

# Deep Learning pour prédire les effets post- radiothérapie

Soutenance finale  
Projet 3A HSB



**Thomas Ménard** et **Clara Cousteix**, encadré.e.s par Mahmoud Bentriou, Sarah Lemler et Véronique Letort-Lechevalier



CentraleSupélec



# Table des matières

**I.**

## **Introduction**

Contexte et enjeux

**II.**

## **Matériel**

Présentation des différents datasets utilisés

**III.**

## **Méthodes**

Analyses de survie, Machine Learning et Deep Learning

**IV.**

## **Résultats**

Prédiction de maladie cardiaque par Machine et Deep Learning



# **Introduction : Contexte et objectifs**

# Quelques chiffres clés



**1750**

Nouveaux cancers de  
l'enfant diagnostiqués  
chaque année en  
France

**80%**

des cancers de l'enfant  
guéris, parfois non sans  
effets nocifs

**11%**

Des patients sont  
diagnostiqués avec une  
maladie cardiaque 40 ans  
après le traitement



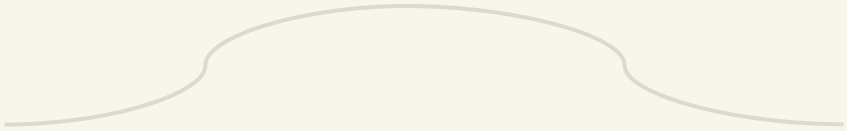
# 3 principaux traitements

## Divers traitements existent :

- La chirurgie
- La radiothérapie
- La chimiothérapie

## peut avoir des effets secondaires graves:

- Radiothérapeutique : L'irradiation détruit les cellules cancéreuses tout en préservant le mieux possible les tissus sains
- Risque de déclencher une maladie cardiaque suite à un traitement anti-cancéreux



## **Objectif : Prédire l'apparition de maladie cardiaque chez des patients traités par radiothérapie pour un cancer de l'enfant**

**Pour cela, on utilise des méthodes d'analyse de données :**

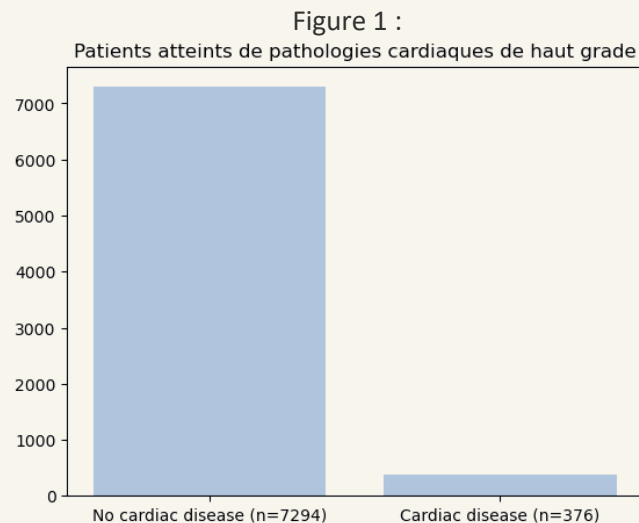
- ✓ **Etablir de premiers résultats par Machine Learning**
  - ✓ **Améliorer la prédiction grâce au Deep Learning**
- 

**II.**

# **Présentation de nos données**

# Cohorte FCCSS : French Childhood Cancer Survivor Study

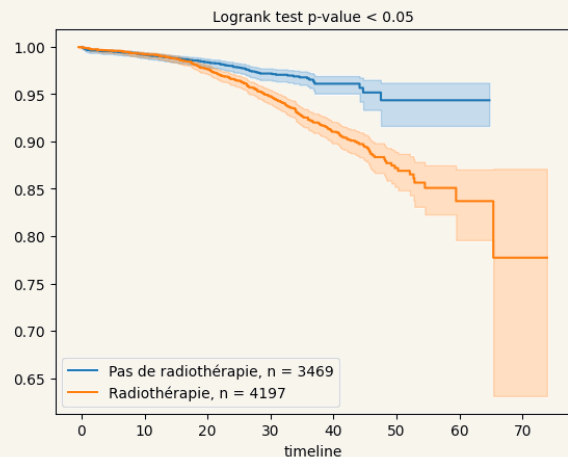
- Etude coordonnée par une équipe épidémiologie des radiations à l'Institut Gustave Roussy
- Rassemble **7670 patients** traités pour un cancer de l'enfant
- **But** : Etudier les effets à long terme des traitements et leurs facteurs de risque. Suivi sur parfois plus de **50 ans**.
- Données très variées :
  - **Données administratives** : centre de soin, identifiant du patient, date de naissance
  - **Données temporelles** : *date de diagnostic, date du dernier suivi*
  - **Données cliniques** : Type de cancer, traitements de reçus, *dose de médicaments de chimiothérapie*, métastases éventuelles, rechute, *maladie cardiaque* (fig1)





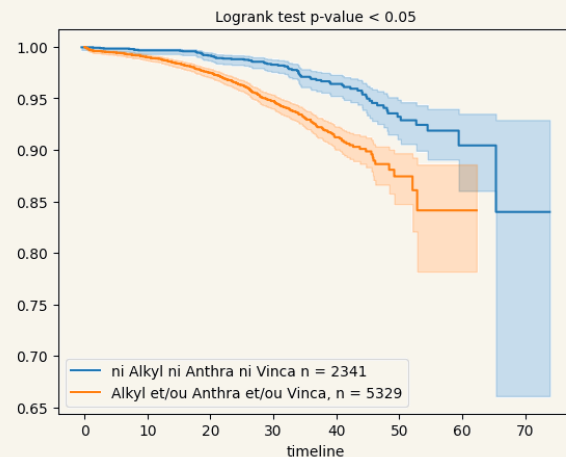
# Courbes de survie Kaplan Meier

Courbes de survie comparant un traitement avec et sans radiothérapie



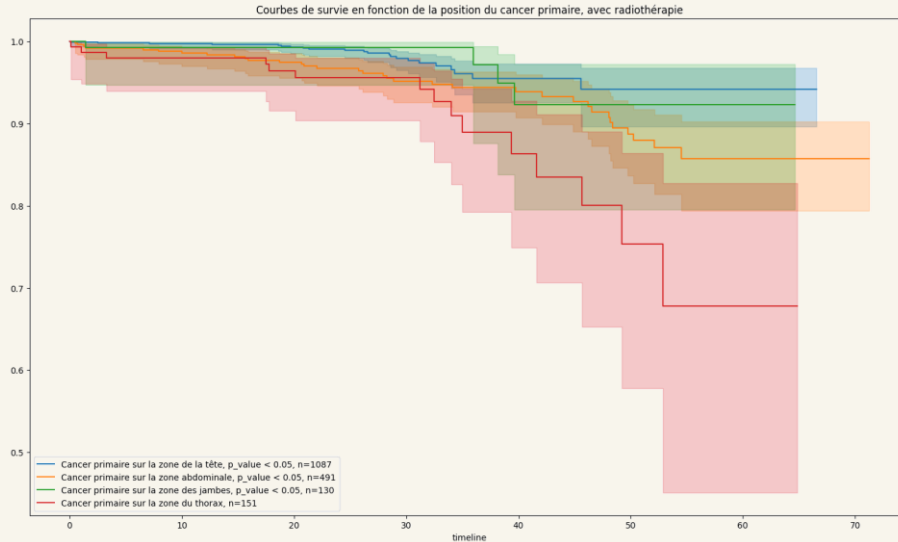
**Fig 1** : un traitement avec radiothérapie a un impact significatif sur l'apparition d'une maladie cardiaque de haut grade (p-value < 0.05)

Traitement avec et sans agents alkylants ou d'anthracyclines ou de vinca-alcaloïdes



**Fig 2** : un traitement avec chimiothérapie a un impact significatif sur l'apparition d'une maladie cardiaque (p-value < 0.05)

# Courbes de survie Kaplan Meier

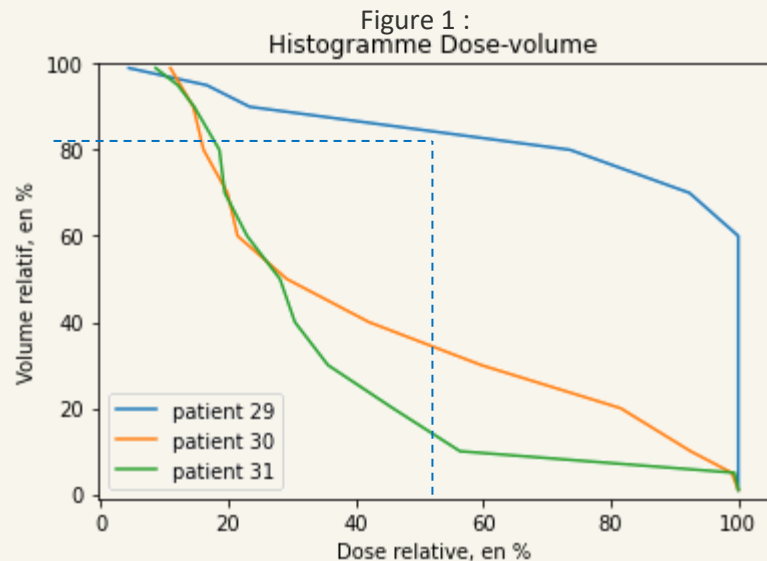


**Résultat :** Les analyses de survie montrent que certains facteurs peuvent avoir un impact significatif pour l'apparition de maladies cardiaques

**Fig 3 :** la position du cancer primaire, dans le cadre de la radiothérapie, semble avoir un impact significatif sur l'apparition d'une maladie cardiaque (p-value < 0.05)

# Histogramme dose-volume

- Représente la **distribution de rayonnement** sous forme de quantiles. Ex : 80% du cœur a reçu au moins 60% de la dose (Fig1) Permet de caractériser un rayonnement
- Plusieurs volumes d'intérêt : cœur, oreillettes G/D, ventricules G/D, myocarde
- Lien entre les indicateurs dose-volume et la cohorte FCCSS : **3943 patients**, dont **282** avec pathologie cardiaque
- Prédiction des effets sur le cœur à partir des indicateurs dose-volume : **Machine Learning**
- **Limite** : pas d'information sur la localisation de la zone irradiée



# Matrice de dose 3D

- **Matrice de dose** : reconstitution 3D du volume irradié (cœur) pendant les séances de radiothérapie. Données + informatives
- **Image NIFTI** : voxels représentés en nuances de gris (Fig 1)
- Format type image : il faudra utiliser des **CNN**
- Lien entre matrices de doses et la cohorte FCCSS : **3943 patients**, dont **282** avec pathologie cardiaque

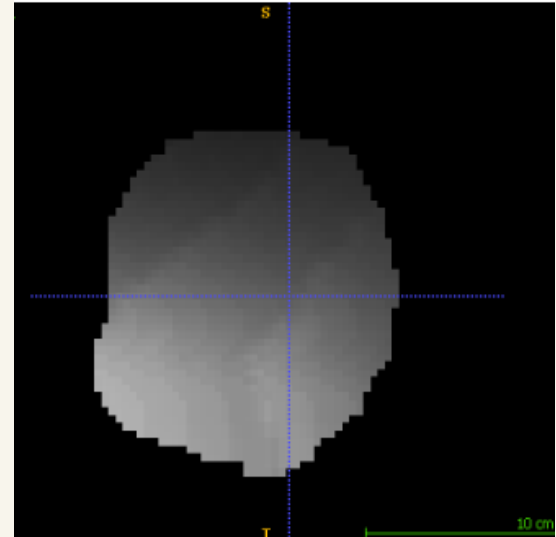
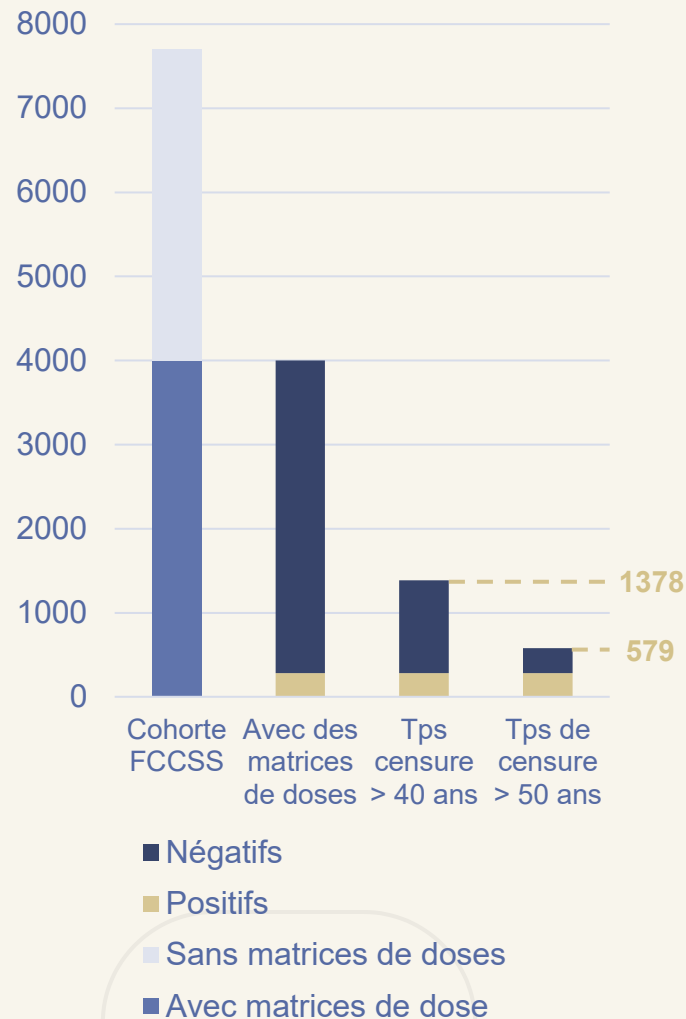
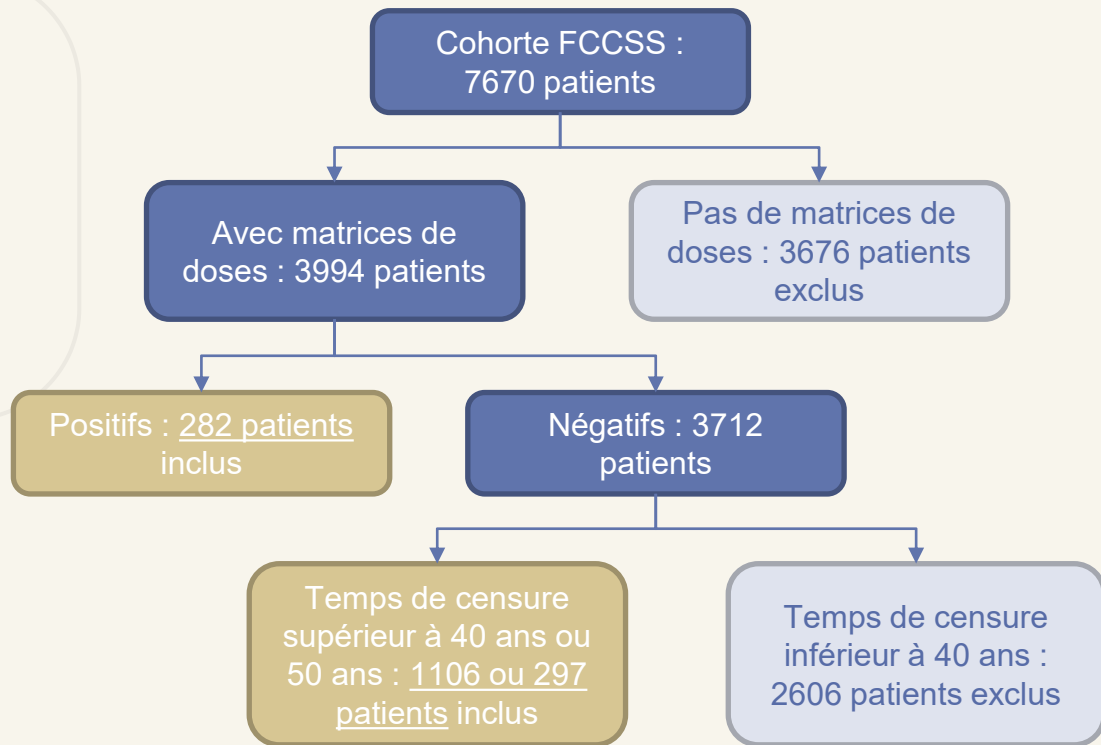


Figure 1 : Matrice de dose 3D visualisée avec ITK-SNAP

# Sélection de patients pour l'étude



# Sélection des patients pour l'étude

- **Deux groupes de patients** sélectionnés pour la suite de l'étude.
  - Temps de censure > 40 ans : 1378 patients
  - Temps de censure > 50 ans : 579 patients
- **Partage en trois datasets** : train/validation/test, stratifiés sur le label :
  - Train = 60 %
  - Validation = 20 %
  - Test = 20 %

# III.



# Méthodes

Machine Learning et Deep Learning

# Dataset déséquilibré

- **Resampling** : oversampling de la classe minoritaire ou undersampling de la classe majoritaire [1]
- **Reweighting**: inclure des poids dans la loss pour prendre en compte le caractère déséquilibré du dataset [1]
- **Bootstrap** : Méthode de resampling poussée, illustrée ci-contre

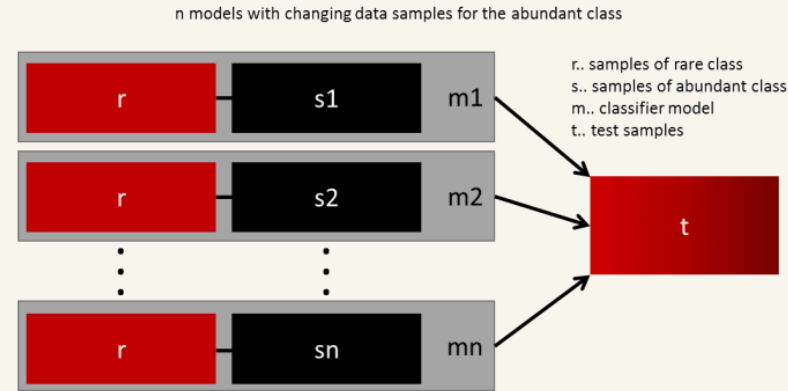


Figure : Schéma Bootstrap

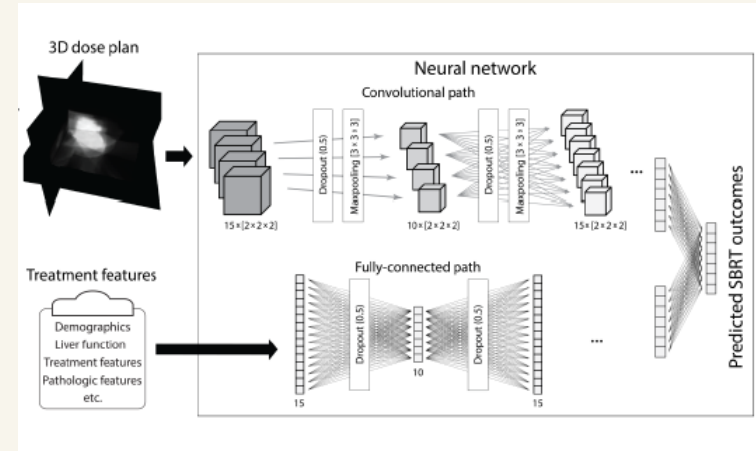


# Méthodes de Machine Learning

- **Données d'entrées** : indicateurs dose-volume
- **Algorithmes de Machine Learning**
  - **Random Forest** : ensemble d'arbres de décision indépendants en bagging
  - **XGBoost** : ensemble d'arbres de décision en boosting
  - **LightGBM** : inspiré de XGBoost, grandit « leaf-wise »

# Réseaux de neurones

- **Données d'entrées** : matrices de doses 3D + variables cliniques
- **Réseaux** :
  - **Réseau linéaire** : fully connected network
  - **Réseau convolutionnel** : CNN à 6 couches
  - **Réseau multi-chemin** : prend en entrée doses 3D et variables cliniques [2], voir schéma ci-contre
- **Loss** : Cross Entropy Loss Weighted
- **Optimizer** : Adam + L2-régularisation



**Fig** : Illustration du réseau multi-chemin [2]

# Métriques de performance

- **Balanced Accuracy** : notée BA, optimisée pour les datasets déséquilibrés.

$$BA = \frac{TPR + TNR}{2}, \text{ avec } TPR = \frac{TP}{TP+FN} \text{ et } TNR = \frac{TN}{TN+FP}$$

- **Recall** : aussi dit Rappel, capacité du modèle à identifier correctement les instances positives

$$Recall = \frac{TP}{TP + FN}$$

Un recall élevé est souvent souhaité dans des applications médicales (détection de maladie)

- **AUC of the ROC curve** : notée AUC, résume la performance d'un modèle de classification, entre spécificité et sensibilité
- **Matrice de confusion**

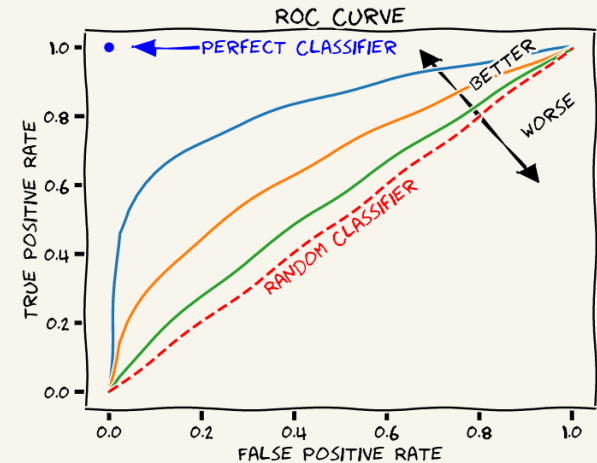


Figure : illustration d'une courbe ROC

# IV.



## Résultats

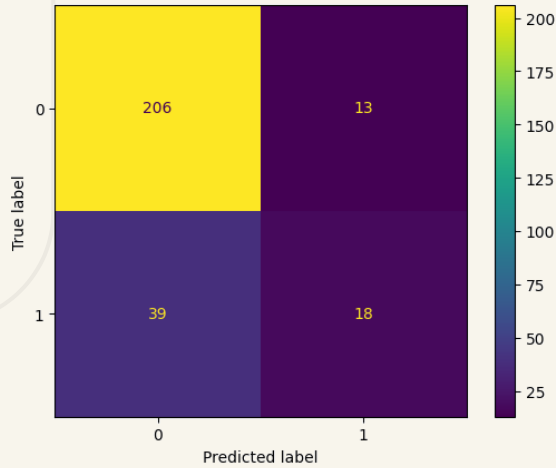
Machine Learning et Deep Learning

# Machine Learning – tous les résultats

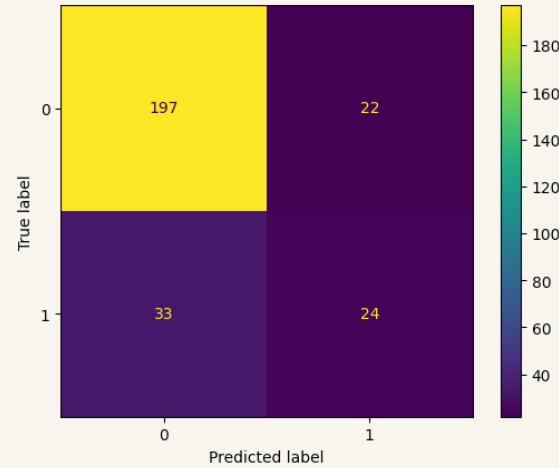
40 ans	Random Forest			XGBoost			LightGBM		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	65%	35%	<b>0,73</b>	<b>69%</b>	<b>52%</b>	0,71	68%	49%	0,72

50 ans	Random Forest			XGBoost			LightGBM		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	<b>81%</b>	<b>70%</b>	<b>0,88</b>	78%	70%	0,83	78%	70%	0,84

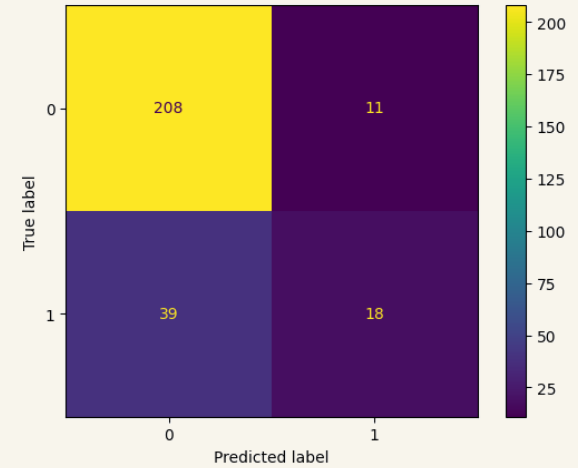
# Machine Learning – Matrices de confusion – 40 ans



Random Forest

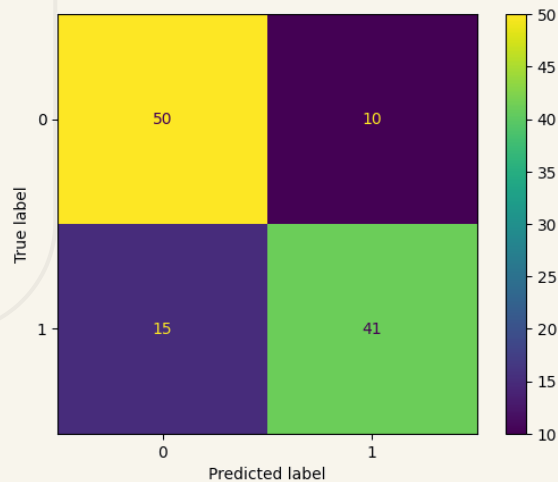


XBG Boost

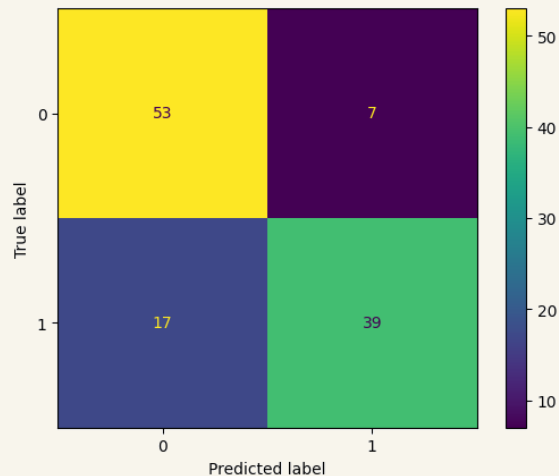


LGBM

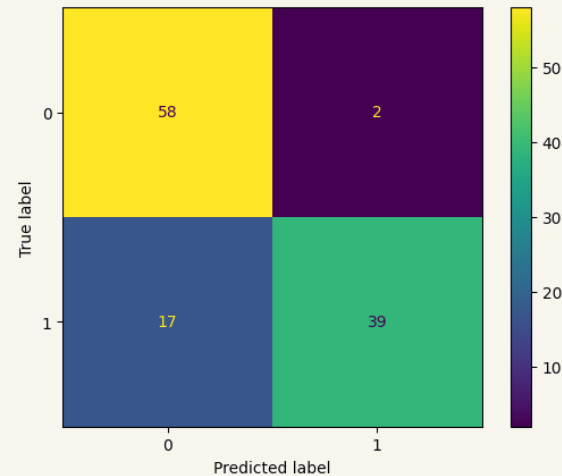
# Machine Learning – Matrices de confusion – 50 ans



Random Forest



XBG Boost



LGBM

# Deep Learning - tous les résultats

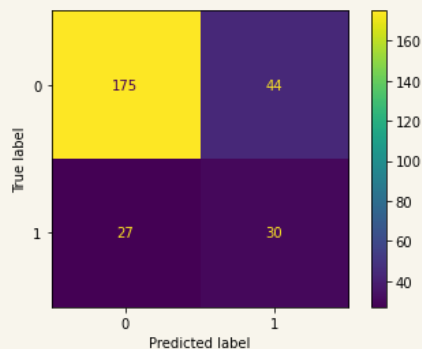
40 ans	Réseau Linéaire			Réseau Convolutionnel			Multipath Network		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	67%	53%	NA	69%	56%	NA	<b>71%</b>	<b>67%</b>	NA

50 ans	Réseau Linéaire			Réseau Convolutionnel			Multipath Network		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	80%	70%	NA	85%	73%	NA	<b>86%</b>	<b>79%</b>	NA

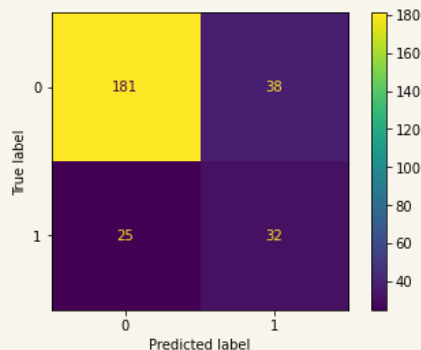


# Deep Learning – Temps de censure 40 ans

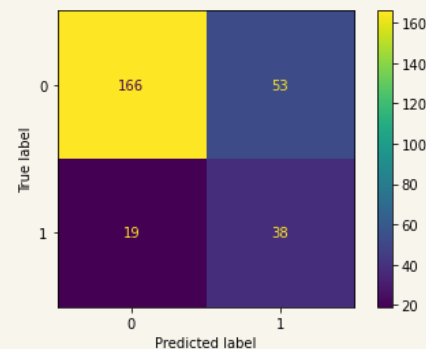
40 ans	Réseau Linéaire			Réseau Convolutionnel			Multipath Network		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	67%	53%	NA	69%	56%	NA	<b>71%</b>	<b>67%</b>	NA



Réseau Linéaire



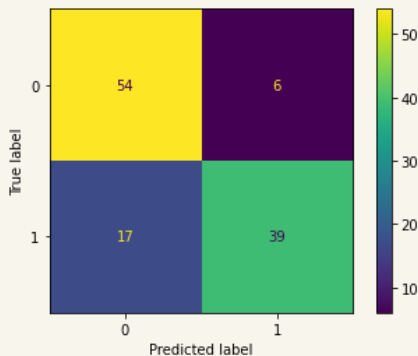
Réseau Convolutionnel



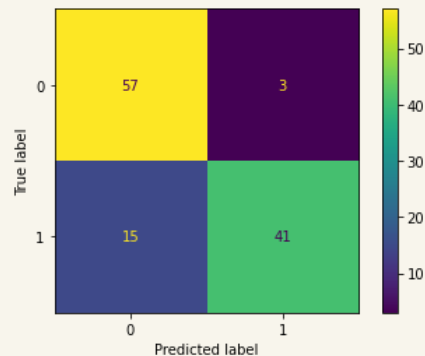
Multipath Network

# Deep Learning – Temps de censure 50 ans

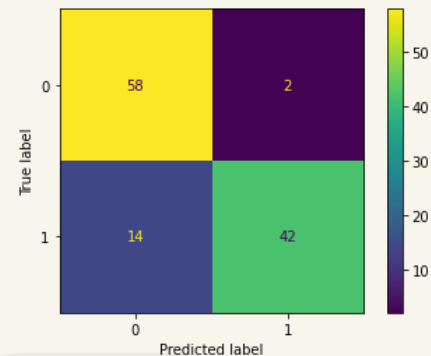
50 ans	Réseau Linéaire			Réseau Convolutionnel			Multipath Network		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	80%	70%	NA	85%	73%	NA	<b>86%</b>	<b>79%</b>	NA



Réseau Linéaire

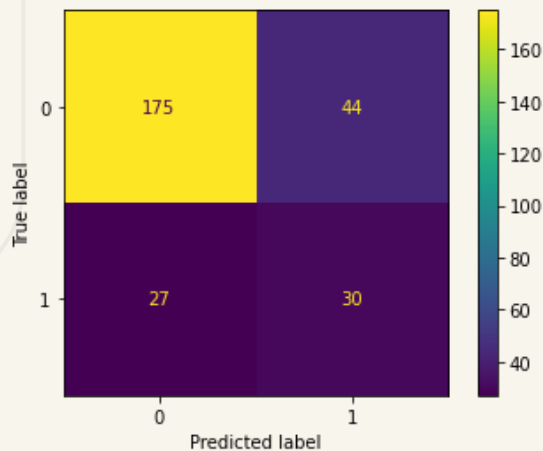


Réseau Convolutionnel

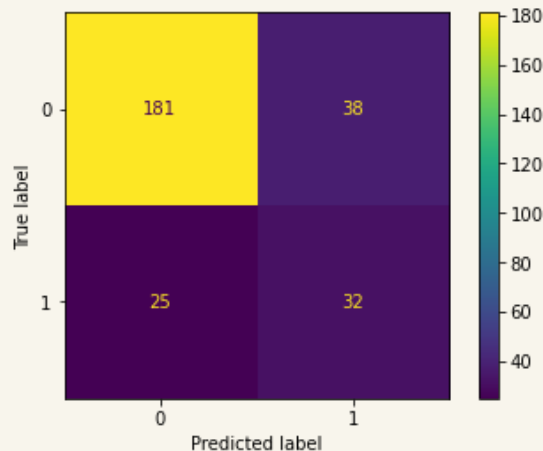


Multipath Network

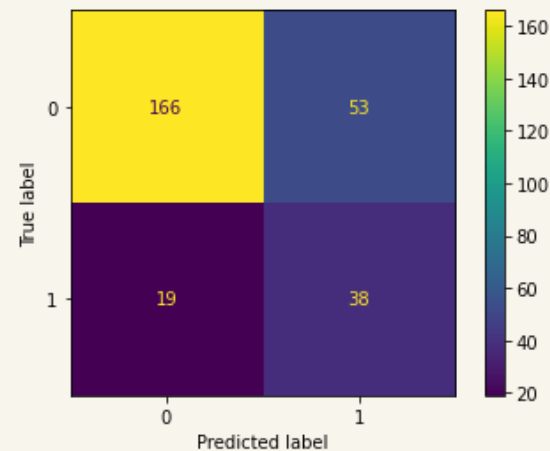
# Deep Learning – Matrices de confusion – 40 ans



Réseau Linéaire

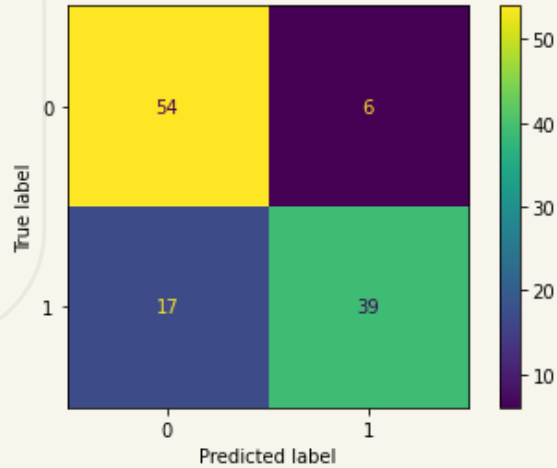


Réseau Convolutionnel

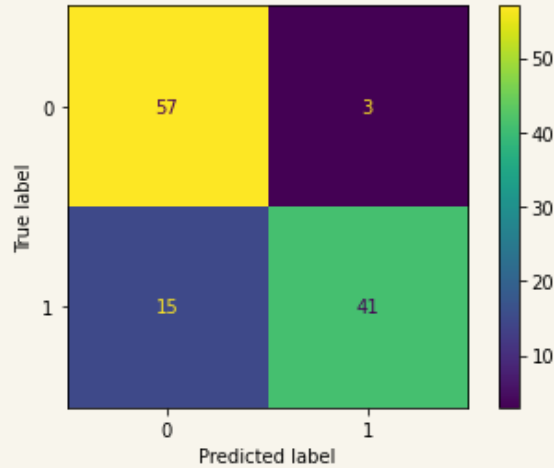


Multipath Network

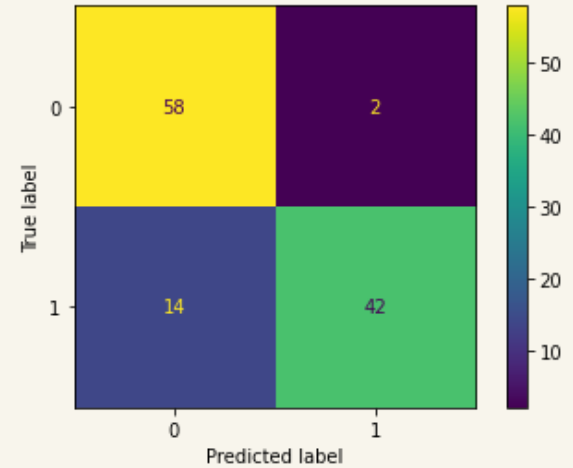
# Deep Learning – Matrices de confusion – 50 ans



Réseau Linéaire



Réseau Convolutionnel



Multipath Network

# Conclusion sur les résultats

- **Machine Learning :**

- Random Forest donne les meilleurs résultats pour le validation set
- Manque de robustesse du dataset de test

- **Deep Learning :**

- **Multipath Network** donne de meilleurs résultats sur le validation set
- **Perte de performance** sur le test set qui questionne

- **Comparaison des deux méthodes**

- La différence en Balanced Accuracy est légère (2-3%), mais différence marquée pour le score de rappel (~10%), à l'avantage des réseaux de neurones
- Meilleure prédiction des individus positifs : primordial dans le médical !
- Les résultats pour  $T_{censure} > 50$  sont globalement meilleurs que pour  $T_{censure} > 40$  car dataset plus équilibré

# Perspectives

## Quelques pistes pour de futurs travaux :

- Comprendre la différence entre le test set et le validation set : étudier le profil des patients mal classés
- Rendre les réseaux de neurones plus généralisables : cross-validation, plus de régularisation, ...
- Inclure plus de variables cliniques : normaliser les données, peut-être faire de la sélection de variables
- Proposer des réseaux de neurones plus profond, avec ou sans transfer learning
- Améliorer la prise en compte du caractère déséquilibré



CentraleSupélec



# Merci

Avez-vous des questions ?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik** and illustrations by **Storyset**



# Bibliographie

- AL Appelt, B Elhaminia, A Gooya, A Gilbert, and M Nix. Deep learning for radiotherapy outcome prediction using dose data—a review – 2022.
- Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote : synthetic minority over-sampling technique – 2002.
- Issam El Naqa and Martin J Murphy. Machine and Deep Learning in Oncology, Medical Physics and Radiology – 1991.
- Bulat Ibragimov, Diego AS Toesca, Yixuan Yuan, Albert C Koong, Daniel T Chang, and Lei Xing. Neural networks for deep radiotherapy dose analysis and prediction of liver sbrt outcomes – 2019.
- David G Kleinbaum, Mitchel Klein, David G Kleinbaum, and Mitchel Klein. Kaplan-meier survival curves and the log-rank test. Survival analysis : a self-learning text – 2012.
- CAO, Kaidi, WEI, Colin, GAIDON, Adrien, et al. Learning imbalanced datasets with label-distribution-aware margin loss. Advances in neural information processing systems, 2019, vol. 32.



# ANNEXES

## 1. Résultats

## Annexe 1 – Résultats test set ML

40 ans	Random Forest			XGBoost			LightGBM		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	65%	35%	<b>0,73</b>	<b>69%</b>	<b>52%</b>	0,71	68%	49%	0,72
Test set	<b>68%</b>	<b>88%</b>	<b>0,76</b>	63%	86%	0,75	62%	86%	0,73

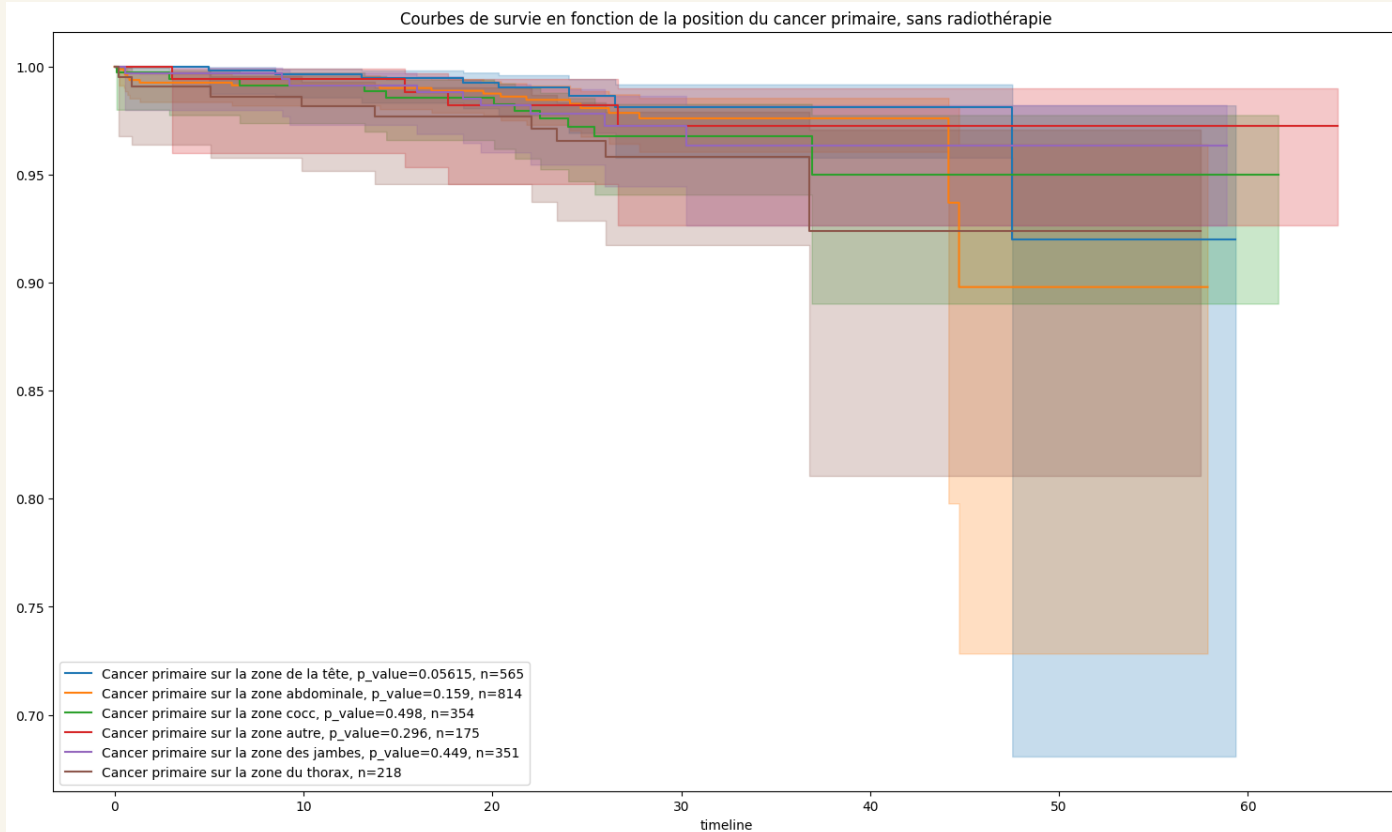
50 ans	Random Forest			XGBoost			LightGBM		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	<b>81%</b>	<b>70%</b>	<b>0,88</b>	78%	70%	0,83	78%	70%	0,84
Test set	54%	42%	0,54	<b>57%</b>	<b>47%</b>	<b>0,54</b>	56%	47%	0,54

## Annexe 2 – Résultats test set DL

40 ans	Réseau Linéaire			Réseau Convolutionnel			Multipath Network		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	67%	53%	NA	69%	56%	NA	<b>71%</b>	<b>67%</b>	NA
Test set	<b>66%</b>	<b>54%</b>	<b>0.69</b>	65%	43%	0.56	54%	52%	0.50

50 ans	Réseau Linéaire			Réseau Convolutionnel			Multipath Network		
	BA	Rappel	AUC	BA	Rappel	AUC	BA	Rappel	AUC
Validation set	80%	70%	NA	85%	73%	NA	<b>86%</b>	<b>79%</b>	NA
Test set	68%	58%	<b>0.72</b>	<b>76%</b>	<b>67%</b>	0.58	60%	65%	0.55

## Annexe 3 – Analyse de survie



## Annexe 4 – Analyse en composantes principales

