding : Standard Scaler

**MODELE :
RANDOM FOREST**

Encoding : Standard Scaler

**CANCER
DU SEIN**

REINS

COEUR

FOIE

DIABETE

**MODELE :
RANDOM FOREST**

Encoding : LabelEncoder et Scaler

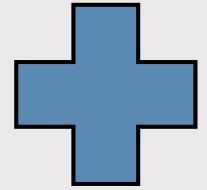
MODELE : NAIVE BAYES

Encoding : Standard scaler

**MODELE :
RANDOM FOREST**

Encoding : Standard Scaler

L'APPLICATION



Nous avons choisi Streamlit comme interface visuelle.

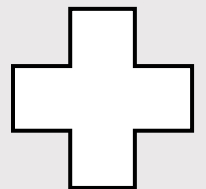
Une réflexion sur l'inclusion en créant une interface spécifique pour les daltoniens a également été appliquée.



AXES D'AMELIORATION



- Implémentation d'un système d'authentification
- Appliquer pour d'autres maladies
- Déploiement de l'application



LIENS TRAVAUX

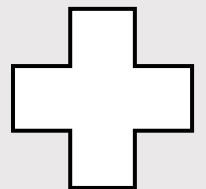


01. [Le AirTable](#)

02. [Veille documentée](#)

03. [Workflow commun](#)

04. [Lien GitHub](#)



**MERCI POUR
VOTRE
CONFIANCE**

