

# **PLAN GÉNÉRAL DU RAPPORT COMBINÉ**

## **INTRODUCTION GÉNÉRALE (2 pages)**

### **1. Contexte médical et scientifique commun**

- Importance des maladies cardiovasculaires et pulmonaires
- Rôle croissant de l'IA en santé (données cliniques & imagerie)
- Enjeux communs : déséquilibre des données, sur-confiance des modèles, risque clinique

### **2. IA en santé : promesses et limites**

- IA comme aide à la décision
- Problèmes de généralisation, biais, sur-apprentissage
- Importance de la fiabilité et de l'incertitude

### **3. Problématique générale du rapport**

Comment développer et comparer des modèles d'intelligence artificielle appliqués à des données hétérogènes (imagerie et données cliniques), capables de fournir des prédictions performantes et fiables, tout en intégrant une estimation de l'incertitude pour limiter les décisions médicales à risque ?

### **4. Objectifs généraux**

- Comparer des approches ML et DL selon le type de données
- Étudier l'impact du déséquilibre des classes
- Évaluer la performance et la fiabilité probabiliste
- Introduire et comparer des méthodes d'incertitude

### **5. Organisation du rapport**

- Partie I : cadre méthodologique commun
- Partie II : application 1 – données cliniques (PARTIE LEA)
- Partie III : application 2 – imagerie pulmonaire (PARTIE Aziz)

## **PARTIE I \_ CADRE MÉTHODOLOGIQUE COMMUN ( 5 pages)**

### **Chapitre 1 Cadre général et principes méthodologiques**

#### **1. Types de données en IA médicale**

- Données tabulaires cliniques
- Données d'imagerie médicale

- Contraintes spécifiques et différences

## **2. Problématique du déséquilibre des classes**

- Effets sur l'apprentissage
- Métriques trompeuses
- Stratégies communes (pondération, seuils, SMOTE)

## **3. Évaluation des modèles**

- Discrimination (ROC, PR)
- Métriques dépendantes de la prévalence
- Limites de l'accuracy

## **4. Fiabilité, calibration et incertitude**

- Pourquoi la performance ne suffit pas
- Calibration, probabilités, confiance
- Incertitude comme outil de réduction du risque

## **PARTIE II \_APPLICATION 1**

### **Prédiction du risque cardiovasculaire à partir de données cliniques ( 10 pages) Partie Léa**

#### **Chapitre 2 Matériels et méthodes**

- Jeu de données (Framingham)
- Prétraitement et imputation
- Gestion du déséquilibre
- Modèles étudiés (LogReg, SVM, RF, XGB)
- Validation, bootstrap
- Calibration, seuil de décision
- Incertitude (bootstrap, conformal prediction)

#### **Chapitre 3 \_Résultats**

- Comparaison des modèles
- Impact du déséquilibre
- Choix du modèle final
- Apport de la calibration
- Analyse de l'incertitude

## **PARTIE III \_APPLICATION 2**

# **Classification de maladies pulmonaires à partir d'images (10 pages) Partie AZIZ**

## **Chapitre 4 — Matériels et méthodes**

- Dataset d'images
- Prétraitement
- MLP + PCA
- CNN
- Gestion du déséquilibre
- Métriques
- Incertitude (softmax, MC Dropout)

## **Chapitre 5 — Résultats**

- Résultats MLP
- Résultats CNN
- Comparaisons
- Analyse des erreurs confiantes
- Visualisation de l'incertitude

## **DISCUSSION GÉNÉRALE TRANSVERSALE (5 pages)**

### **1. Comparaison des deux approches**

- Données cliniques vs imagerie
- ML interprétable vs DL complexe

### **2. Gestion du déséquilibre : enseignements communs**

- Pondération vs seuil
- Robustesse des approches

### **3. Apport central de l'incertitude**

- Différences bootstrap / MC Dropout
- Intérêt clinique partagé

### **4. Limites communes**

- Données
- Généralisation
- Absence de validation clinique réelle

### **5. Apports académiques et professionnels**

- Compétences méthodologiques

- Vision critique de l'IA en santé

## CONCLUSION GÉNÉRALE ( $\leq 2$ pages)

- Synthèse des deux travaux
- Message central : **performance  $\neq$  fiabilité**
- Positionnement professionnel
- Ouverture vers l'IA responsable en santé