République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université du 20 Août 1955 – Skikda




Faculté des Sciences

Département d'Informatique

## Mémoire de fin d'études

Présenté en vue de l'obtention du diplôme de Master en Informatique

**Spécialité : Intelligence Artificielle (IA)**

---

# Prédiction des maladies cardiovasculaires à l'aide du Deep Learning

Étude comparative sur données cliniques tabulaires

---

**Réalisé par :**
Lounissi BOUDIS
Yousra SOLTANI

**Encadré par :**
Seloua HADIBY

Année universitaire : 2025 – 2026

*Page laissée intentionnellement blanche.*

# Dédicace

*Au nom d'Allah, le Tout Miséricordieux, le Très Miséricordieux.*

*À celle dont les prières ont été le secret de ma réussite et dont la lumière a éclairé mon chemin,*

*à ma très chère mère, je dédie le fruit de mes efforts ; c'est à toi que revient tout le mérite après Allah.*

*À mes chères sœurs, Rahma et Abir, pour votre soutien constant et votre présence précieuse.*

*À mes frères, Lamine et Mehdi, dont je suis fier, merci d'avoir toujours été une force et un appui pour moi.*

*À la joie de notre famille et à l'innocence incarnée,*

*à mes petits : Nouh, Yaqine, Adam et Daoud, je dédie cette réussite.*

*À mes chers amis : Ahmed, Yaacoub, Raouf, Mohamed et Abderrahmane, merci pour votre amitié sincère et votre soutien.*

*À Abdelbasset, qui a été pour moi comme un enseignant et un conseiller, je lui exprime ma profonde gratitude et mon respect.*

*À tous ceux qui ont partagé ma joie, de près ou de loin, je vous adresse mes sincères remerciements.*

*À ma promotrice, Docteure Seloua HADIBY, pour son encadrement, ses précieux conseils et son soutien constant, je lui exprime ma profonde reconnaissance.*

*Je dédie ce modeste travail à vous tous, en espérant qu'il soit le début d'un parcours plein de réussite et de succès.*

**Lounissi**

# Dédicace

*Avec tout mon respect et mon amour, je dédie ce travail à :*

*À mes parents bien-aimés, pour leur soutien inconditionnel et dont le seul objectif dans la vie est le bonheur et la réussite de leurs enfants. Je ne les remercierai jamais assez pour tous leurs sacrifices, leur amour et leurs encouragements.*

*À mes frères : Walid, Mokhtar et Bilal, pour leur présence, leur affection et leurs encouragements tout au long de mon parcours.*

*À mes sœurs : Amel, Halima, Seham et Chafia, pour leur soutien, leur tendresse et leurs précieux conseils.*

*À toutes mes amies, notamment : Rayan Achouak et Nour El Houda, en souvenir des plus beaux moments que nous avons partagés ensemble et pour leur amitié sincère.*

*Enfin, à toutes les personnes qui m'ont donné la force de continuer et à tous ceux qui m'ont soutenue, même par un simple mot d'encouragement.*

*Merci à tous.*

**Yousra**

# Remerciements

Nous tenons à exprimer notre profonde gratitude à notre encadrante, **Dr. Seloua HADIBY**, pour sa disponibilité, ses orientations méthodologiques et ses remarques constructives tout au long de la réalisation de ce mémoire.

Nous remercions également l'ensemble des enseignants du Département d'Informatique pour la qualité de leur formation, ainsi que toutes les personnes qui nous ont apporté leur aide, leurs conseils ou leur soutien durant ce travail.

Enfin, nous adressons nos remerciements les plus sincères à nos familles et à nos amis pour leur patience, leur confiance et leurs encouragements permanents.

# ملخص

تُعد الأمراض القلبية الوعائية من أهم أسباب المراضة والوفيات عالميا، مما يجعل التنبؤ المبكر بالمخاطر أداة مهمة لدعم الوقاية والقرار الطبي، ولا سيما عندما تكون النماذج المستعملة قابلة للفهم والتفسير من منظور سريري.

تكمن إشكالية هذا العمل في تقييم القيمة العملية لنماذج التعلم العميق مقارنة بالنماذج الكلاسيكية عند التعامل مع معطيات سريرية جدولية محدودة وغير متوازنة، مع مراعاة الحساسية، جودة التمييز، وقابلية تفسير التنبؤات.

تقترح هذه الدراسة إطارا مقارنا وقائما على قابلية التفسير، يجمع بين نماذج كلاسيكية ونماذج تعلم عميق ملائمة للبيانات الجدولية. ولا يقتصر هذا الإطار على قياس الأداء التنبئي، بل يربط النتائج بمؤشرات ذات معنى طبي وبآليات تفسير تساعد على فهم العوامل الأكثر تأثيرا في القرار النموذجي.

أظهرت النتائج أن الشبكة العصبية متعددة الطبقات (MLP) بلغت أفضل قدرة تمييزية إلى جانب الانحدار اللوجستي، مع أفضل توازن عام بين الفئات، بينما قدم ResNetTab أعلى حساسية، وهي خاصية مهمة في سياقات الكشف المبكر. وتبرز هذه النتائج أن نماذج التعلم العميق توفر بدائل واعدة ومتكاملة عندما تكون أولوية التطبيق هي الحساسية، توازن الأداء، وقابلية التفسير، خاصة مع إمكانية تعزيزها مستقبلا ببيانات أكثر غنى.

**الكلمات المفتاحية :** الأمراض القلبية الوعائية؛ التنبؤ؛ التعلم الآلي؛ التعلم العميق؛ البيانات السريرية الجدولية؛ الذكاء الاصطناعي القابل للتفسير.

# Résumé

Les maladies cardiovasculaires constituent une cause majeure de morbidité et de mortalité, ce qui rend la prédiction précoce du risque particulièrement utile pour la prévention, l'orientation du dépistage et l'aide à la décision médicale explicable.

La problématique étudiée est d'évaluer la valeur pratique des modèles d'apprentissage profond par rapport aux modèles classiques pour des données cliniques tabulaires limitées et déséquilibrées, tout en tenant compte de la sensibilité, de la discrimination et de l'interprétabilité.

La solution proposée repose sur un cadre comparatif et reproductible associant des modèles classiques et des architectures profondes adaptées aux données tabulaires. L'évaluation ne se limite pas à la performance prédictive : elle intègre également la sensibilité, l'équilibre entre les classes et une analyse explicable des prédictions afin de relier les sorties des modèles aux facteurs de risque cliniquement pertinents.

Les résultats montrent que le Perceptron Multicouche (MLP) atteint la meilleure discrimination au même niveau que la régression logistique, tout en obtenant le meilleur équilibre global entre les classes, tandis que ResNetTab fournit la sensibilité la plus élevée, critère important pour le dépistage. Ces résultats positionnent les modèles profonds comme des alternatives prometteuses et complémentaires lorsque la priorité porte sur la sensibilité, l'équilibre des performances et l'analyse des prédictions, avec un potentiel renforcé par des données plus riches.

**Mots-clés :** maladies cardiovasculaires ; prédiction ; apprentissage automatique ; apprentissage profond ; données cliniques tabulaires ; intelligence artificielle explicable.

# Abstract

Cardiovascular diseases are a major cause of morbidity and mortality, making early risk prediction important for prevention, screening prioritization, and explainable medical decision support.

This work evaluates the practical value of deep learning models compared with classical approaches for limited and imbalanced tabular clinical data, while considering sensitivity, discrimination, and interpretability.

The proposed solution adopts a reproducible comparative framework that brings together classical models and deep architectures suited to tabular clinical data. Beyond predictive performance, the evaluation emphasizes sensitivity, class-balance behaviour, and explainable prediction analysis, linking model outputs to clinically meaningful risk factors and supporting more transparent decision support.

The results show that Multi-Layer Perceptron (MLP) reaches the strongest discrimination at the same level as logistic regression while providing the best overall class balance, whereas ResNetTab achieves the highest sensitivity, an important criterion for screening. These findings position deep learning models as promising and complementary alternatives when sensitivity, balanced performance, and prediction analysis are prioritized, with further potential as richer clinical data become available.

**Keywords:** cardiovascular disease; prediction; machine learning; deep learning; tabular clinical data; explainable artificial intelligence.

# Table des matières

**Dédicace de Lounissi BOUDIS** i

**Dédicace de Yousra SOLTANI** ii

**Remerciements** iii

**Résumé en arabe** iv

**Résumé** v

**Abstract** vi

**Liste des figures** x

**Liste des tableaux** xiii

**Liste des abréviations** xiv

**Introduction générale** 1

**1 Généralités sur les maladies cardiaques** 5

1.1 Introduction . . . 5
1.2 Structure et rôle du cœur . . . 6
1.3 Fonctionnement du cœur . . . 7
1.4 Les maladies cardiaques . . . 8
1.5 Diagnostique des maladies cardiaques . . . 11
1.6 Les principaux symptômes des maladies cardiaques . . . 12
1.7 Facteurs de risque des maladies cardiaques . . . 13
1.7.1 Facteurs de risque non modifiables . . . 13
1.7.2 Facteurs de risque modifiables . . . 14
1.8 Prévention des maladies cardiaques . . . 15
1.8.1 Prévention individuelle . . . 16
1.8.2 Prévention collective . . . 16
1.9 Extraction et sélection des caractéristiques . . . 16

1.10 Conclusion . . . . . 17

**2 Apprentissage automatique : concepts et méthodes 18**
2.1 Introduction . . . . . 18
2.2 Apprentissage supervisé et non supervisé . . . . . 19
2.2.1 L'apprentissage supervisé . . . . . 19
2.2.2 Types de l'apprentissage supervisé . . . . . 21
2.2.3 Algorithmes de classification . . . . . 22
2.2.4 Algorithmes de régression . . . . . 26
2.2.5 Comparaison entre classification et régression . . . . . 30
2.2.6 L'apprentissage non supervisé . . . . . 31
2.2.7 Types de l'apprentissage non supervisé . . . . . 32
2.2.8 Algorithmes de clustering . . . . . 33
2.2.9 Algorithmes de réduction de dimension . . . . . 35
2.3 Méthodes de Deep Learning . . . . . 37
2.3.1 Réseaux de neurones artificiels (ANN) . . . . . 37
2.3.2 Réseaux de neurones convolutifs (CNN) . . . . . 38
2.3.3 Réseaux de neurones récurrents (RNN) . . . . . 39
2.3.4 Long Short-Term Memory (LSTM) . . . . . 40
2.3.5 Autoencoders . . . . . 40
2.3.6 Transformers . . . . . 41
2.3.7 Comparaison des méthodes de Deep Learning . . . . . 42
2.3.8 Fonction de perte (Loss function) . . . . . 43
2.4 Interprétabilité et intelligence artificielle explicable . . . . . 46
2.5 Évaluation des modèles . . . . . 48
2.5.1 Matrice de confusion . . . . . 48
2.5.2 Accuracy (Exactitude) . . . . . 48
2.5.3 Précision et Recall (Sensibilité) . . . . . 48
2.5.4 F1-score . . . . . 49
2.5.5 Spécificité . . . . . 49
2.5.6 Courbe ROC et AUC . . . . . 49
2.6 Conclusion . . . . . 50

**3 Méthodologie, Expérimentation et Résultats 52**
3.1 Introduction . . . . . 52
3.2 Méthodologie expérimentale . . . . . 53
3.2.1 Description du jeu de données (dataset) . . . . . 53
3.2.2 Prétraitement des données . . . . . 55
3.2.3 Modèles de référence . . . . . 56
3.2.4 Architectures d'apprentissage profond . . . . . 58

3.2.5 Stratégie d'entraînement . . . . . 62
3.2.6 Optimisation des hyperparamètres . . . . . 64
3.2.7 Seuils de décision . . . . . 64
3.2.8 Métriques d'évaluation . . . . . 65
3.2.9 Méthodes d'interprétabilité . . . . . 67
3.2.10 Configuration expérimentale . . . . . 67
3.3 Résultats et analyse . . . . . 68
3.3.1 Comparaison globale des performances . . . . . 68
3.3.2 Analyse des matrices de confusion . . . . . 70
3.3.3 Courbes ROC et précision-rappel . . . . . 71
3.3.4 Dynamique d'entraînement . . . . . 72
3.3.5 Calibration des probabilités . . . . . 74
3.3.6 Interprétabilité et importance des variables . . . . . 75
3.4 Discussion . . . . . 76
3.5 Prototype d'application interactive . . . . . 77
3.6 Conclusion . . . . . 82

**Conclusion générale et perspectives** **83**

**Références bibliographiques** **86**

# Liste des figures


1.1 Structure anatomique du cœur . . . . . 6
1.2 Fonctionnement du cœur . . . . . 7
1.3 Infarctus-du-myocarde . . . . . 8
1.4 Accident vasculaire cérébral . . . . . 9
1.5 Athérosclérose . . . . . 9
1.6 Angine de poitrine . . . . . 10
1.7 La différence entre les deux cœurs . . . . . 10
1.8 Arythmie cardiaque . . . . . 11
1.9 Hypertension artérielle . . . . . 11
2.1 Workflow de l'apprentissage supervisé appliqué aux données médicales 20
2.2 Schéma du processus de rétropropagation . . . . . 21
2.3 Frontière de décision et fonction sigmoïde en régression logistique . 23
2.4 Hyperplan optimal et marge maximale en SVM . . . . . 24
2.5 Structure d'un arbre de décision . . . . . 25
2.6 Classification par les K plus proches voisins . . . . . 26
2.7 Droite de régression linéaire par moindres carrés . . . . . 27
2.8 Régression SVR avec tube $\varepsilon$ . . . . . 28
2.9 Arbre de régression et prédictions par segments . . . . . 29
2.10 Forêt aléatoire (Random Forest) et moyenne des prédictions . . . . . 30
2.11 Processus de l'apprentissage non supervisé sur données brutes non étiquetées . . . . . 32
2.12 Regroupement des données par K-Means . . . . . 33
2.13 Dendrogramme du clustering hiérarchique . . . . . 34
2.14 Clustering basé sur la densité (DBSCAN) . . . . . 34
2.15 Projection des données avec PCA . . . . . 35

2.16 Visualisation des données avec t-SNE . . . . . 36
2.17 Architecture d'un autoencoder . . . . . 36
2.18 Réseau de neurones artificiel . . . . . 38
2.19 Architecture d'un CNN . . . . . 39
2.20 Réseau de neurones récurrent . . . . . 39
2.21 Cellule LSTM . . . . . 40
2.22 Autoencoder . . . . . 41
2.23 Architecture Transformer . . . . . 42
2.24 Fonction de perte . . . . . 43
2.25 Courbe ROC . . . . . 50
3.1 Schéma de la régression logistique utilisée comme modèle de référence 57
3.2 Schéma de fonctionnement du modèle XGBoost utilisé . . . . . 58
3.3 Architecture utilisée du perceptron multicouche (MLP) . . . . . 59
3.4 Architecture retenue du réseau résiduel tabulaire ResNetTab . . . . . 60
3.5 Architecture adoptée de TabNet . . . . . 62
3.6 Comparaison globale des performances des modèles sur l'ensemble de test . . . . . 69
3.7 Matrices de confusion des modèles de référence sur l'ensemble de test 70
3.8 Matrices de confusion des architectures d'apprentissage profond sur l'ensemble de test . . . . . 70
3.9 Courbes ROC et précision-rappel des cinq modèles sur l'ensemble de test . . . . . 71
3.10 Courbes de perte d'entraînement et de validation pour le MLP et ResNetTab . . . . . 72
3.11 Convergence de la recherche d'hyperparamètres Optuna pour les modèles d'apprentissage profond . . . . . 73
3.12 Courbes de calibration des probabilités sur l'ensemble de test . . . . 74
3.13 Importance des variables pour les modèles d'apprentissage profond sur l'ensemble de test . . . . . 75
3.14 Vue générale initiale de l'application avec le modèle MLP sélectionné par défaut, illustrant la prédiction effectuée pour un échantillon de patient. L'application permet également de changer de langue (arabe, français, anglais) et de basculer entre les thèmes clair et sombre. . . . 78

3.15 Onglet des métriques du modèle MLP, incluant les indicateurs de performance, la matrice de confusion et le rapport de classification . . . 78
3.16 Prédiction d'un échantillon patient avec le modèle ResNetTab, affichant la probabilité estimée et le niveau de risque associé . . . . . 79
3.17 Prédiction d'un échantillon patient avec le modèle TabNet, avec estimation du risque et synthèse de décision . . . . . 79
3.18 Comparaison globale des modèles profonds à l'aide de graphiques radar et de diagrammes en barres des métriques de performance . . 80
3.19 Prédiction issue d'une saisie manuelle des caractéristiques cliniques du patient . . . . . 80
3.20 Interprétabilité du modèle MLP par importance locale des variables (avec SHAP) . . . . . 81
3.21 Interprétabilité du modèle TabNet à partir des importances de variables issues du mécanisme d'attention . . . . . 81

# Liste des tableaux


2.1 Comparaison entre classification et régression . . . . . 30
2.2 Comparaison entre clustering et réduction de dimension . . . . . 37
2.3 Comparaison globale entre apprentissage supervisé et non supervisé 37
2.4 Comparaison des méthodes de Deep Learning . . . . . 42
2.5 Principales familles de méthodes d'interprétabilité et d'explicabilité . 47
3.1 Caractéristiques générales du jeu de données . . . . . 54
3.2 Description des variables du jeu de données . . . . . 54
3.3 Performance des modèles sur l'ensemble de test. Les deux premières lignes correspondent aux modèles de référence (baselines), placés en haut du tableau. Les meilleurs scores de chaque métrique sont indiqués en gras. . . . . 69

# Liste des abréviations

| | |
|---|---|
| **ANN** | Artificial Neural Network |
| **AUC** | Area Under the Curve |
| **AVC** | Accident Vasculaire Cérébral |
| **BoW** | Bag of Words |
| **CNN** | Convolutional Neural Network |
| **DBSCAN** | Density-Based Spatial Clustering of Applications with Noise |
| **ECG** | Electrocardiogram |
| **FN** | False Negative |
| **FP** | False Positive |
| **FPR** | False Positive Rate |
| **GANs** | Generative Adversarial Networks |
| **HbA1c** | Hemoglobin A1c |
| **HDL** | High-Density Lipoprotein |
| **HTA** | Hypertension Artérielle |
| **IDM** | Infarctus Du Myocarde |
| **IRM** | Imagerie par Résonance Magnétique |
| **IV** | Intravenous |
| **KNN** | K-Nearest Neighbors |
| **L1** | L1 Regularization |
| **L2** | L2 Regularization |
| **LDL** | Low-Density Lipoprotein |
| **LSTM** | Long Short-Term Memory |
| **MAE** | Mean Absolute Error |
| **MSE** | Mean Squared Error |
| **PAD** | Pression Artérielle Diastolique |
| **PAS** | Pression Artérielle Systolique |
| **PCA** | Principal Component Analysis |
| **RMSE** | Root Mean Squared Error |
| **RNN** | Recurrent Neural Network |
| **ROC** | Receiver Operating Characteristic |
| **SHAP** | SHapley Additive exPlanations |
| **SVM** | Support Vector Machine |
| **SVR** | Support Vector Regression |
| **TDM** | Tomodensitométrie |
| **TEP** | Tomographie par Émission de Positons |
| **TF-IDF** | Term Frequency–Inverse Document Frequency |
| **t-SNE** | t-distributed Stochastic Neighbor Embedding |
| **TN** | True Negative |
| **TP** | True Positive |
| **TPR** | True Positive Rate |
| **XAI** | Explainable Artificial Intelligence |

# Introduction générale

## Contexte

Les maladies cardiovasculaires constituent un problème majeur de santé publique. Elles regroupent plusieurs atteintes du cœur et des vaisseaux sanguins, notamment les cardiopathies coronariennes, les accidents vasculaires cérébraux, les maladies vasculaires périphériques et certaines affections cardiaques structurelles. Selon l'Organisation mondiale de la Santé, elles restent la première cause de mortalité dans le monde, avec environ 19,8 millions de décès en 2022, soit près d'un tiers de l'ensemble des décès mondiaux (World Health Organization 2025). Cette charge sanitaire justifie le développement d'outils capables d'améliorer la prévention, le dépistage précoce et la stratification du risque.

La prédiction du risque cardiovasculaire repose historiquement sur des facteurs cliniques bien établis, tels que l'âge, le sexe, la pression artérielle, le cholestérol, le tabagisme et le diabète. Les travaux issus de l'étude de Framingham ont joué un rôle central dans la construction des scores de risque cardiovasculaire utilisés en prévention primaire (D'Agostino et al. 2008). Ces modèles ont l'avantage d'être simples, interprétables et faciles à appliquer. Toutefois, ils reposent souvent sur des hypothèses de linéarité et sur un nombre limité de variables, ce qui peut limiter leur capacité à représenter certaines relations non linéaires ou interactions complexes entre facteurs de risque.

Dans ce contexte, l'apprentissage automatique offre une voie complémentaire. Les algorithmes de machine learning peuvent apprendre des relations à partir des données, comparer plusieurs familles de modèles et exploiter des mesures de performance adaptées à des tâches de classification médicale. En médecine, leur intérêt est particulièrement important lorsque les données sont structurées, hétérogènes et issues de mesures cliniques de routine (Rajkomar, Dean et Kohane 2019). Des études ont montré que certaines méthodes d'apprentissage automatique peuvent améliorer la prédiction du risque cardiovasculaire par rapport à des approches statistiques classiques, tout en posant des questions importantes de validation, de calibration et

d'interprétabilité (Weng et al. 2017).

Au-delà de la seule performance, l'intégration de tels modèles dans un contexte médical exige une lecture transparente des prédictions. L'intelligence artificielle explicable (XAI) constitue à cet égard un enjeu central : elle vise à rendre les décisions algorithmiques plus auditables, à identifier les variables qui orientent le risque estimé et à vérifier la cohérence des résultats avec les connaissances cliniques. Cette dimension ne remplace pas l'expertise médicale, mais elle renforce la confiance, facilite l'analyse des erreurs et soutient une utilisation plus responsable des modèles prédictifs.

## Problématique

La prédiction automatique des maladies cardiovasculaires à partir de données cliniques tabulaires est une tâche utile, mais difficile. D'une part, les données disponibles sont souvent de taille modérée, avec un nombre limité de variables et une classe positive minoritaire. D'autre part, l'objectif clinique ne se limite pas à maximiser une mesure globale comme l'AUROC. Il faut aussi tenir compte de la sensibilité, afin de réduire les cas positifs manqués, de la spécificité, afin de limiter les fausses alertes, et de la qualité des probabilités produites par les modèles.

L'apprentissage profond est devenu dominant dans plusieurs domaines comme l'image, le texte et la parole. Pour les données tabulaires cliniques, son intérêt doit être apprécié de manière plus fine qu'un simple classement global, car ces architectures peuvent apporter des profils utiles en sensibilité, équilibre des classes, représentation des interactions et interprétabilité. Des architectures spécialisées, comme TabNet, ont été proposées pour combiner performance et interprétabilité grâce à des mécanismes d'attention (Arik et Pfister 2021). D'autres travaux ont montré que des réseaux profonds bien réglés peuvent être compétitifs sur certains jeux de données tabulaires (Gorishniy et al. 2021). Les études comparatives indiquent aussi que les modèles classiques, notamment les arbres de décision en ensemble et les modèles de boosting, constituent des références solides sur des données tabulaires de taille faible ou moyenne (Chen et Guestrin 2016 ; Shwartz-Ziv et Armon 2022 ; Grinsztajn, Oyallon et Varoquaux 2022).

La problématique centrale de ce mémoire peut donc être formulée ainsi : comment les modèles d'apprentissage profond se positionnent-ils par rapport à des modèles classiques pour la prédiction des maladies cardiovasculaires à partir de données cliniques tabulaires, en tenant compte à la fois de la discrimination, du déséquilibre des classes, de la sensibilité et de l'explicabilité des prédictions ?

## Objectifs du mémoire

L'objectif général de ce travail est de concevoir et d'évaluer une chaîne expérimentale de prédiction des maladies cardiovasculaires fondée sur des données cliniques tabulaires. Plus précisément, ce mémoire vise à :

- présenter les notions médicales essentielles liées aux maladies cardiovasculaires et à leurs facteurs de risque ;
- rappeler les principes des méthodes d'apprentissage supervisé et des architectures profondes utiles pour la classification ;
- préparer un jeu de données clinique en traitant les valeurs aberrantes, la normalisation et le déséquilibre entre classes ;
- comparer des modèles classiques et profonds (deep learning) selon des métriques adaptées au contexte médical ;
- analyser les résultats obtenus en termes de performance, de sensibilité, de calibration et d'explicabilité des prédictions ;
- discuter les limites de l'étude et proposer des pistes d'amélioration pour de futurs travaux.

## Approche méthodologique

Le travail expérimental utilise le jeu de données composite *Heart_Disease_2024*. Associé au BIG-Dataset (Tajnim 2024 ; Haque et al. 2024), il comprend 4 238 observations et 14 variables cliniques et démographiques issues notamment de la cohorte de Framingham et du jeu de données de Cleveland. La variable cible indique la présence ou l'absence d'une maladie cardiovasculaire. Comme la classe positive est minoritaire, l'évaluation des modèles doit tenir compte du déséquilibre des données et ne pas se limiter à une seule mesure globale.

La méthodologie adoptée suit une chaîne expérimentale rigoureuse et reproductible : préparation du jeu de données, séparation contrôlée des ensembles d'apprentissage et de test, entraînement comparatif des modèles, puis évaluation finale sur des données indépendantes. Cette organisation permet de comparer les approches dans des conditions homogènes, de limiter le risque d'une évaluation trop optimiste et de conserver une lecture adaptée aux exigences de la recherche médicale.

Cinq modèles sont étudiés : la régression logistique, XGBoost, un perceptron multicouche (MLP), TabNet et une architecture résiduelle pour données tabulaires, appelée ResNetTab. L'évaluation repose sur des métriques complémentaires, notamment l'AUROC, l'AUPRC, la sensibilité, la spécificité et le F1 macro. L'analyse accorde également une place explicite à la calibration et à l'explicabilité des prédictions, en mobilisant des outils XAI pour interpréter les facteurs qui contribuent aux décisions des modèles et apprécier leur cohérence clinique.

## Organisation du mémoire

La suite du mémoire est structurée de manière progressive, depuis le contexte médical jusqu'à l'analyse expérimentale et aux perspectives finales :

- *Introduction générale* — présentation du contexte, de la problématique, des objectifs et de l'approche méthodologique.
- Chapitre **1** — *Généralités sur les maladies cardiaques* — notions médicales essentielles, principales formes, facteurs de risque et prévention.
- Chapitre **2** — *Apprentissage automatique : concepts et méthodes* — fondements de l'apprentissage automatique et architectures adaptées aux données tabulaires.
- Chapitre **3** — *Méthodologie, Expérimentation et Résultats* — protocole expérimental, modèles évalués, métriques retenues et analyse des performances.
- *Conclusion générale* — synthèse des contributions, réponse à la problématique, limites et perspectives de recherche.

CHAPITRE 1

# Généralités sur les maladies cardiaques

## 1.1 Introduction

Les maladies cardiaques représentent aujourd'hui l'une des principales causes de mortalité dans le monde. Elles regroupent un ensemble de troubles affectant le cœur et les vaisseaux sanguins, tels que les maladies coronariennes, l'insuffisance cardiaque, les troubles du rythme (arythmies) et les cardiopathies congénitales. Ces pathologies peuvent avoir des conséquences graves sur la santé humaine et nécessitent une prise en charge précoce et efficace. Au cours des dernières années, l'augmentation des facteurs de risque tels que l'hypertension artérielle, le diabète, le tabagisme, la sédentarité et l'obésité a contribué à la propagation des maladies cardiovasculaires. Par conséquent, la détection précoce et la prévention jouent un rôle essentiel dans la réduction du taux de mortalité et l'amélioration de la qualité de vie des patients. Dans ce contexte, les avancées technologiques, notamment dans le domaine de l'intelligence artificielle et de l'apprentissage profond (Deep Learning), offrent de nouvelles perspectives prometteuses pour la prédiction des maladies cardiaques. L'analyse des données médicales permet d'identifier des modèles complexes et de fournir des outils d'aide à la décision pour les professionnels de santé. Ce chapitre a pour objectif de présenter des généralités sur les maladies cardiaques, en abordant la structure et le fonctionnement du cœur, les différents types de maladies, leurs symptômes, les facteurs de risque, les méthodes de diagnostic, ainsi que les approches de prévention. Il introduit également la notion d'extraction et de sélection des caractéristiques, qui constitue une étape essentielle pour la mise en œuvre des modèles de prédiction dans les chapitres suivants.

## 1.2 Structure et rôle du cœur

Le cœur est un organe musculaire creux qui permet la circulation du sang dans le corps et l'apport d'oxygène et de nutriments à l'ensemble des cellules de l'organisme. Le cœur est situé dans la partie médiane de la cage thoracique (le médiastin) délimité par les 2 poumons, le sternum et la colonne vertébrale, comme le montre la figure 1.1. Il est un peu plus gros chez l'homme que chez la femme et pompe chaque jour en moyenne 8000 litres de sang grâce à environ 100 000 battements quotidiens (soit jusqu'à 2 milliards de battements au cours de la vie). Le cœur est composé de 4 cavités : les oreillettes (ou atria) sur la partie supérieure et les ventricules sur la partie inférieure. Les oreillettes et ventricules sont séparés de chaque côté par une épaisse paroi musculaire, le septum. Ainsi, aucun échange de sang entre la partie supérieur et la partie inférieure n'est possible. Le passage est unidirectionnel entre oreillette vers ventricule et cela via les valves cardiaques. La paroi du cœur est composée de muscle en trois couches distinctes : l'épicarde (cellules épithéliales et tissu conjonctif), le myocarde ou muscle cardiaque et à l'intérieur, l'endocarde (cellules épithéliales et tissu conjonctif). Le sang qui circule dans le cœur va trop vite pour y être absorbé, si bien qu'il dispose de son propre système de vaisseaux, appelé artères coronaires, le vascularisant pour apporter aux cellules cardiaques oxygène et nutriments. Ce sont les ventricules qui assurent la fonction de pompes du sang vers le corps ou vers les poumons et leurs parois sont plus épaisses et leurs contractions plus fortes que les oreillettes (Gaut 2026).

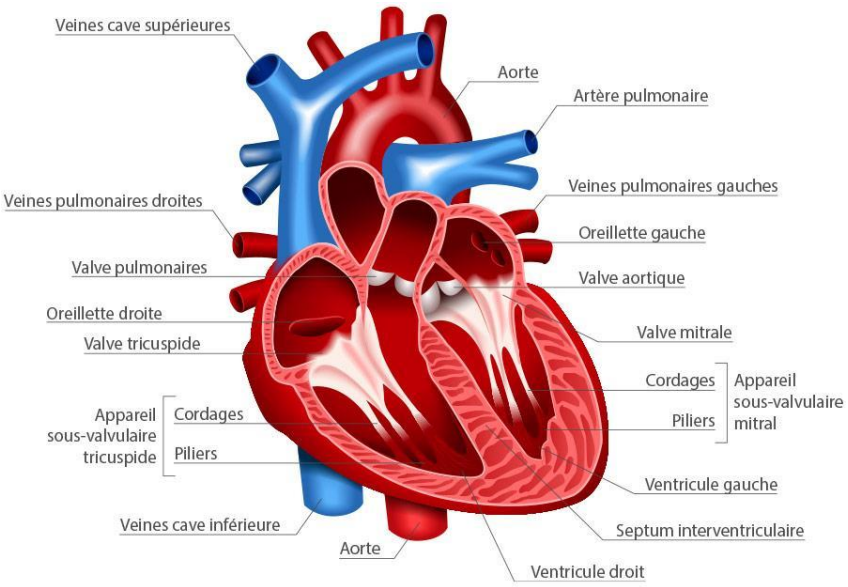



**Fig 1.1 :** Structure anatomique du cœur (Santé sur le Net 2026)

## 1.3 Fonctionnement du cœur

La figure 1.2 schématise le fonctionnement du cœur.

1. Le sang pauvre en oxygène arrive dans l'oreillette droite par les veines caves ;
2. L'oreillette droite se contracte et le sang est expulsé dans le ventricule droit par la valve tricuspide ;
3. Le ventricule droit se contracte et expulse le sang dans les artères pulmonaires par la valve pulmonaire ;
4. Ce sang va aller directement vers les poumons pour se charger en oxygène ;
5. Le sang oxygéné arrive dans l'oreillette gauche par les veines pulmonaires ;
6. La contraction de l'oreillette gauche permet le passage du sang vers le ventricule gauche par la valve mitrale ;
7. Le ventricule gauche se contracte à son tour pour expulser le sang dans l'aorte par la valve aortique ;
8. Le sang riche en oxygène peut alors être distribué à tous les organes, muscles et tissus de l'organisme.

Ce cycle recommence environ 60 fois par minute au repos (Gaut 2026).

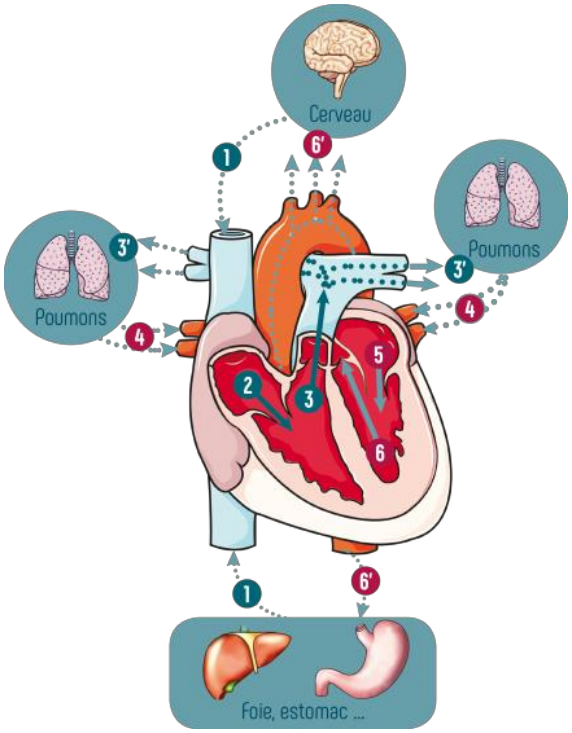



**Fig 1.2 :** Fonctionnement du cœur (Centre hospitalier universitaire de Rouen 2026)

## 1.4 Les maladies cardiaques

Les maladies cardiques englobent plusieurs types de troubles de l'appareil circulatoire, la pathologie du coeur et les vaissaux sanguins, soit les maladies congénitales, ischémiques ou coronariennes, les maladies cérébrovasculaires et vasculaires périphériques ainsi que l'insuffisance et le rhumatisme cardiaque. La maladie ischémique ou coronarienne est la maladie cardiaque la plus répandue. Elle touche les vaisseaux sanguins qui irriguent le muscle cardiaque. La maladie cérébrovasculaire est un problème au niveau de la circulation sanguine dans les vaisseaux du cerveau alors que la maladie vasculaire périphérique affecte principalement les vaisseaux qui alimentent les bras et les jambes. L'insuffisance cardiaque survient lorsque le coeur ne pompe pas suffisamment de sang pour atteindre le niveau de circulation sanguine nécessaire aux besoins énergétiques du corps. Le rhumatisme cardiaque est une maladie infectieuse qui affecte les articulations et les valvules cardiaques alors que la maladie congénitale est une malformation du cœur qui découle d'une anomalie présente à la naissance.

- **Infarctus du myocarde (IDM) :** L'infarctus du myocarde est une nécrose du myocarde se manifestant lorsqu'une ou plusieurs artères coronaires s'obstruent. De ce fait une partie du cœur n'est plus approvisionnée en sang et en oxygène (Ornstein et al. 2004). La figure 1.3 illustre cette pathologie.

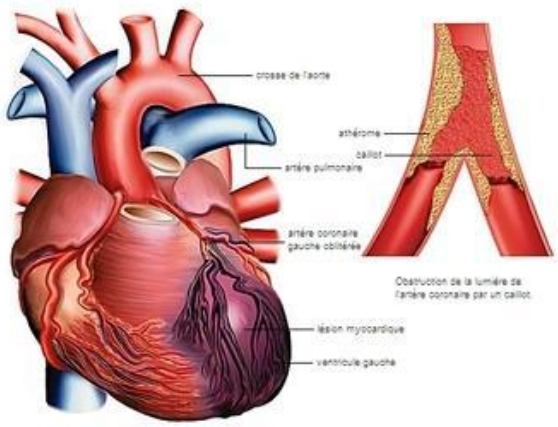



**Fig 1.3 :** Infarctus-du-myocarde (Benahmed, Beutler et Chabloz 2015)

- **Accident vasculaire cérébral (AVC) :** Un accident vasculaire cérébral est l'arrêt soudain de la circulation sanguine vers une ou plusieurs parties du cerveau. La rapidité de prise en charge est essentielle. Le symptôme le plus courant d'un accident vasculaire cérébral est une sensation soudaine de faiblesse au visage, au bras ou à la jambe, généralement d'un côté du corps (Ornstein et al. 2004).

La figure 1.4 en donne une représentation.



**Fig 1.4 :** Accident vasculaire cérébral (Ray 2017)

- **Athériosclérose :** L'athériosclérose est un épaississement et un durcissement de la paroi des artères (Ornstein et al. 2004). La figure 1.5 présente ce mécanisme.

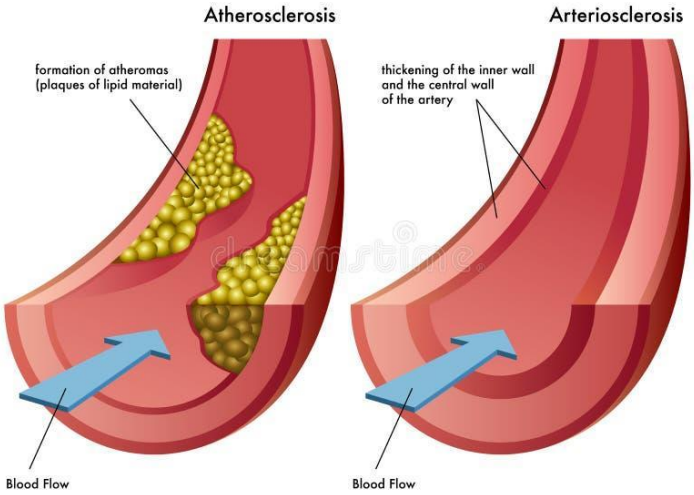



**Fig 1.5 :** Athérosclérose (Dreamstime 2026)

- **Angine de poitrine :** Angor L'angine de poitrine d'écrit une douleur violente localisée dans la poitrine (Ornstein et al. 2004). La figure 1.6 illustre cette manifestation.



**Fig 1.6 :** Angine de poitrine (Docteurclic 2026)

- **Insuffisance cardiaque :** On parle d'insuffisance cardiaque lorsque le cœur n'est plus capable d'effectuer correctement son travail de pompe. Il est possible d'agir sur certains facteurs de risque (tabagisme, sur-poids, diabète...) en modifiant durablement ses habitudes de vie (Ornstein et al. 2004). La figure 1.7 compare un cœur normal et un cœur atteint.

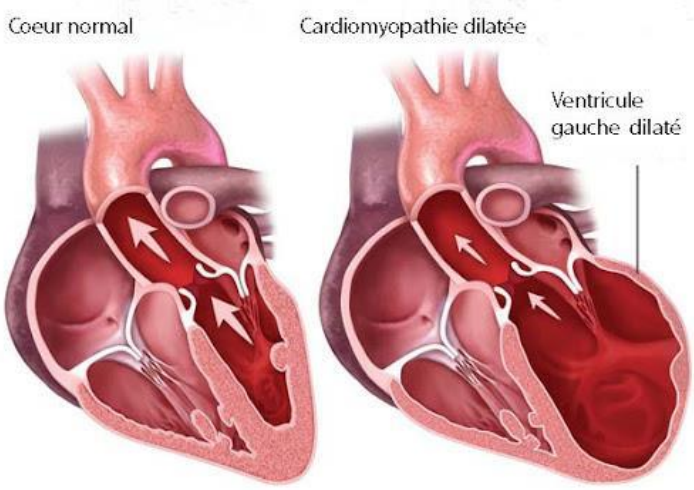



**Fig 1.7 :** La différence entre les deux cœurs (Google Images 2026)

- **Arythmie cardiaque :** L'arythmie cardiaque est un trouble du rythme cardiaque. Il en existe plusieurs types, de gravité variable (Ornstein et al. 2004). La figure 1.8 illustre ce trouble.

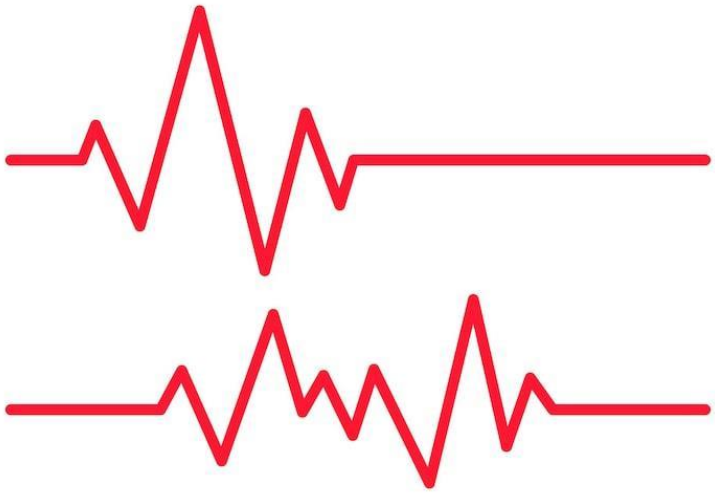


**Fig 1.8 :** Arythmie cardiaque (Freepik 2026)

- **Hypertension artérielle :** L'hypertension artérielle est une augmentation de la pression du sang dans les artères (Ornstein et al. 2004). La figure 1.9 présente ce phénomène.

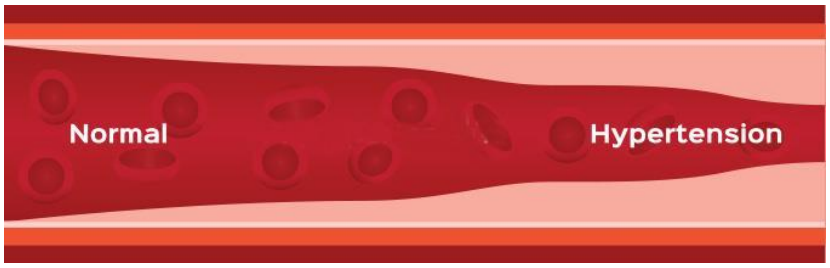



**Fig 1.9 :** Hypertension artérielle (Chirurgie Vasculaire Lyon 2026)

## 1.5 Diagnostique des maladies cardiaques

Si des symptômes évocateurs (fatigue, essoufflement, œdème, etc.) sont présents, une évaluation est nécessaire. Un examen physique permet aux médecins d'identifier l'insuffisance cardiaque, d'évaluer sa gravité et d'en trouver la cause (séquelles d'infarctus du myocarde, hypertension (HTA), valvulopathie, cardiomyopathie, etc.) (Cascino et Shea 2023). Les procédures diagnostiques peuvent être :

- Non invasives ;
- Minimalement invasives.

Les tests non invasifs ne nécessitent pas d'incision ou de ponction à l'aiguille,

parfois simplement en prélevant du sang ou en insérant un cathéter veineux court (IV) standard dans une veine du bras. Ces épreuves comprennent :

- Tomodensitométrie (TDM) ;
- Électrocardiographie (ECG) ;
- Imagerie par résonance magnétique (IRM) ;
- Tomographie par émission de positons (TEP) ;
- Scintigraphie ;
- Épreuve d'effort ;
- Test de la table basculante ;
- Échographie (y compris échocardiographie) ;
- Radiographies.

Les tests mini-invasifs nécessitent généralement un long cathéter flexible qui est inséré dans un vaisseau sanguin du poignet, du cou ou de la cuisse, puis à travers le vaisseau sanguin jusqu'au cœur. Ces épreuves comprennent :

- Angiographie ;
- Cathétérisme cardiaque ;
- Cathétérisme veineux central ;
- Examen électrophysiologique.

La plupart de ces procédures comportent peu ou pas de risque, mais le risque augmente avec la complexité de la procédure, la gravité de la maladie cardiaque et toute autre condition médicale dont souffre une personne. Le traitement peut parfois être administré lors de certains tests diagnostiques mini-invasifs. Par exemple, les patients atteints de maladie coronarienne peuvent subir une intervention coronarienne percutanée lors d'un cathétérisme cardiaque, tandis que les patients souffrant d'arythmies peuvent subir une ablation par radiofréquence lors d'examens électrophysiologiques (Cascino et Shea 2023).

## 1.6 Les principaux symptômes des maladies cardiaques

Cette liste de symptômes n'est pas exhaustive, mais fournit des informations sur les symptômes les plus courants chez les personnes atteintes de maladies car-

diaques. A l'inverse, la présence d'un de ces signes cliniques n'indique pas systématiquement la présence d'une cardiopathie. Les symptômes de chaque maladie sont sensiblement différents (Thompson et Shea 2024) :

- Douleur thoracique ;
- Essoufflement ;
- Fatigue ;
- Palpitations, perception des battements du cœur qui sont lents, rapides ou irréguliers ;
- Évanouissement, perte de connaissance, syncope ;
- Sensation de vertige ;
- Nausées et vomissements ;
- Douleur dans un membre, engourdissement ou crampe d'un muscle ;
- Gonflement des membres inférieurs (jambes, chevilles, pieds) ;
- Changement de la couleur de la peau au niveau d'un membre.

## 1.7 Facteurs de risque des maladies cardiaques

Les facteurs de risque cardiovasculaire regroupent des éléments qui augmentent la probabilité de développer une maladie cardiaque ou vasculaire. Ils sont généralement distingués en facteurs non modifiables, liés aux caractéristiques propres de l'individu, et en facteurs modifiables, sur lesquels la prévention et la prise en charge peuvent agir.

### 1.7.1 Facteurs de risque non modifiables

Ces facteurs ne peuvent pas être supprimés, mais leur identification permet d'estimer le niveau de risque global et d'adapter le suivi médical.

**Âge :** C'est un facteur de risque persistant qui augmente progressivement l'incidence de l'athérosclérose aortique, coronarienne et carotidienne et des complications de l'insuffisance cardiaque. Ce risque devient significatif à partir de 50 ans pour les hommes et 60 ans pour les femmes.

**Sexe masculin :** Avant l'âge de 70 ans, les deux tiers des crises cardiaques surviennent chez les hommes. Cette différence diminue chez les femmes ménopausées et disparaît après 75 ans.

**Hérédité :** Les antécédents familiaux de maladie cardiaque, coronarienne, d'accident vasculaire cérébral ou de mort subite sont des facteurs de risque, surtout s'ils surviennent chez un ou plusieurs parents du premier degré, à un âge jeune ($< 55$ ans pour le père ou $< 65$ ans pour la mère) (Laroche et al. 2005).

### 1.7.2 Facteurs de risque modifiables

Les facteurs modifiables correspondent aux habitudes de vie ou aux paramètres biologiques pouvant être corrigés ou contrôlés afin de réduire le risque cardiovasculaire.

**Tabagisme :** Il augmente les lésions athérosclérotiques en modifiant la fonction endothéliale, en perturbant la vasomotion, en activant l'agrégation plaquettaire et en réduisant le cholestérol des lipoprotéines de haute densité. Elle est athérogène et prothrombotique. Le risque relatif de crise cardiaque était de 5, et le risque relatif de maladie artérielle des membres inférieurs était de 2. Ce risque relatif existe également lors du tabagisme passif. Le risque est proportionnel à l'exposition au tabac et s'évalue en paquets-années. Les bénéfices de l'arrêt du tabac sont rapides : l'augmentation du risque relatif disparaît en 3 ans, et le risque de récidive est réduit de 50% chez les patients coronariens (Laroche et al. 2005).

**Hypertension artérielle :** Elle est définie par une valeur de pression de $> 140$ mmHg pour la pression artérielle systolique (PAS) ou de $> 90$ mmHg pour la pression artérielle diastolique (PAD). Tous les types d'hypertension sont des facteurs de risque : hypertension permanente, hypertension paroxystique, traitée ou non. Le risque relatif était de 7 points pour les accidents vasculaires cérébraux, 3 points pour les maladies coronariennes et 2 points pour les maladies artérielles des membres inférieurs. Avant 55 ans, ce risque était associé aux valeurs de tension artérielle systolique et diastolique. Après 60 ans, la corrélation est plus forte avec la pression différentielle (PAS-PAD), et donc la pression artérielle systolique surtout chez les personnes âgées. Le traitement de l'hypertension artérielle réduit le risque d'AVC de 40% et le risque d'infarctus de 15% (Laroche et al. 2005).

**Dyslipidémies :** Parmi les dyslipidémies circulantes, le principal facteur de risque de maladie cardiovasculaire est l'élévation du cholestérol LDL, avec un choles-

térol lié au LDL 1,60 g/L (4,1 mmol/L). Le cholestérol LDL était positivement associé au risque de maladie cardiovasculaire, tandis que le cholestérol HDL était associé négativement s'il était > 0,40 g/L (1 mmol/L). Des taux élevés de triglycérides (> 2,0 g/L) seuls ne sont pas un facteur de risque (indépendant), mais peuvent être un facteur de risque en combinaison avec d'autres facteurs (voir Syndrome métabolique). Le cholestérol à lipoprotéines de basse densité a un rôle direct dans la croissance des plaques d'athérosclérose et leur rupture due à l'instabilité. Le risque relatif d'hypercholestérolémie pour les maladies coronariennes était 3 fois plus élevé que pour les maladies artérielles et les accidents vasculaires cérébraux. L'efficacité du traitement de l'hypercholestérolémie est un facteur majeur de réduction de la mortalité cardiovasculaire (30% à 20 ans) (Laroche et al. 2005).

**Diabète :** Le diabète a été défini comme deux mesures à jeun > 1,26 g/L (7 mmol/L) ou une seule mesure de glycémie > 2 g/L (11 mmol/L). Le diabète de type I ou de type II est associé à un risque cardiovasculaire accru. Les complications cardiovasculaires du diabète I commencent plus tôt dans la trentaine, mais l'apparition rapide du diabète II en fait un facteur de risque très préoccupant. Il a un risque relatif de > 2, causant principalement plus de maladies artérielles que de maladies coronariennes et d'accidents vasculaires cérébraux. Mais le diabète se complique plus souvent de microangiopathie (rétinopathie et maladie rénale). Ce risque relatif augmente avec les anomalies rénales. Le traitement du diabète avec un objectif de 6,5% d'hémoglobine glyquée (HbA1c) réduit les complications cardiovasculaires (Laroche et al. 2005).

**Insuffisance rénale :** L'insuffisance rénale chronique est associée à un taux élevé de complications cardiovasculaires, comparable à la sévérité du diabète sur le système cardiovasculaire (Laroche et al. 2005 ; Lachance 2015).

## 1.8 Prévention des maladies cardiaques

La prévention des maladies cardiaques repose sur des actions complémentaires menées à l'échelle individuelle et collective. Elle vise à limiter l'exposition aux facteurs de risque, à favoriser le dépistage précoce et à réduire la fréquence des complications cardiovasculaires.

### 1.8.1 Prévention individuelle

L'objectif est que chacun élimine ou réduise au maximum les facteurs de risque modifiables. Cela comprend l'élimination du tabac, le respect du mode de vie et des mesures diététiques et la pratique d'une activité physique régulière. Des médicaments peuvent également être prescrits en respectant les conseils selon Janin (2015) :

- Le résultat de ces premières mesures ;
- Le nombre de facteurs de risque ;
- Le calcul du risque absolu ;
- La présence d'antécédents cardiovasculaires.

### 1.8.2 Prévention collective

Son champ d'application est considérable, car il touche toute la population et comprend notamment les actions suivantes (Janin 2015) :

- La réglementation anti-tabac ;
- La limitation réglementée de la teneur en sel dans l'industrie alimentaire ;
- L'éducation et l'alimentation scolaires, principaux axes pour freiner l'obésité croissante ;
- L'information de la population par diverses campagnes nationales ou régionales ;
- L'accès aux équipements sportifs et zones de plein air.

## 1.9 Extraction et sélection des caractéristiques

L'extraction et la sélection des caractéristiques sont des étapes importantes dans le prétraitement des données en apprentissage automatique. L'extraction des caractéristiques consiste à transformer les données brutes en une représentation numérique exploitable par les algorithmes de classification. Dans le cas de l'analyse des sentiments, les textes sont convertis en vecteurs numériques à l'aide de techniques telles que Bag of Words (BoW) ou TF-IDF.

La sélection des caractéristiques vise ensuite à identifier les attributs les plus pertinents pour la tâche de classification. Cette étape permet d'éliminer les caractéristiques redondantes ou peu informatives, de réduire la dimension des données et d'améliorer les performances du modèle. Elle contribue également à diminuer le temps d'apprentissage et à limiter le risque de surapprentissage (overfitting).

Dans notre travail, l'extraction des caractéristiques permet de représenter les textes du dialecte algérien sous une forme numérique, tandis que la sélection des caractéristiques permet de conserver uniquement les termes les plus significatifs pour la prédiction des sentiments (Guyon et Elisseeff 2003).

## 1.10 Conclusion

Dans cette section, nous avons présenté les notions d'extraction et de sélection des caractéristiques, qui constituent des étapes importantes dans le prétraitement des données. Ces techniques permettent de transformer les données brutes en informations exploitables et de conserver uniquement les caractéristiques les plus pertinentes pour améliorer les performances des modèles d'apprentissage automatique. Dans le chapitre suivant, nous présenterons les méthodes et les algorithmes utilisés pour la classification et l'analyse

CHAPITRE 2

# Apprentissage automatique : concepts et méthodes

## 2.1 Introduction

Le Deep Learning constitue aujourd'hui l'une des approches les plus avancées de l'intelligence artificielle, particulièrement dans le domaine de l'analyse des données biomédicales. Grâce à l'augmentation de la puissance de calcul et à la disponibilité de grandes bases de données médicales, les modèles profonds ont démontré une capacité remarquable à extraire automatiquement des caractéristiques pertinentes à partir de données complexes et hétérogènes.

Dans le contexte des maladies cardiovasculaires, qui représentent l'une des principales causes de mortalité dans le monde, l'utilisation de techniques de Deep Learning offre de nouvelles perspectives pour l'aide au diagnostic précoce et la prédiction du risque chez les patients. Contrairement aux approches traditionnelles de machine learning, qui nécessitent souvent une ingénierie manuelle des caractéristiques, les réseaux de neurones profonds permettent un apprentissage hiérarchique des représentations directement à partir des données brutes.

Les architectures de Deep Learning, telles que les réseaux de neurones artificiels profonds (Artificial Neural Networks – ANN), les réseaux convolutionnels (Convolutional Neural Networks – CNN) et les réseaux récurrents (Recurrent Neural Networks – RNN), ont été largement exploitées dans plusieurs applications médicales, notamment la classification des patients, l'analyse des signaux physiologiques et la détection automatisée des pathologies. Leur capacité à modéliser des relations non linéaires complexes entre les variables cliniques les rend particulièrement adaptées aux problèmes de prédiction des maladies cardiaques. Par ailleurs, l'intégration du Deep Learning dans les systèmes d'aide à la décision médicale contribue à améliorer la précision diagnostique, à réduire le temps d'analyse et à assister les

professionnels de santé dans la prise de décision clinique. Toutefois, malgré leurs performances prometteuses, ces méthodes nécessitent une évaluation rigoureuse à l'aide de métriques appropriées afin de garantir leur fiabilité et leur généralisabilité dans des contextes réels. Dans ce chapitre, nous présentons d'abord les notions fondamentales de l'apprentissage supervisé et non supervisé. Ensuite, nous décrivons les principales architectures de Deep Learning utilisées pour la prédiction des maladies cardiaques. Enfin, nous exposons les métriques d'évaluation couramment employées pour mesurer la performance des modèles proposés.

## 2.2 Apprentissage supervisé et non supervisé

L'apprentissage supervisé et l'apprentissage non supervisé constituent les deux paradigmes fondamentaux du Deep Learning. Ces approches se distinguent principalement par la présence ou l'absence d'étiquettes (labels) dans les données d'entraînement. L'apprentissage supervisé repose sur des exemples annotés, permettant au modèle d'apprendre à associer les entrées à des sorties connues. En revanche, l'apprentissage non supervisé explore les données brutes afin de découvrir des structures et des motifs cachés sans supervision extérieure. Comme le soulignent Esteva et al. (2019) dans leur guide sur le Deep Learning en santé, la majorité des avancées actuelles s'appuient sur l'apprentissage supervisé, qui nécessite de grandes quantités de données étiquetées. Dans le domaine médical, l'annotation des données par des experts représente toutefois un coût élevé en termes de temps et de ressources. Face à cette limitation, l'apprentissage non supervisé et les approches semi-supervisées apparaissent comme des solutions prometteuses, en particulier lorsqu'un volume important de données non étiquetées est disponible. Dans le cadre d'une thèse sur la prédiction des maladies cardiovasculaires, le choix entre ces paradigmes doit tenir compte de la disponibilité réelle des données annotées ainsi que des contraintes cliniques et économiques inhérentes au domaine médical.

### 2.2.1 L'apprentissage supervisé

L'apprentissage supervisé constitue la méthode la plus couramment utilisée en Deep Learning appliqué à la santé, dès lors que l'on dispose de données étiquetées par des experts médicaux (Esteva et al. 2019). Dans cette approche, chaque exemple d'entrée est associé à une étiquette cible validée, et le modèle apprend à prédire correctement ces étiquettes sur de nouvelles données non vues. Le processus d'apprentissage repose sur l'ajustement itératif des paramètres internes du modèle (notamment les poids des connexions neuronales) afin de minimiser l'erreur entre les pré-

dictions du modèle et les étiquettes réelles. Cet ajustement est réalisé principalement par l'algorithme de rétropropagation du gradient (backpropagation) et l'utilisation de fonctions de perte adaptées au type de tâche, telles que l'entropie croisée pour les problèmes de classification ou l'erreur quadratique moyenne pour les problèmes de régression (Goodfellow, Bengio et Courville 2016). Comme le soulignent Esteva et al. (2019), la grande majorité des avancées actuelles en intelligence artificielle médicale s'appuie sur l'apprentissage supervisé, car il permet d'obtenir des prédictions directes et mesurables. Cette méthode s'avère particulièrement adaptée aux applications médicales où la précision et la fiabilité revêtent une importance capitale. Dans le cadre d'une thèse sur la prédiction des maladies cardiovasculaires, l'apprentissage supervisé fournit un cadre méthodologique solide pour exploiter les données cliniques annotées, qu'il s'agisse de signaux physiologiques, d'images médicales ou de variables clinico-biologiques, et pour développer des modèles susceptibles d'assister les cliniciens dans la détection précoce des anomalies et la stratification du risque. Toutefois, l'efficacité de cette approche dépend fortement de la qualité et surtout de la quantité des données étiquetées disponibles. Dans le contexte médical, l'annotation des données par des experts représente un investissement considérable en temps et en ressources, constituant l'une des principales limitations de l'apprentissage supervisé (Esteva et al. 2019).

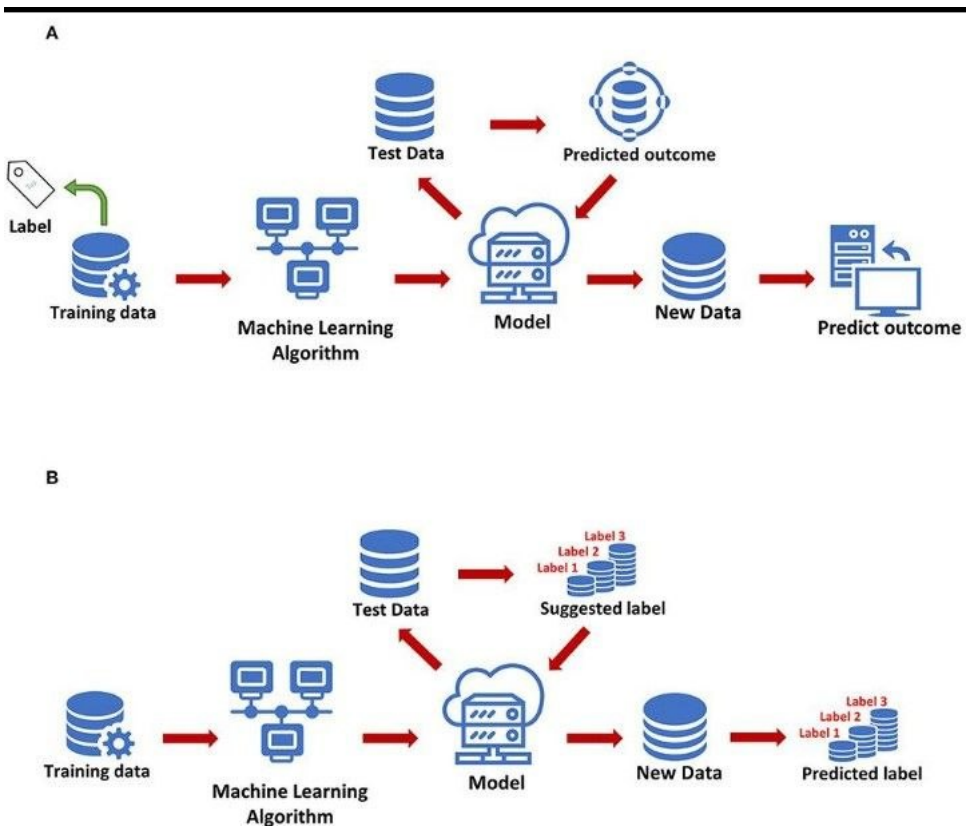



**Fig 2.1 :** Workflow de l'apprentissage supervisé appliqué aux données médicales (adapté et inspiré de Esteva et al. (2019))

La figure 2.1 illustre les étapes principales de l'apprentissage supervisé en contexte médical, incluant le prétraitement des données, l'entraînement du modèle et la phase de prédiction.

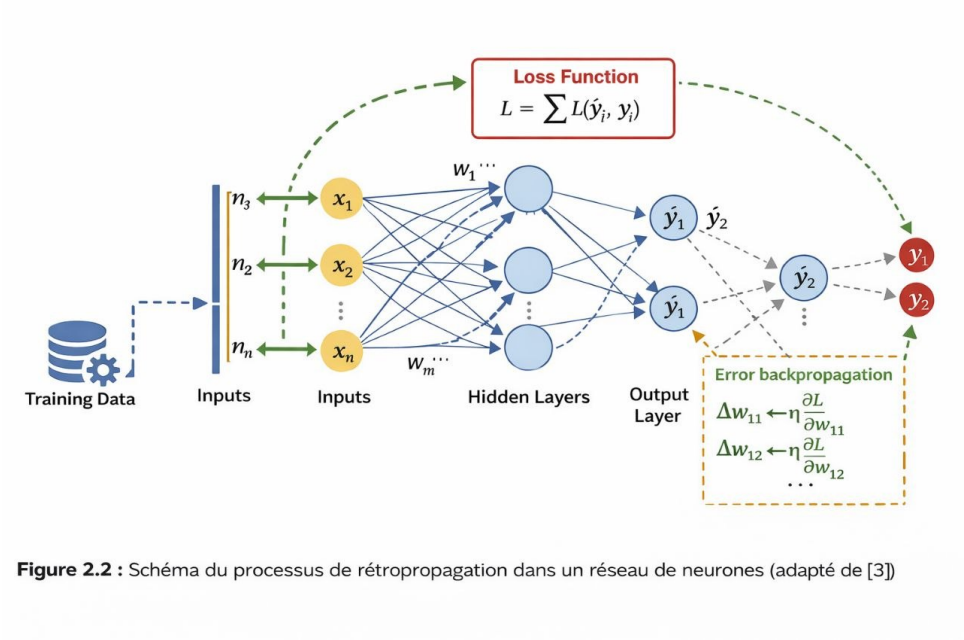



**Figure 2.2 :** Schéma du processus de rétropropagation dans un réseau de neurones (adapté de [3])

**Fig 2.2 :** Schéma du processus de rétropropagation (adapté de Goodfellow, Bengio et Courville (2016))

La rétropropagation (ou backpropagation) est l'algorithme fondamental qui permet d'entraîner un réseau de neurones. Comme le montre la figure 2.2, le processus se déroule en deux phases principales : Passage avant (Forward pass) : Les données d'entrée ($x_1$, $x_2$, ..., $x_n$) traversent les couches cachées jusqu'à la couche de sortie pour produire les prédictions $\hat{y}$. On calcule ensuite la fonction de perte $L = \sum_i L(\hat{y}_i, y_i)$ qui mesure l'erreur entre les prédictions et les valeurs réelles y. Passage arrière (Backward pass) : L'erreur est propagée en sens inverse à travers le réseau (flèche verte). En utilisant la règle de dérivation en chaîne, on calcule les gradients ($\partial L/\partial w$) pour chaque poids. Ces gradients servent ensuite à mettre à jour les poids selon la formule : $\Delta w \leftarrow -\eta \partial L/\partial w$ où $\eta$ est le taux d'apprentissage (learning rate). Ce mécanisme permet au réseau d'ajuster progressivement ses poids afin de minimiser l'erreur et d'améliorer ses performances.

### 2.2.2 Types de l'apprentissage supervisé

L'apprentissage supervisé est une méthode de machine learning où le modèle apprend à partir de données étiquetées, c'est-à-dire des entrées associées à des sorties connues. L'objectif est d'apprendre une fonction capable de prédire correctement les sorties à partir de nouvelles données. Ce processus d'apprentissage permet au

modèle de construire une relation mathématique entre les variables d'entrée et la variable cible. Cette relation peut être linéaire ou non linéaire selon la complexité du modèle utilisé. De plus, la performance dépend fortement de la qualité des données d'entraînement ainsi que du choix des caractéristiques pertinentes (Bishop 2006; Hastie, Tibshirani et Friedman 2009).

**Classification** La classification est une tâche qui consiste à prédire une catégorie discrète à partir des données d'entrée. Chaque observation est assignée à une classe prédéfinie. Exemples : spam / non spam, malade / sain.

La classification peut être réalisée à l'aide de modèles linéaires ou non linéaires. Elle repose généralement sur l'apprentissage de frontières de décision permettant de séparer les différentes classes dans l'espace des données. Dans le domaine médical, elle est largement utilisée pour le diagnostic assisté (Bishop 2006; Hastie, Tibshirani et Friedman 2009).

B)Régression

La régression consiste à prédire une valeur numérique continue à partir des variables d'entrée. Elle est utilisée lorsque la sortie n'est pas une classe mais une valeur réelle. Exemples : prix d'une maison, température. Les modèles de régression cherchent à minimiser une fonction de perte telle que l'erreur quadratique moyenne (MSE). Ils permettent également d'évaluer l'influence des variables explicatives sur la variable cible (Bishop 2006; Montgomery, Peck et Vining 2012).

### 2.2.3 Algorithmes de classification

Plusieurs algorithmes peuvent être utilisés pour attribuer une observation à une classe prédéfinie. Le choix du modèle dépend notamment de la nature des données, de la complexité des relations entre variables et du besoin d'interprétabilité.

#### 2.2.3.1 Régression logistique

La régression logistique est utilisée pour la classification binaire. Elle modélise la probabilité d'appartenance à une classe à l'aide d'une fonction logistique. Elle peut être étendue aux problèmes multiclasse à travers des stratégies comme "one-vs-rest". Elle est appréciée pour sa simplicité et son interprétabilité (Bishop 2006; Hastie, Tibshirani et Friedman 2009).

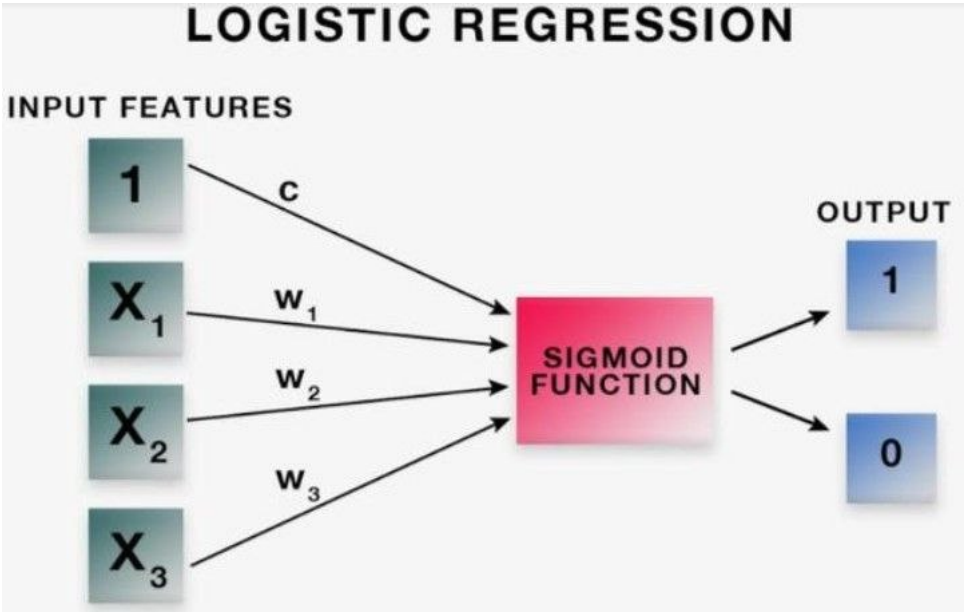



**Fig 2.3 :** Frontière de décision et fonction sigmoïde en régression logistique (adapté de Bishop (2006) et Hastie, Tibshirani et Friedman (2009))

**Explication :** Comme le montre la figure 2.3, la régression logistique utilise la fonction sigmoïde pour transformer une entrée en probabilité entre 0 et 1 afin de réaliser une classification binaire (Bishop 2006 ; Hastie, Tibshirani et Friedman 2009).

#### 2.2.3.2 Support Vector Machine (SVM)

Le SVM est un algorithme qui cherche à maximiser la marge entre les classes en utilisant un hyperplan optimal. Il peut utiliser des fonctions noyaux (kernels) pour traiter des données non linéaires, en projetant les données dans un espace de dimension plus élevée (Vapnik 1995).

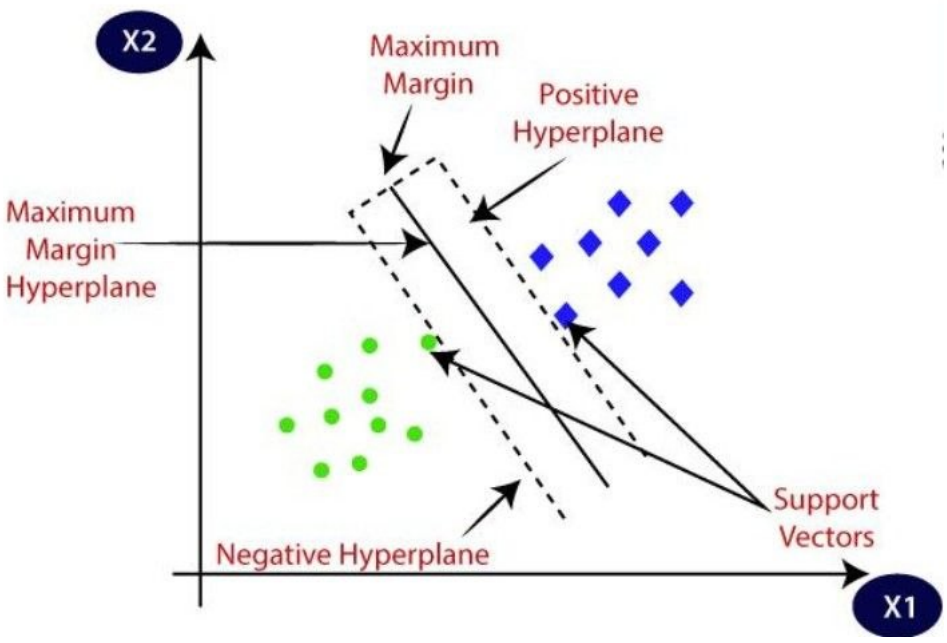



**Fig 2.4 :** Hyperplan optimal et marge maximale en SVM (adapté de Vapnik (1995))

**Explication :** Comme l'illustre la figure 2.4, le SVM cherche l'hyperplan optimal qui sépare les classes avec la marge la plus large possible afin d'améliorer la généralisation (Vapnik 1995).

#### 2.2.3.3 Arbre de décision

Un arbre de décision est une structure hiérarchique composée de nœuds représentant des décisions basées sur les caractéristiques. Il est facile à interpréter mais peut être sensible au surapprentissage si sa profondeur n'est pas contrôlée (Hastie, Tibshirani et Friedman 2009).

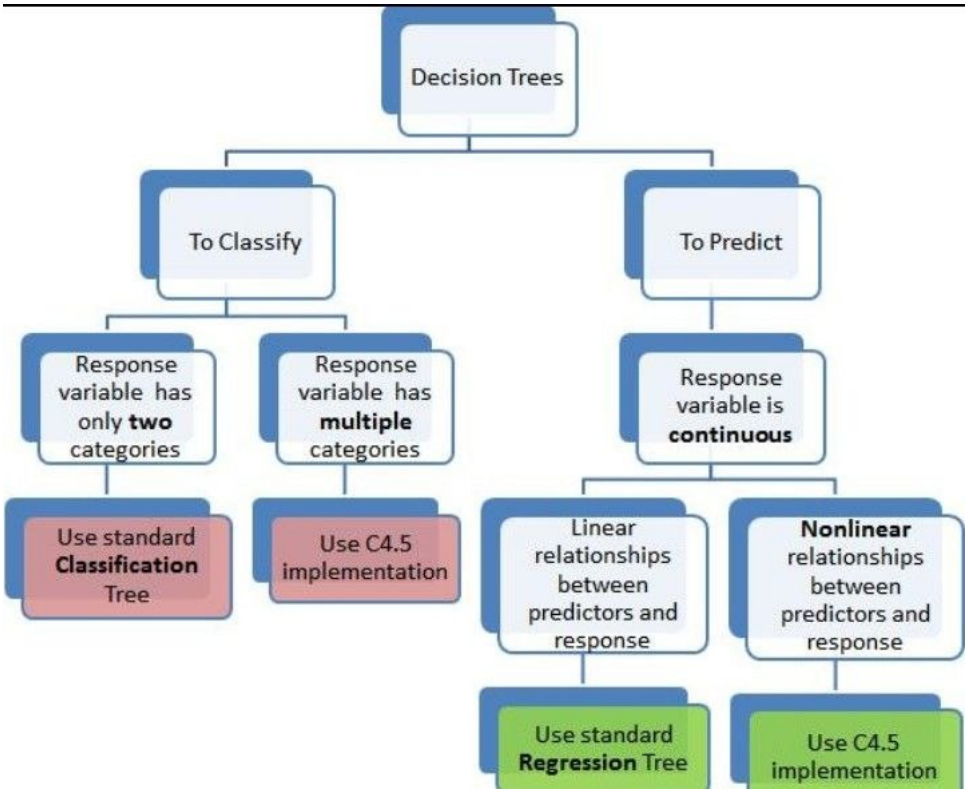



**Fig 2.5 :** Structure d'un arbre de décision (adapté de Hastie, Tibshirani et Friedman (2009))

**Explication :** Comme le montre la figure 2.5, un arbre de décision fonctionne en posant des questions successives (if/else) afin d'aboutir à une décision finale (Hastie, Tibshirani et Friedman 2009).

#### 2.2.3.4 K plus proches voisins (KNN)

Le KNN est un algorithme basé sur la similarité entre les données. Il classe un point en fonction des k plus proches voisins. C'est une méthode non paramétrique qui ne nécessite pas de phase d'entraînement explicite, mais son coût de calcul peut être élevé lors de la prédiction (Cover et Hart 1967).

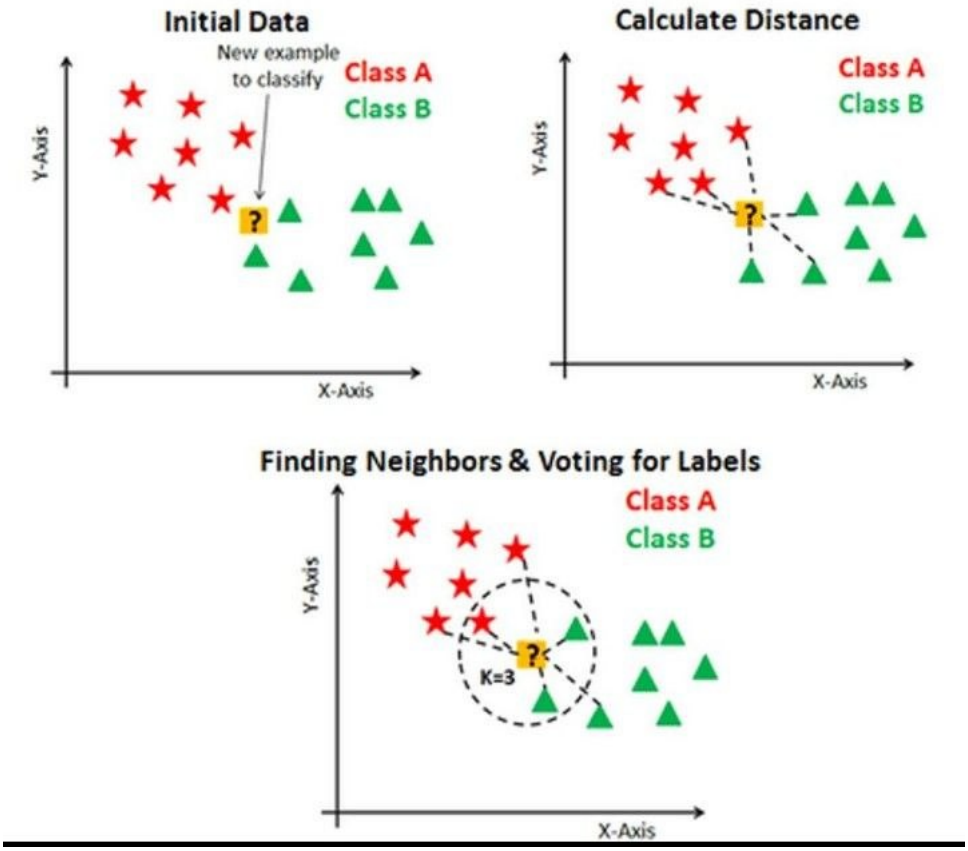



**Fig 2.6 :** Classification par les K plus proches voisins (adapté de Cover et Hart (1967))

**Explication :** Comme l'illustre la figure 2.6, KNN classe un point en fonction des K voisins les plus proches dans l'espace des caractéristiques (Cover et Hart 1967).

### 2.2.4 Algorithmes de régression

Les algorithmes de régression visent à estimer une variable numérique continue à partir d'un ensemble de variables explicatives. Ils sont utilisés lorsque l'objectif n'est pas de prédire une classe, mais une valeur mesurable.

#### 2.2.4.1 Régression linéaire

La régression linéaire modélise la relation entre les variables sous forme d'une équation linéaire. Elle peut être étendue à des modèles plus complexes comme la régression polynomiale. Elle est largement utilisée pour sa simplicité (Montgomery, Peck et Vining 2012).

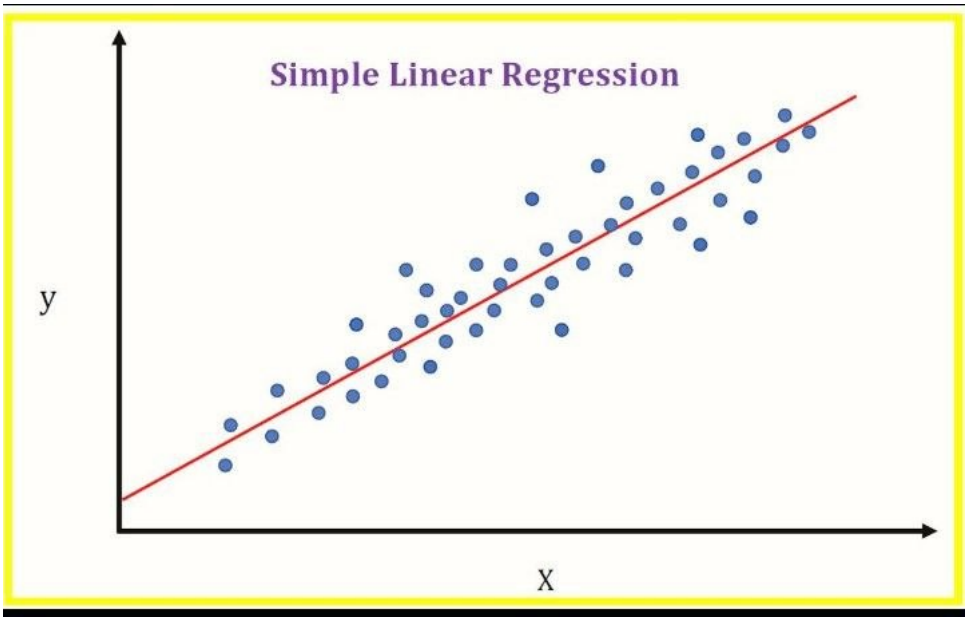



**Fig 2.7 :** Droite de régression linéaire par moindres carrés (adapté de Montgomery, Peck et Vining (2012))

**Explication :** La figure 2.7 illustre la manière dont la régression linéaire modélise une relation linéaire entre variables pour prédire une valeur continue (Montgomery, Peck et Vining 2012).

#### 2.2.4.2 Support Vector Regression (SVR)

Le SVR est une extension du SVM pour les problèmes de régression. Il introduit une marge d'erreur $\varepsilon$ permettant d'ignorer les petites erreurs, ce qui le rend robuste face au bruit (Vapnik 1998).

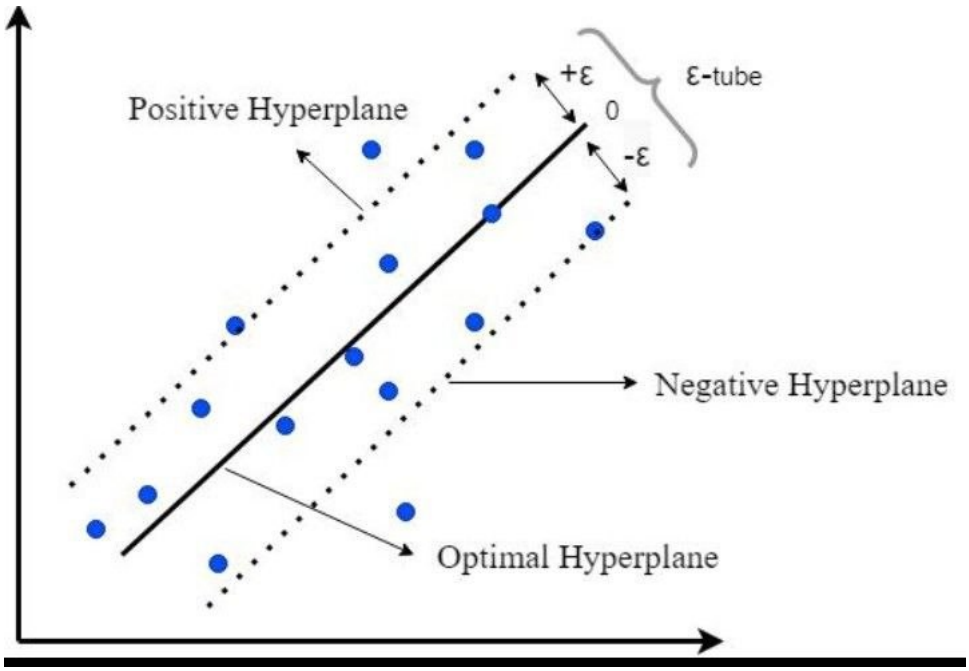



**Fig 2.8 :** Régression SVR avec tube $\varepsilon$ (adapté de Vapnik (1998))

**Explication :** Comme le montre la figure 2.8, SVR est une extension de SVM pour la régression, utilisant un tube d'erreur $\varepsilon$ autour de la courbe (Vapnik 1998).

#### 2.2.4.3 Decision Tree Regression

Cette méthode utilise un arbre de décision pour prédire des valeurs continues. Elle est capable de modéliser des relations non linéaires mais peut être instable sans régularisation (Hastie, Tibshirani et Friedman 2009).

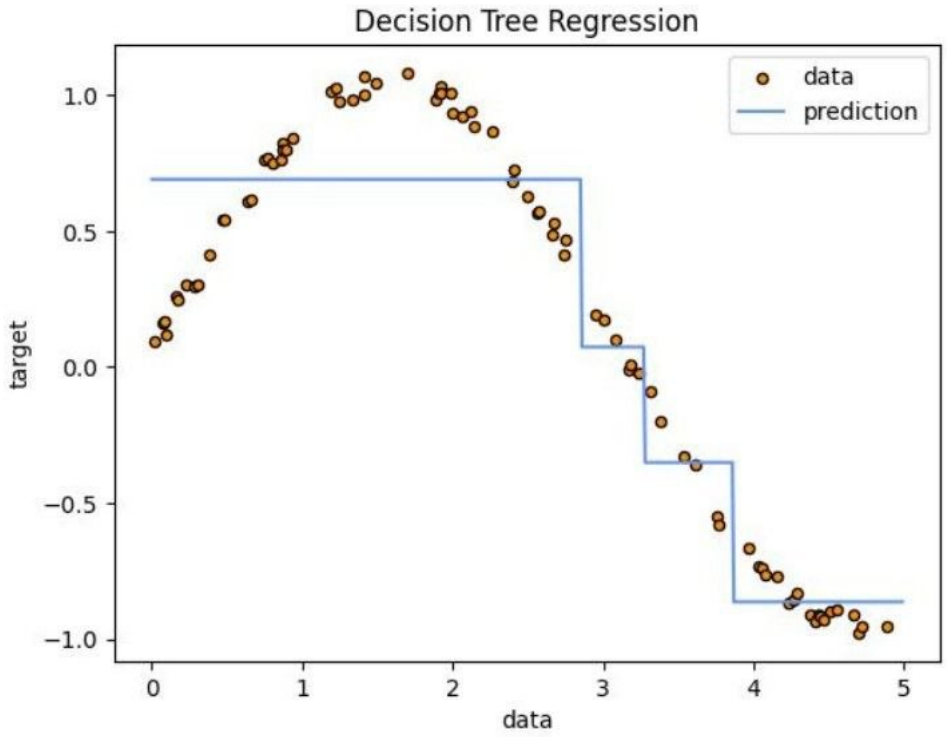



**Fig 2.9 :** Arbre de régression et prédictions par segments (adapté de Hastie, Tibshirani et Friedman (2009))

**Explication :** Comme l'illustre la figure 2.9, un arbre de régression prédit des valeurs numériques en divisant l'espace en régions (Hastie, Tibshirani et Friedman 2009).

#### 2.2.4.4 Random Forest Regression

Random Forest est une méthode d'ensemble qui combine plusieurs arbres de décision. Elle améliore la précision et réduit le surapprentissage en utilisant des sous-ensembles de données aléatoires (Breiman 2001).

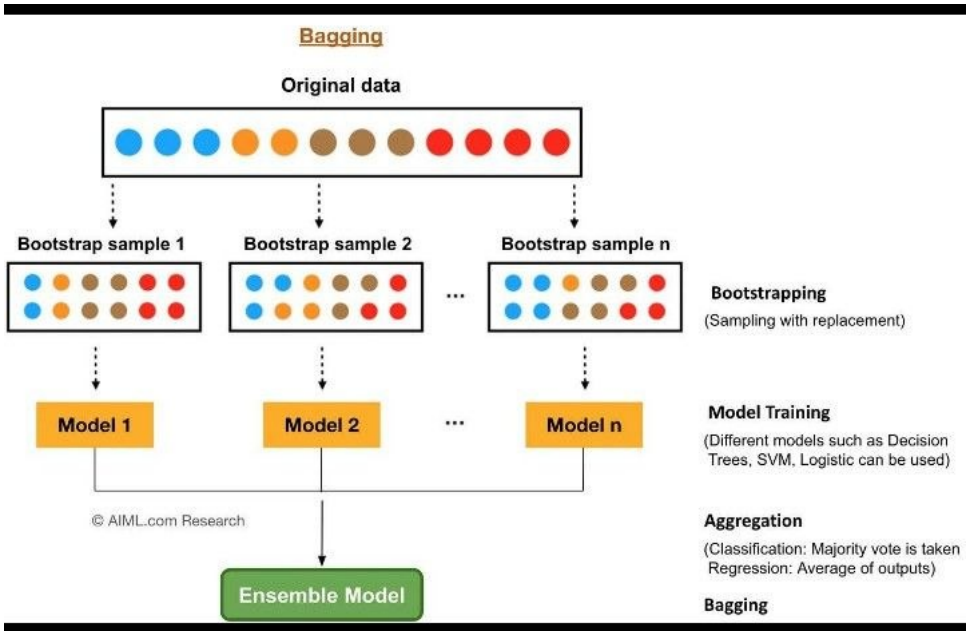



**Fig 2.10 :** Forêt aléatoire (Random Forest) et moyenne des prédictions (adapté de Breiman (2001))

**Explication :** La figure 2.10 montre comment Random Forest combine plusieurs arbres de décision et calcule la moyenne pour améliorer la précision et réduire le surapprentissage (Breiman 2001).

### 2.2.5 Comparaison entre classification et régression

Le tableau 2.1 résume les différences principales entre classification et régression selon le type de sortie attendu, les objectifs et les exemples d'algorithmes utilisés.

**Tab 2.1 :** Comparaison entre classification et régression

| Critère | Classification | Régression |
|---|---|---|
| Objectif | Prédire une catégorie (classe) | Prédire une valeur numérique continue |
| Type de sortie | Discrète | Continue |
| Malade / Sain | Malade / Sain | Prix d'une maison |
| Nature | Classes | Valeurs numériques |
| Algorithmes | SVM, KNN, Logistic Regression | Linear Regression, SVR |
| Représentation | Frontières de décision | Courbe / droite |
| Évaluation | Accuracy, Recall | MSE, MAE |

### 2.2.6 L'apprentissage non supervisé

L'apprentissage non supervisé fonctionne sans aucune étiquette préalable (Bishop 2006). Contrairement à l'apprentissage supervisé qui nécessite des données annotées par des experts, le modèle explore les données brutes pour découvrir par lui-même des structures cachées, telles que des regroupements (clustering) de patients présentant des profils similaires ou la détection d'anomalies rares dans des signaux physiologiques complexes. Cette approche repose sur des techniques comme les auto-encodeurs, les réseaux génératifs antagonistes (GANs) ou l'apprentissage auto-supervisé, qui permettent au modèle d'apprendre des représentations riches et utiles à partir de données non étiquetées. Elle est particulièrement précieuse en cardiologie, où l'annotation médicale est souvent limitée, coûteuse et chronophage. C'est le cas des données massives et continues issues des dispositifs portables (wearables) et des enregistrements Holter de longue durée, qui génèrent des volumes énormes de signaux ECG difficiles à annoter manuellement de manière exhaustive (Esteva et al. 2019). Dans ce contexte, l'apprentissage non supervisé (ou semi-supervisé) permet de regrouper les patients selon des patterns électrocardiographiques similaires, de détecter des anomalies subtiles ou intermittentes (comme certaines arythmies), et même de prétraiter les données pour améliorer les performances des modèles supervisés ultérieurs. Ces avancées ouvrent la voie à une surveillance cardiaque plus accessible, personnalisée et proactive, surtout face à l'explosion des données provenant des technologies portables.

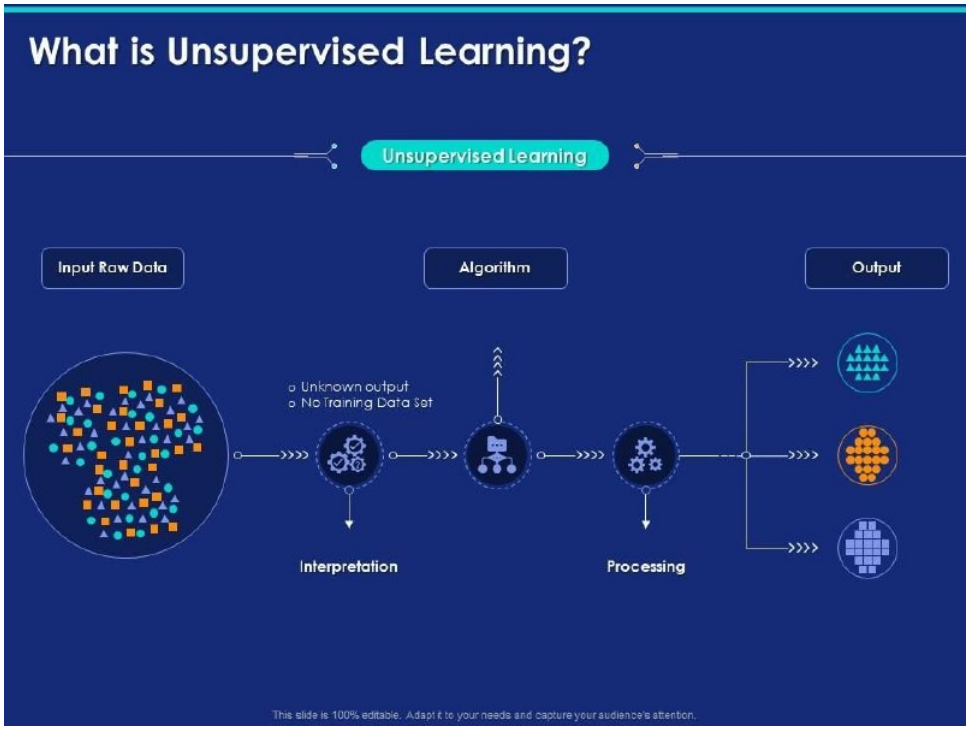



**Fig 2.11 :** Processus de l'apprentissage non supervisé sur données brutes non étiquetées

La figure 2.11 illustre le processus de l'apprentissage non supervisé. Le modèle reçoit des données brutes non étiquetées (Input Raw Data) et, sans aucune annotation préalable, il découvre par lui-même les structures et les patterns cachés présents dans ces données, tels que le regroupement (clustering). Cette approche est particulièrement utile lorsque l'annotation des données par des experts est coûteuse ou difficile.

### 2.2.7 Types de l'apprentissage non supervisé

L'apprentissage non supervisé fonctionne sans aucune étiquette préalable. Le modèle analyse les données afin de découvrir des structures cachées ou des relations implicites. Dans cette approche, le modèle cherche à identifier des regroupements naturels ou des patterns dans les données. Elle est particulièrement utile lorsque les données ne sont pas annotées (Bishop 2006 ; Hastie, Tibshirani et Friedman 2009).

**Clustering (Regroupement)** . Le clustering consiste à regrouper les données similaires en ensembles appelés clusters. La qualité du regroupement dépend du choix de la métrique de distance et des paramètres utilisés. Cette technique est largement utilisée pour l'analyse exploratoire (Bishop 2006 ; Hastie, Tibshirani et Friedman 2009).

**Réduction de dimension** . La réduction de dimension consiste à réduire le nombre de variables tout en conservant l'information essentielle. Elle permet de simplifier les données, réduire le bruit et améliorer les performances des modèles (Bishop 2006; Jolliffe 2002).

### 2.2.8 Algorithmes de clustering

Les algorithmes de clustering cherchent à regrouper des observations similaires sans utiliser d'étiquettes préalables. Ils sont utiles pour explorer la structure des données et identifier des profils ou sous-groupes potentiels.

#### 2.2.8.1 K-Means

K-Means est un algorithme qui partitionne les données en K groupes en minimisant la distance intra-cluster. Il nécessite de définir le nombre de clusters à l'avance, ce qui peut être une limitation (MacQueen 1967).

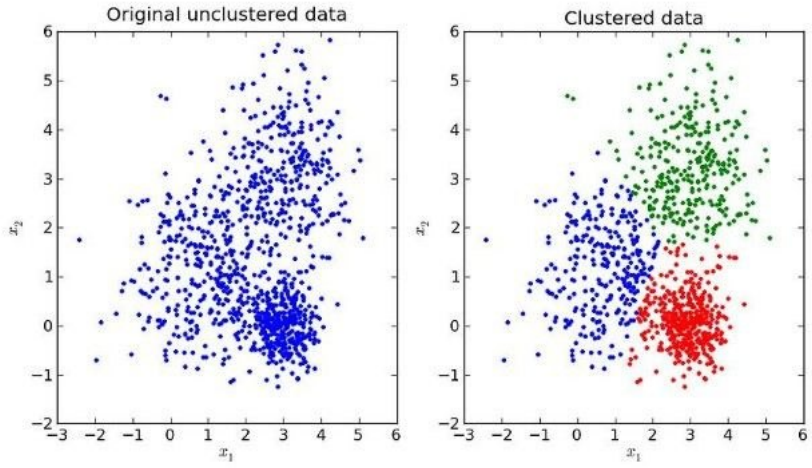



**Fig 2.12 :** Regroupement des données par K-Means (adapté de MacQueen (1967))

**Explication :** Comme le montre la figure 2.12, K-Means partitionne les données en K clusters en minimisant la distance entre les points et leurs centroïdes (MacQueen 1967).

#### 2.2.8.2 Clustering hiérarchique

Cette méthode construit une hiérarchie de clusters sous forme d'arbre. Elle permet une analyse visuelle intuitive des relations entre les données (Hastie, Tibshirani et Friedman 2009).

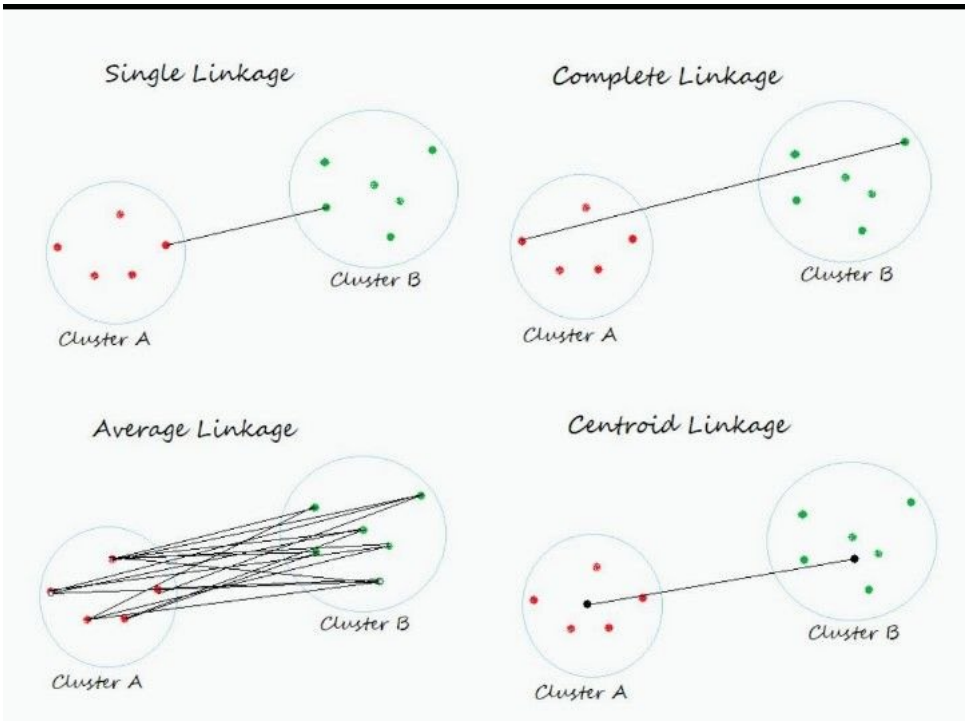



**Fig 2.13 :** Dendrogramme du clustering hiérarchique (adapté de Hastie, Tibshirani et Friedman (2009))

**Explication :** La figure 2.13 présente cette hiérarchie de clusters sous forme d'arbre (dendrogramme) (Hastie, Tibshirani et Friedman 2009).

#### 2.2.8.3 DBSCAN

DBSCAN regroupe les points denses et identifie les points isolés comme anomalies. Il est efficace pour détecter des structures complexes et gérer le bruit (Ester et al. 1996).

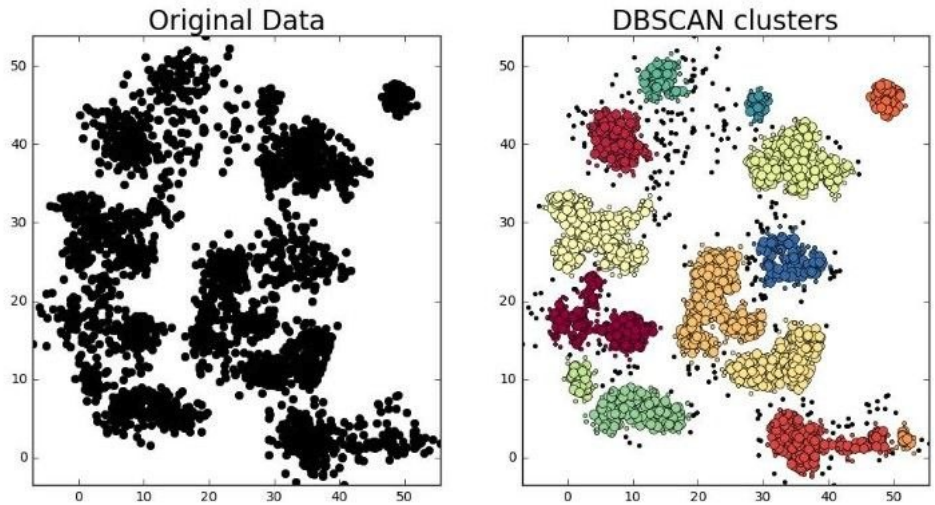



**Fig 2.14 :** Clustering basé sur la densité (DBSCAN) (adapté de Ester et al. (1996))

**Explication :** La figure 2.14 montre que DBSCAN regroupe les points selon la densité et identifie aussi les points aberrants (bruit) (Ester et al. 1996).

### 2.2.9 Algorithmes de réduction de dimension

Les méthodes de réduction de dimension transforment les données afin de conserver l'information essentielle dans un espace plus compact. Elles facilitent la visualisation, réduisent la redondance et peuvent améliorer l'efficacité des modèles.

#### 2.2.9.1 PCA (Analyse en composantes principales)

PCA transforme les données en un ensemble de composantes principales en maximisant la variance. Elle permet de réduire la dimension tout en conservant l'essentiel de l'information (Jolliffe 2002).

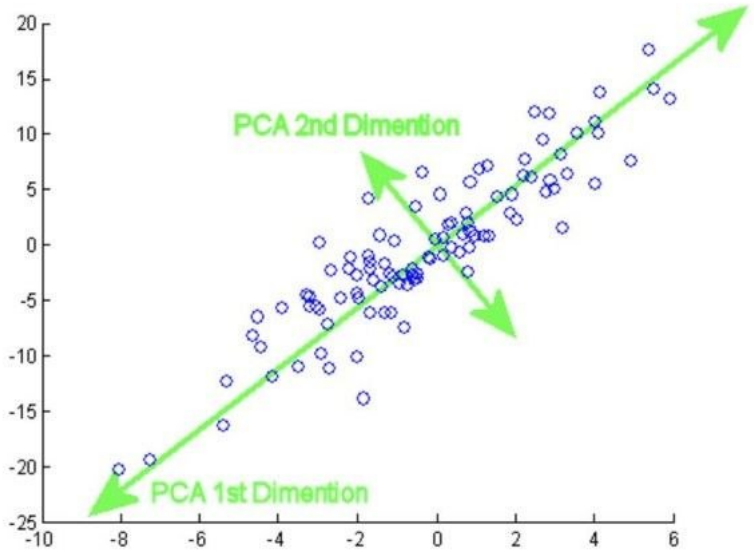



**Fig 2.15 :** Projection des données avec PCA (adapté de Jolliffe (2002))

**Explication :** La figure 2.15 illustre la projection des données dans un nouvel espace où PCA maximise la variance expliquée (Jolliffe 2002).

#### 2.2.9.2 t-SNE

t-SNE est une technique de visualisation qui projette les données dans un espace de faible dimension. Elle est efficace pour représenter des structures complexes mais coûteuse en calcul (Maaten et G. Hinton 2008).

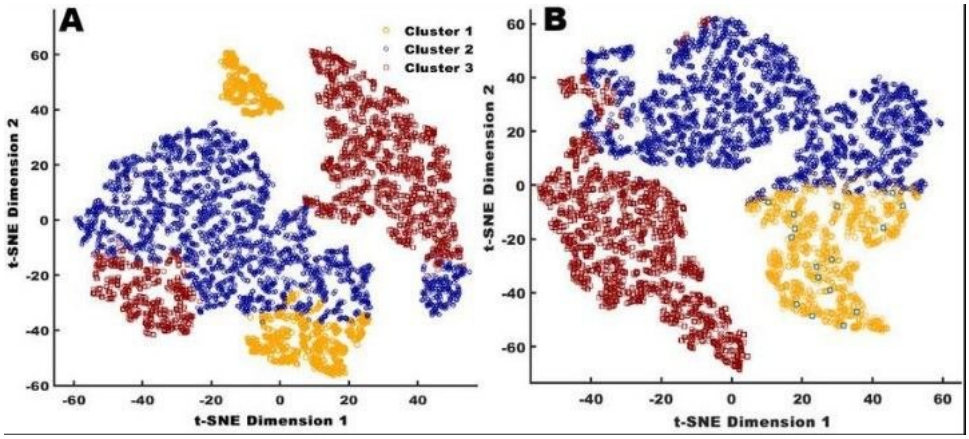



**Fig 2.16 :** Visualisation des données avec t-SNE (adapté de Maaten et G. Hinton (2008))

**Explication :** Comme l'illustre la figure 2.16, t-SNE est utilisé pour visualiser des données de haute dimension en 2D ou 3D en conservant les relations locales (Maaten et G. Hinton 2008).

#### 2.2.9.3 Autoencoders

Les autoencoders sont des réseaux de neurones utilisés pour apprendre une représentation compacte des données. Ils sont utilisés pour la compression, la réduction de dimension et la détection d'anomalies (Goodfellow, Bengio et Courville 2016).

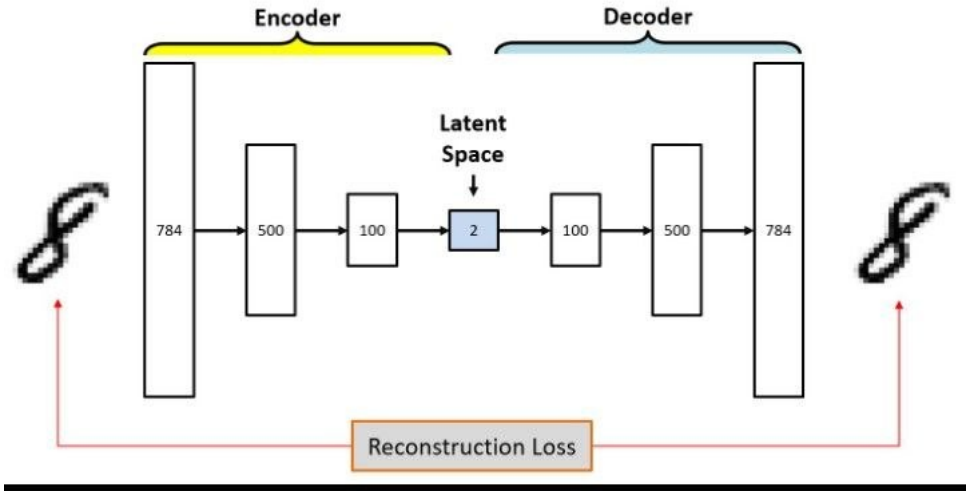



**Fig 2.17 :** Architecture d'un autoencoder (adapté de Goodfellow, Bengio et Courville (2016))

**Explication :** La figure 2.17 présente l'architecture d'un autoencoder, dans laquelle le réseau apprend une représentation compressée des données (Goodfellow, Bengio

et Courville 2016).

Le tableau 2.2 compare les deux grandes familles de méthodes non supervisées présentées précédemment.

**Tab 2.2 :** Comparaison entre clustering et réduction de dimension

| Critère | Clustering | Réduction de dimension |
|---|---|---|
| Objectif | Regrouper les données | Réduire les variables |
| Sortie | Groupes (clusters) | Nouvelles variables |
| Exemple | Segmentation clients | Visualisation données |
| Nature | Structure cachée | Compression information |
| Algorithmes | K-Means, DBSCAN | PCA, t-SNE |
| Utilisation | Analyse exploratoire | Prétraitement |

Le tableau 2.3 synthétise les différences globales entre apprentissage supervisé et apprentissage non supervisé.

**Tab 2.3 :** Comparaison globale entre apprentissage supervisé et non supervisé

| Critère | Apprentissage supervisé | Apprentissage non supervisé |
|---|---|---|
| Objectif | Prédire une sortie connue | Découvrir des structures cachées |
| Données | Données étiquetées (input + output) | Données non étiquetées |
| Sortie | Classe ou valeur numérique | Clusters ou nouvelles représentations |
| Type de problème | Classification / Régression | Clustering / Réduction de dimension |
| Exemples | Diagnostic médical, prédiction de prix | Segmentation clients, analyse exploratoire |
| Algorithmes | SVM (Vapnik 1995), KNN (Cover et Hart 1967), Régression linéaire (Montgomery, Peck et Vining 2012), Random Forest (Breiman 2001) | K-Means (MacQueen 1967), DBSCAN (Ester et al. 1996), PCA (Jolliffe 2002), t-SNE (Maaten et G. Hinton 2008) |
| Évaluation | Accuracy, MSE, RMSE | Silhouette score, visualisation |
| Utilisation | Prédiction | Exploration des données |

## 2.3 Méthodes de Deep Learning

Les méthodes de Deep Learning regroupent plusieurs architectures adaptées à différents types de données. Certaines sont conçues pour des données tabulaires, d'autres pour les images, les séries temporelles ou les séquences, ce qui justifie leur présentation séparée.

### 2.3.1 Réseaux de neurones artificiels (ANN)

Les méthodes de Deep Learning reposent sur des architectures de réseaux de neurones profonds capables d'extraire automatiquement des représentations hiérarchiques à partir des données brutes. Contrairement aux méthodes classiques, ces modèles apprennent des caractéristiques pertinentes sans intervention humaine explicite, ce qui les rend particulièrement efficaces pour les tâches complexes telles

que la vision par ordinateur, le traitement du langage naturel et la prédiction médicale (Goodfellow, Bengio et Courville 2016 ; LeCun, Bengio et G. Hinton 2015).

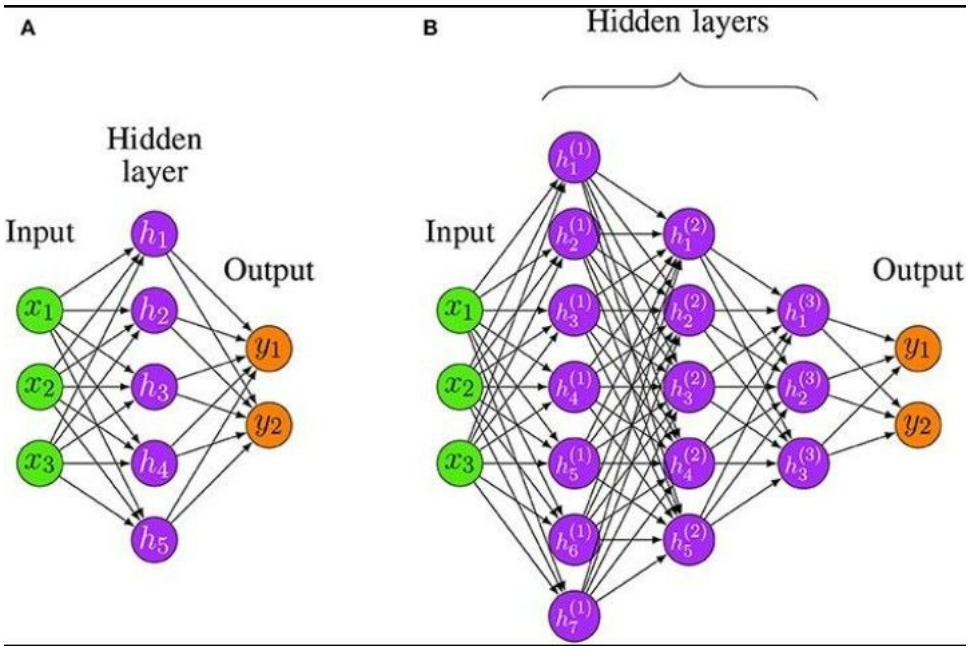



**Fig 2.18 :** Réseau de neurones artificiel (adapté de Goodfellow, Bengio et Courville (2016) et LeCun, Bengio et G. Hinton (2015))

**Explication :** La figure 2.18 présente une architecture générale de réseau de neurones artificiel. Les réseaux de neurones artificiels (ANN), notamment les perceptrons multicouches, sont composés de couches de neurones interconnectés. Chaque neurone applique une transformation suivie d'une fonction d'activation non linéaire. L'apprentissage repose sur l'algorithme de rétropropagation qui ajuste les poids du réseau en minimisant une fonction de perte. Ces modèles permettent d'approximer des relations complexes entre les variables (Goodfellow, Bengio et Courville 2016 ; LeCun, Bengio et G. Hinton 2015).

### 2.3.2 Réseaux de neurones convolutifs (CNN)

Les CNN sont particulièrement adaptés aux données présentant une structure spatiale, comme les images médicales, car ils apprennent des motifs locaux à l'aide de filtres convolutionnels.

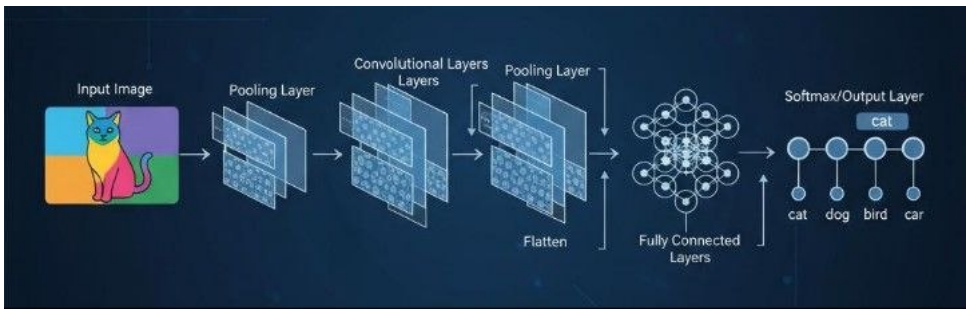



**Fig 2.19 :** Architecture d'un CNN (adapté de LeCun, Bengio et G. Hinton (2015))

**Explication :** La figure 2.19 illustre l'organisation typique d'un CNN. Les CNN sont conçus pour traiter des données structurées comme les images. Ils utilisent des filtres pour extraire des caractéristiques locales, suivis de couches de pooling pour réduire la dimension tout en conservant l'information essentielle (LeCun, Bengio et G. Hinton 2015).

### 2.3.3 Réseaux de neurones récurrents (RNN)

Les RNN sont conçus pour traiter des données séquentielles en tenant compte de l'ordre des observations. Ils sont donc pertinents pour des signaux ou séries temporelles où l'information précédente influence l'interprétation de l'état courant.

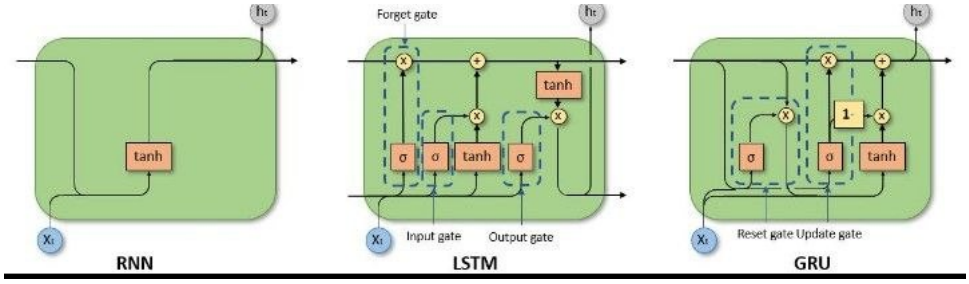



**Fig 2.20 :** Réseau de neurones récurrent (adapté de Goodfellow, Bengio et Courville (2016) et Sarker (2021))

**Explication :** La figure 2.20 met en évidence le principe de récurrence utilisé par les RNN. Les RNN sont adaptés aux données séquentielles. Ils introduisent des connexions récurrentes permettant de conserver une information sur les états précédents. Toutefois, ils peuvent rencontrer le problème du gradient disparaissant lors de l'apprentissage de longues dépendances (Goodfellow, Bengio et Courville 2016 ; Sarker 2021).

### 2.3.4 Long Short-Term Memory (LSTM)

Les LSTM constituent une variante des RNN destinée à mieux conserver l'information sur de longues séquences. Leur mécanisme de portes permet de contrôler ce qui doit être mémorisé, oublié ou transmis.

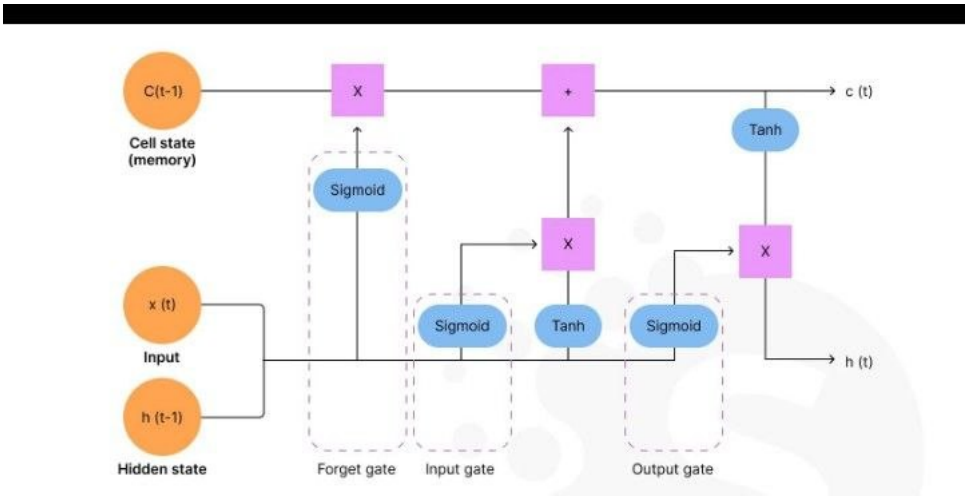



**Fig 2.21 :** Cellule LSTM (adapté de Goodfellow, Bengio et Courville (2016) et Sarker (2021))

**Explication :** La figure 2.21 détaille la structure interne d'une cellule LSTM. Les LSTM sont une extension des RNN permettant de mieux gérer les dépendances à long terme. Grâce à un système de portes, ils contrôlent le flux d'information et réduisent les problèmes liés au gradient disparaissant (Goodfellow, Bengio et Courville 2016 ; Sarker 2021).

### 2.3.5 Autoencoders

Les autoencoders apprennent à reconstruire leurs entrées après les avoir compressées dans une représentation latente. Cette propriété les rend utiles pour la réduction de dimension, la détection d'anomalies et l'apprentissage non supervisé.

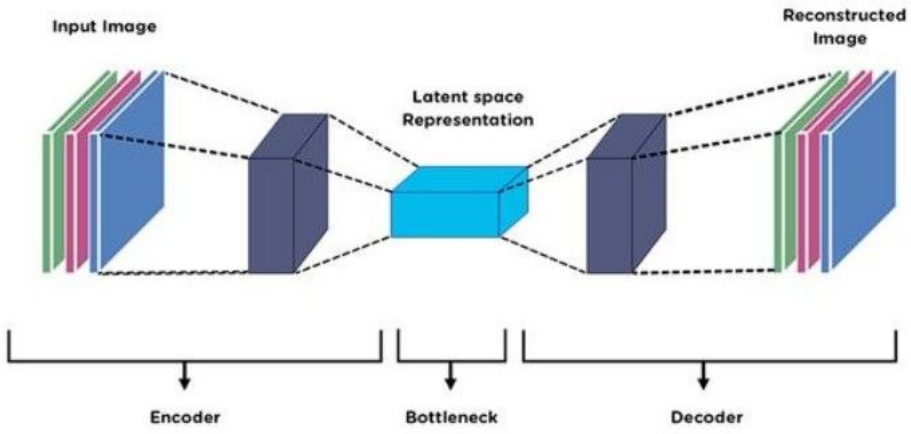



**Fig 2.22 :** Autoencoder (adapté de Goodfellow, Bengio et Courville (2016))

**Explication :** La figure 2.22 met en évidence le couple encodeur-décodeur caractéristique des autoencoders. Les autoencoders sont des modèles d'apprentissage non supervisé qui apprennent une représentation compressée des données via un encodeur et un décodeur. Leur objectif est de minimiser l'erreur de reconstruction entre l'entrée et la sortie (Goodfellow, Bengio et Courville 2016).

### 2.3.6 Transformers

Les Transformers reposent sur le mécanisme d'attention, qui permet de pondérer l'importance des éléments d'une séquence. Ils sont largement utilisés pour modéliser des dépendances complexes entre observations.

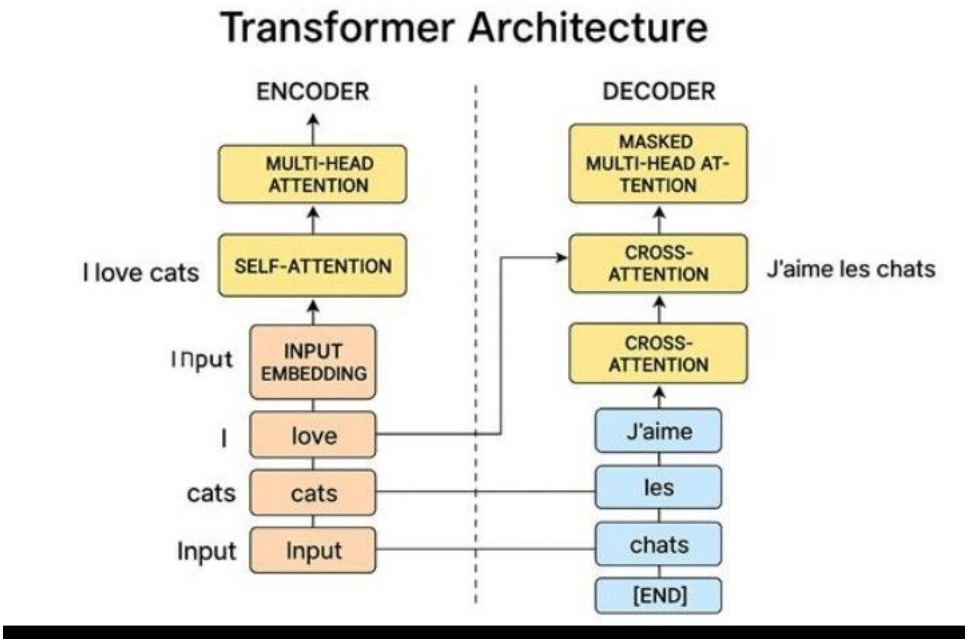



**Fig 2.23 :** Architecture Transformer (adapté de Vaswani et al. (2017))

**Explication :** La figure 2.23 présente l'architecture Transformer et son organisation autour du mécanisme d'attention. Les Transformers utilisent un mécanisme d'attention pour modéliser les relations entre les éléments d'une séquence. Contrairement aux RNN, ils permettent un traitement parallèle des données, ce qui améliore l'efficacité et les performances sur de nombreuses tâches (Vaswani et al. 2017).

### 2.3.7 Comparaison des méthodes de Deep Learning

Le tableau 2.4 compare les architectures présentées selon le type de données généralement traité, leurs principaux avantages et leurs limites.

**Tab 2.4 :** Comparaison des méthodes de Deep Learning

| Méthode | Type de données | Avantages | Limites |
|---|---|---|---|
| ANN | Données générales | Modélisation flexible | Limité sur données complexes |
| CNN | Images | Extraction automatique | Besoin de données volumineuses |
| RNN | Séquences | Modélisation temporelle | Gradient disparaissant |
| LSTM | Séquences longues | Mémoire longue durée | Complexité |
| Autoencoder | Données complexes | Compression | Perte d'information |
| Transformer | Texte | Attention globale | Coût élevé |

### 2.3.8 Fonction de perte (Loss function)

La fonction de perte sert de critère quantitatif pour mesurer l'écart entre les prédictions du modèle et les valeurs attendues. Elle constitue le signal utilisé par l'algorithme d'optimisation pour ajuster les paramètres du réseau.

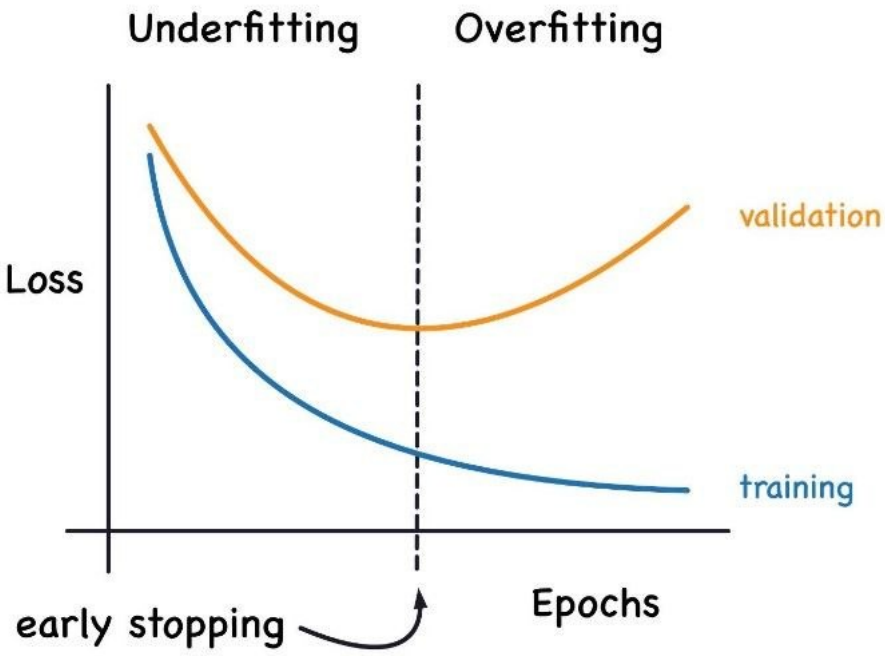



**Fig 2.24 :** Fonction de perte (adapté de Goodfellow, Bengio et Courville (2016) et Chollet (2021))

#### 2.3.8.1 Définition :

La figure 2.24 schématise ce principe : la fonction de perte (Loss function) est une mesure permettant de quantifier l'écart entre les prédictions du modèle et les valeurs réelles. Elle joue un rôle central dans l'apprentissage des modèles de Deep Learning. Le processus d'entraînement consiste à minimiser cette fonction afin d'améliorer la précision du modèle. Cette minimisation est réalisée à l'aide d'algorithmes d'optimisation tels que la descente de gradient (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

#### 2.3.8.2 Formulation mathématique :

De manière générale, la perte totale correspond à la moyenne des pertes calculées sur les exemples d'entraînement.

$$L(\theta) = \frac{1}{n} \sum_{i=1}^{n} \ell(y_i, \hat{y}_i)$$

où $y_i$ désigne la valeur réelle, $\hat{y}_i$ la valeur prédite, $\ell$ la fonction de perte élémentaire et $n$ le nombre d'exemples. Cette formulation générale permet de représenter la majorité des fonctions de perte utilisées en apprentissage automatique (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

#### 2.3.8.3 Types de fonctions de perte

Le choix de la fonction de perte dépend de la nature de la tâche, notamment selon qu'il s'agit d'un problème de régression, de classification binaire ou de classification multiclasse.

**a) Mean Squared Error (MSE) :**

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Cette fonction est utilisée pour les régression et pénalise fortement les grandes erreurs (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

**b) Mean Absolute Error (MAE) :**

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Elle est plus robuste face aux valeurs aberrantes que MSE (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

**c) Binary Cross-Entropy :**

$$L = -\frac{1}{n} \sum_{i=1}^{n} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Utilisée pour les problèmes de classification binaire (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

**d) Categorical Cross-Entropy :**

$$L = -\sum_{i=1}^{n} y_i \log(\hat{y}_i)$$

Adaptée aux problèmes multiclasse (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

#### 2.3.8.4 Rôle dans l'apprentissage

La fonction de perte guide l'optimisation du modèle en indiquant la direction dans laquelle ajuster les paramètres. Une diminution progressive de la perte indique une amélioration des performances. Elle permet également :

d'évaluer la qualité du modèle, de suivre la convergence et d'ajuster les hyperparamètres (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

#### 2.3.8.5 Surapprentissage et sous-apprentissage

Overfitting : perte faible sur les données d'entraînement mais élevée sur validation

Underfitting : perte élevée dans les deux cas Ces phénomènes peuvent être analysés à partir des courbes de perte (Goodfellow, Bengio et Courville 2016 ; Chollet 2021).

#### 2.3.8.6 Techniques d'amélioration

Plusieurs techniques peuvent être utilisées pour améliorer la généralisation et stabiliser l'entraînement des modèles profonds. Elles agissent soit sur la complexité du modèle, soit sur le processus d'apprentissage, soit sur la diversité des données.

**a) Dropout :** Le dropout consiste à désactiver aléatoirement une fraction des neurones pendant l'entraînement.

$$h_i^{(drop)} = h_i \cdot r_i, \qquad r_i \sim \text{Bernoulli}(p)$$

**Effets :** réduction du surapprentissage amélioration de la généralisation

**b) Régularisation L1 / L2 :**

$$L_{total} = L + \lambda \sum |w| \quad (L1)$$

$$L_{total} = L + \lambda \sum w^2 \quad (L2)$$

**Objectif :** limiter la complexité éviter le surapprentissage

**c) Early Stopping :** Technique consistant à arrêter l'entraînement lorsque la perte sur validation augmente. Avantages : amélioration de la généralisation réduction du temps de calcul

**d) Batch Normalization :**

$$x_{norm} = \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}}$$

**Effets :** stabilisation de l'apprentissage accélération de la convergence

**e) Data Augmentation :** Technique consistant à générer de nouvelles données à partir des données existantes.

**Objectifs :** augmenter la diversité améliorer la robustesse

## 2.4 Interprétabilité et intelligence artificielle explicable

Les modèles d'apprentissage automatique et d'apprentissage profond peuvent atteindre de bonnes performances prédictives, mais leur complexité rend parfois difficile la compréhension des raisons qui conduisent à une décision. Cette difficulté est généralement désignée par le problème de la « boîte noire ». L'intelligence artificielle explicable, ou XAI (*Explainable Artificial Intelligence*), regroupe les méthodes qui visent à rendre le comportement d'un modèle plus compréhensible pour l'utilisateur, sans se limiter au seul résultat numérique de la prédiction (Doshi-Velez et Kim 2017 ; Arrieta et al. 2020).

L'interprétabilité désigne la capacité à comprendre le fonctionnement ou la logique d'un modèle, tandis que l'explicabilité renvoie plutôt à la production d'explications utilisables après ou pendant la prédiction. On distingue souvent deux ni-

veaux complémentaires. L'explication globale décrit le comportement moyen du modèle sur l'ensemble des données, par exemple en identifiant les variables les plus influentes. L'explication locale s'intéresse à une observation précise et cherche à montrer pourquoi le modèle a produit une prédiction particulière. Ces deux niveaux sont utiles : le premier aide à contrôler la cohérence générale du modèle, tandis que le second facilite l'analyse d'un cas individuel.

Une autre distinction importante oppose les modèles intrinsèquement interprétables aux méthodes post-hoc. Les premiers sont conçus pour être directement lisibles, comme la régression logistique, les arbres de décision simples ou certains modèles à mécanisme d'attention. Les secondes sont appliquées après l'entraînement pour expliquer un modèle déjà appris, y compris lorsqu'il est complexe. Dans les domaines à enjeu élevé, comme la santé, cette distinction est importante, car l'explication ne doit pas donner une impression trompeuse de certitude : elle aide à analyser la décision, mais ne prouve pas à elle seule que le modèle est correct ou causalement valide (Rudin 2019 ; Rajkomar, Dean et Kohane 2019).

Le tableau 2.5 résume quelques familles de méthodes couramment utilisées.

**Tab 2.5 :** Principales familles de méthodes d'interprétabilité et d'explicabilité

| Famille | Principe | Intérêt principal |
|---|---|---|
| Modèles intrinsèquement interprétables | Utiliser une structure directement lisible, comme une régression, un arbre simple ou des règles. | Compréhension directe du modèle, utile lorsque la transparence est prioritaire. |
| Méthodes locales par substitution | Approximer localement un modèle complexe par un modèle simple autour d'une observation, comme LIME. | Expliquer une prédiction individuelle de manière agnostique au modèle (Ribeiro, Singh et Guestrin 2016). |
| Méthodes par attribution | Attribuer à chaque variable une contribution positive ou négative à la prédiction, comme les valeurs SHAP. | Relier les variables d'entrée à la sortie du modèle, localement et globalement (Lundberg et Lee 2017). |
| Mécanismes d'attention | Pondérer les variables ou parties de l'entrée utilisées par le modèle lors de la décision. | Identifier les informations mises en avant par certaines architectures, notamment TabNet (Arik et Pfister 2021). |
| Explications contrefactuelles | Rechercher les changements minimaux qui auraient modifié la décision du modèle. | Rendre l'explication plus actionnable, sous réserve de plausibilité métier (Arrieta et al. 2020). |

Dans le cadre de ce mémoire, l'objectif de l'explicabilité n'est pas de remplacer l'expertise médicale, mais de vérifier si les modèles exploitent des facteurs cohérents avec les connaissances cliniques et d'accompagner la lecture des prédictions. Deux approches sont donc retenues dans l'étude expérimentale : les valeurs SHAP pour expliquer les modèles MLP et ResNetTab, et les masques d'attention pour analyser TabNet. Ces méthodes seront détaillées dans le chapitre suivant, où elles seront appliquées aux variables du jeu de données cardiovasculaire.

## 2.5 Évaluation des modèles

L'évaluation des modèles permet de mesurer leur performance sur des données non utilisées pendant l'entraînement. Dans un contexte médical, elle doit tenir compte non seulement du taux global de bonnes prédictions, mais aussi de la capacité du modèle à limiter les erreurs cliniquement importantes, notamment les faux négatifs.

### 2.5.1 Matrice de confusion

La matrice de confusion permet de comparer les prédictions du modèle aux valeurs réelles. Elle se compose de quatre éléments : TP (True Positive) TN (True Negative) FP (False Positive) FN (False Negative)

Elle constitue la base de calcul des principales métriques d'évaluation (Géron 2022).

### 2.5.2 Accuracy (Exactitude)

L'accuracy mesure la proportion totale de prédictions correctes parmi l'ensemble des observations évaluées.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Cette métrique mesure la proportion de prédictions correctes. Toutefois, elle peut être trompeuse en présence de données déséquilibrées (Géron 2022).

### 2.5.3 Précision et Recall (Sensibilité)

La précision et le recall décrivent deux aspects complémentaires de la performance sur la classe positive.

$$\text{Precision} = \frac{TP}{TP + FP}, \qquad \text{Recall} = \frac{TP}{TP + FN}$$

La précision évalue la fiabilité des prédictions positives, tandis que la sensibilité mesure la capacité à détecter les cas positifs. Dans le domaine médical, la sensibilité est cruciale afin de réduire les faux négatifs (patients non détectés) (Géron 2022).

### 2.5.4 F1-score

Le F1-score combine la précision et le recall en une seule mesure, ce qui permet d'évaluer leur équilibre.

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Il représente un compromis entre précision et recall, particulièrement utile lorsque les classes sont déséquilibrées (Géron 2022).

### 2.5.5 Spécificité

La spécificité évalue la capacité du modèle à reconnaître correctement les cas négatifs.

$$\text{Specificite} = \frac{TN}{TN + FP}$$

Elle mesure la capacité du modèle à identifier correctement les cas négatifs (individus sains). En complément de la sensibilité, elle permet une évaluation complète en diagnostic médical (Géron 2022).

### 2.5.6 Courbe ROC et AUC

La courbe ROC analyse le compromis entre la sensibilité et le taux de faux positifs pour différents seuils de décision. L'AUC synthétise cette capacité de discrimination sur l'ensemble des seuils, comme l'illustre la figure 2.25.

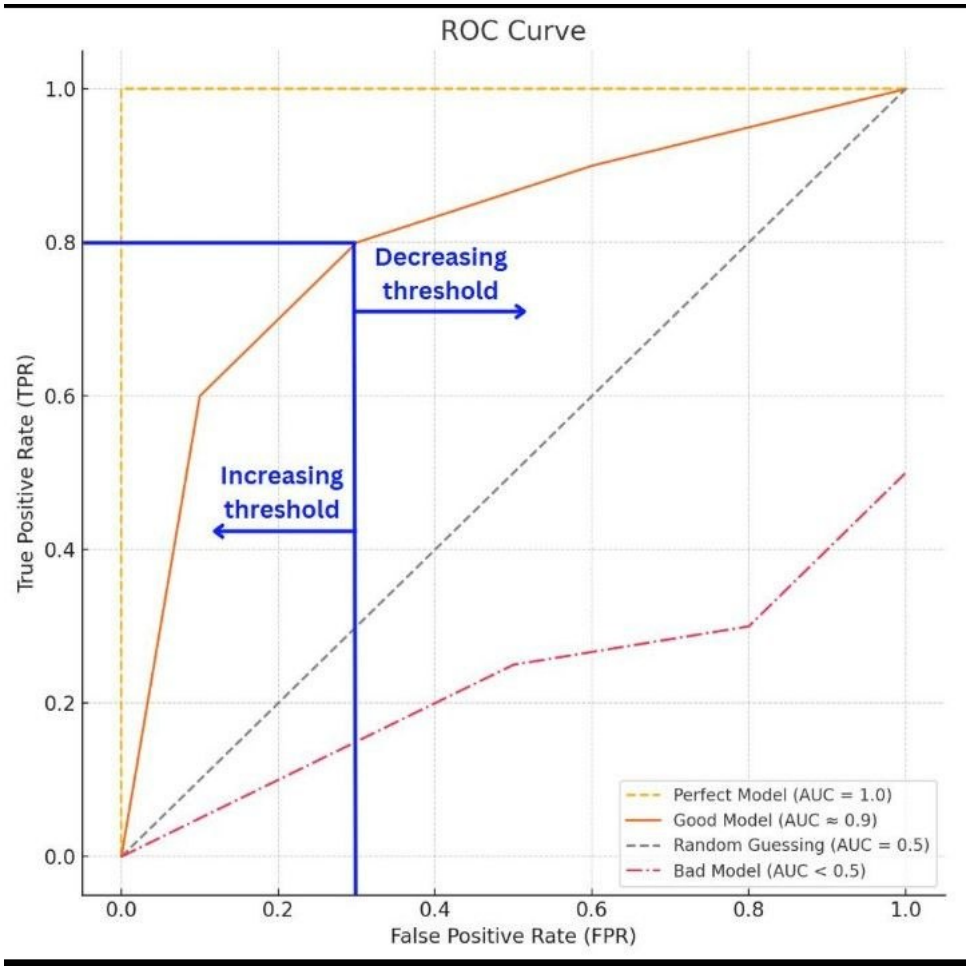



**Fig 2.25 :** Courbe ROC (adapté de Géron (2022))

**Explication :**

$$TPR = \frac{TP}{TP + FN}, \qquad FPR = \frac{FP}{FP + TN}$$

La courbe ROC représente le taux de vrais positifs (TPR) en fonction du taux de faux positifs (FPR). Une valeur d'AUC proche de 1 indique un modèle performant, tandis qu'une valeur proche de 0,5 correspond à un modèle aléatoire.

## 2.6 Conclusion

Dans ce chapitre, nous avons présenté les principales méthodes de Deep Learning utilisées pour la prédiction, en mettant en évidence leurs caractéristiques ainsi que leurs domaines d'application. Les architectures telles que les réseaux de neurones artificiels (ANN), les réseaux de neurones convolutifs (CNN) et les réseaux de neurones récurrents (RNN) jouent un rôle essentiel dans le traitement des données complexes, particulièrement dans le domaine médical où elles excellent dans l'analyse d'images médicales, de signaux temporels et de données

structurées (Goodfellow, Bengio et Courville 2016 ; LeCun, Bengio et G. Hinton 2015 ; Sarker 2021). Nous avons également exploré les approches d'apprentissage supervisé et non supervisé. La première permet d'accomplir des tâches de classification et de régression à partir de données étiquetées, tandis que la seconde vise à découvrir des structures et des patterns cachés dans des données non étiquetées. Ces paradigmes constituent un socle fondamental pour l'analyse et l'exploitation des données médicales (Bishop 2006 ; Hastie, Tibshirani et Friedman 2009 ; Géron 2022). Par ailleurs, l'évaluation des modèles a été abordée à travers différentes métriques de performance telles que l'accuracy, la précision, le recall (sensibilité), le F1-score et la courbe ROC (avec son aire sous la courbe - AUC). Ces indicateurs permettent une évaluation rigoureuse et multidimensionnelle de la performance des modèles, tout en mettant en lumière leurs forces et leurs limites. Nous avons également introduit les notions d'interprétabilité et d'intelligence artificielle explicable, indispensables pour analyser les décisions des modèles complexes et préparer l'étude des méthodes SHAP et des mécanismes d'attention utilisées dans le chapitre expérimental (Doshi-Velez et Kim 2017 ; Arrieta et al. 2020). Dans le contexte spécifique de la prédiction des maladies cardiaques, le choix des métriques d'évaluation revêt une importance particulière. La minimisation des faux négatifs est essentielle afin d'éviter des diagnostics manqués aux conséquences potentiellement graves pour le patient. Ainsi, la sensibilité (recall) constitue un indicateur clé dans ce type d'application médicale (Esteva et al. 2019 ; Zhou et al. 2024). Enfin, les concepts théoriques et méthodologiques exposés dans ce chapitre fournissent une base solide pour le développement du modèle proposé dans le chapitre suivant. Ce dernier mettra en œuvre ces approches sur des données réelles afin de concevoir un système de prédiction fiable, performant et adapté aux exigences cliniques (Goodfellow, Bengio et Courville 2016 ; Sarker 2021).

CHAPITRE 3

# Méthodologie, Expérimentation et Résultats

## 3.1 Introduction

Ce chapitre présente la mise en œuvre expérimentale du système de prédiction des maladies cardiovasculaires. Il traduit les notions théoriques exposées précédemment en un protocole complet, allant de la préparation des données jusqu'à l'évaluation comparative des modèles. L'objectif principal est d'analyser dans quelle mesure des architectures d'apprentissage profond peuvent être exploitées sur des données cliniques tabulaires, tout en tenant compte des contraintes propres au domaine médical, notamment le déséquilibre des classes et l'importance de la sensibilité.

L'étude s'appuie sur le jeu de données public *Heart_Disease_2024*, disponible sur Kaggle (Tajnim 2024) et associé au BIG-Dataset décrit par Haque et al. (2024), composé de variables démographiques et cliniques associées à une cible binaire indiquant la présence ou l'absence de maladie cardiovasculaire. Après le prétraitement, plusieurs modèles sont comparés : deux modèles de référence, la régression logistique et XGBoost (Chen et Guestrin 2016), ainsi que trois architectures profondes, MLP, ResNetTab (Gorishniy et al. 2021) et TabNet (Arik et Pfister 2021). Le chapitre détaille également la stratégie d'entraînement, l'optimisation des hyperparamètres, le choix des seuils de décision et les métriques utilisées pour évaluer les performances.

Au-delà des scores globaux, l'analyse porte sur les matrices de confusion, les courbes ROC et précision-rappel, la calibration des probabilités et l'interprétabilité des modèles. Les valeurs SHAP (Lundberg et Lee 2017) et les masques d'attention permettent d'examiner les variables qui influencent le plus les prédictions. Enfin, un prototype interactif illustre la manière dont les modèles peuvent être présentés dans

une interface de démonstration destinée à l'analyse et à l'explication des résultats.

## 3.2 Méthodologie expérimentale

Cette section décrit le protocole expérimental retenu, depuis la caractérisation du jeu de données jusqu'à la configuration d'exécution. Elle regroupe les choix de préparation des données, les modèles comparés, les paramètres d'entraînement, les critères d'évaluation et les méthodes d'interprétabilité utilisées avant l'analyse des résultats.

### 3.2.1 Description du jeu de données (dataset)

Les expériences menées dans le cadre de ce travail reposent sur le jeu de données public *Heart_Disease_2024* (Tajnim 2024), qui correspond au BIG-Dataset présenté de manière formelle par Haque et al. (2024). Il est important de préciser que ce jeu de données est un ensemble composite fusionnant deux sources historiques majeures. Les variables physiologiques et de mode de vie (telles que l'IMC, la fréquence cardiaque et le statut tabagique) sont issues de la cohorte de l'étude de Framingham (D'Agostino et al. 2008). En revanche, les variables d'examen clinique et diagnostique (notamment *oldpeak*, le type de douleur thoracique *cp*, les résultats de l'ECG au repos *restecg*, l'angine induite par l'effort *exang*, le cholestérol *chol* et la pression artérielle au repos *trestbps*) proviennent historiquement du jeu de données de Cleveland (Cleveland Heart Disease Dataset) (Datta et Gang 2021). Le jeu de données contient 4 238 observations de patients, chacune décrite par 14 variables cliniques et démographiques, ainsi qu'une variable cible binaire indiquant la présence ou l'absence de maladie cardiovasculaire.

Les variables d'entrée se répartissent en deux catégories. La première regroupe huit variables continues : l'âge du patient (en années), la dépression du segment ST induite par l'effort par rapport au repos (`oldpeak`, en millimètres), le taux de cholestérol sérique (en mg/dL), la pression artérielle systolique au repos (en mmHg), la pression artérielle diastolique (en mmHg), l'indice de masse corporelle (en $kg/m^2$), la fréquence cardiaque maximale atteinte (en battements par minute), et le taux de glucose sanguin (en mg/dL). La seconde catégorie comprend six variables binaires ou catégorielles : le sexe du patient, le statut de fumeur actuel, la glycémie à jeun supérieure à 120 mg/dL, le type de douleur thoracique (variable catégorielle à cinq modalités), les résultats de l'électrocardiogramme au repos (variable catégorielle à trois modalités), et la présence d'angine de poitrine induite par l'effort.

La variable cible prend la valeur 1 lorsque le patient présente une maladie cardiovasculaire et 0 dans le cas contraire. La distribution des classes est nettement déséquilibrée : 3 594 observations (84,8 %) appartiennent à la classe négative et 644 (15,2 %) à la classe positive, soit un ratio de déséquilibre d'environ 5,6 pour 1. Ce déséquilibre constitue un défi majeur pour l'apprentissage, car les modèles tendent à favoriser la classe majoritaire au détriment de la détection des cas positifs.

Le tableau 3.1 résume les caractéristiques principales du jeu de données et le tableau 3.2 détaille l'ensemble des variables utilisées.

**Tab 3.1 :** Caractéristiques générales du jeu de données

| Caractéristique | Valeur |
|---|---|
| Nombre d'observations | 4 238 |
| Nombre de variables d'entrée | 14 (8 continues, 6 binaires/catégorielles) |
| Variable cible | Présence de maladie cardiovasculaire (binaire) |
| Classe 0 (absence de maladie) | 3 594 (84,8 %) |
| Classe 1 (présence de maladie) | 644 (15,2 %) |
| Ratio de déséquilibre | 5,6 : 1 |

**Tab 3.2 :** Description des variables du jeu de données

| Variable | Type | Description |
|---|---|---|
| `age` | Continue | Âge du patient (années) |
| `oldpeak` | Continue | Dépression du segment ST induite par l'effort (mm) |
| `chol` | Continue | Cholestérol sérique (mg/dL) |
| `trestbps` | Continue | Pression artérielle systolique au repos (mmHg) |
| `diaBP` | Continue | Pression artérielle diastolique (mmHg) |
| `BMI` | Continue | Indice de masse corporelle (kg/m$^2$) |
| `heartRate` | Continue | Fréquence cardiaque maximale (bpm) |
| `glucose` | Continue | Glucose sanguin (mg/dL) |
| `sex` | Binaire | Sexe du patient (0 = féminin, 1 = masculin) |
| `Smkr` | Binaire | Fumeur actuel (0 = non, 1 = oui) |
| `fbs` | Binaire | Glycémie à jeun > 120 mg/dL (0 = non, 1 = oui) |
| `cp` | Catégorielle | Type de douleur thoracique (5 modalités) |
| `restecg` | Catégorielle | Résultats de l'ECG au repos (3 modalités) |
| `exang` | Binaire | Angine de poitrine à l'effort (0 = non, 1 = oui) |

### 3.2.2 Prétraitement des données

#### 3.2.2.1 Traitement des valeurs aberrantes

Les variables continues peuvent contenir des valeurs extrêmes susceptibles de perturber la convergence des modèles. Pour limiter leur influence, un écrêtage a été appliqué sur chacune des huit variables continues : toute valeur inférieure au 1$^{er}$ percentile ou supérieure au 99$^{e}$ percentile de sa distribution a été ramenée à la borne correspondante. Cette approche conserve la quasi-totalité de la distribution naturelle tout en supprimant les valeurs aberrantes les plus extrêmes.

#### 3.2.2.2 Partition des données

Le jeu de données a été divisé en trois sous-ensembles selon une partition stratifiée qui préserve la proportion relative des deux classes dans chaque sous-ensemble : 70 % pour l'entraînement, 10 % pour la validation, et 20 % pour le test. Cette répartition a été réalisée en deux étapes successives de partitionnement stratifié. L'ensemble d'entraînement contient environ 2 966 observations (avant augmentation), l'ensemble de validation 424 observations, et l'ensemble de test 848 observations.

L'ensemble de test n'intervient à aucun moment dans les phases d'entraînement, de validation ou de sélection des hyperparamètres. Il est strictement réservé à l'évaluation finale et n'a été utilisé qu'une seule fois, après que tous les choix de modélisation ont été arrêtés.

#### 3.2.2.3 Normalisation

Les huit variables continues ont été normalisées par centrage-réduction (standardisation). Les paramètres de normalisation (moyenne et écart-type) ont été estimés exclusivement sur l'ensemble d'entraînement, puis appliqués par simple transformation aux ensembles de validation et de test, afin d'éviter toute fuite d'information. Les variables binaires et catégorielles, dont les valeurs discrètes ne nécessitent pas de mise à l'échelle, n'ont pas été normalisées.

#### 3.2.2.4 Gestion du déséquilibre des classes

Le déséquilibre des classes a été traité par la technique SMOTE (*Synthetic Minority Over-sampling Technique*), proposée par Chawla et al. (2002). SMOTE génère des observations synthétiques de la classe minoritaire en interpolant linéairement

entre des paires d'observations existantes de cette classe dans l'espace des variables. L'augmentation a été appliquée uniquement à l'ensemble d'entraînement, après normalisation. Les ensembles de validation et de test sont restés intacts afin de garantir une évaluation sur des données réelles.

L'interpolation linéaire opérée par SMOTE peut cependant produire des valeurs fractionnaires pour les variables discrètes (par exemple, une valeur de 0,4 pour une variable binaire). Pour corriger cet artefact, les variables binaires ont été arrondies à l'entier le plus proche puis écrêtées dans l'intervalle $[0, 1]$ après l'application de SMOTE. Pour les variables catégorielles encodées numériquement, le même principe de projection vers les modalités valides doit être appliqué lorsqu'elles sont incluses dans l'espace de SMOTE. Cette précaution vise à maintenir la plausibilité clinique des observations synthétiques.

Après augmentation, l'ensemble d'entraînement contient un nombre équilibré d'observations positives et négatives.

### 3.2.3 Modèles de référence

Pour disposer de points de comparaison solides, nous avons utilisé deux modèles d'apprentissage automatique classique comme lignes de base (*baselines*). Leur rôle est de situer les performances des architectures profondes par rapport à des approches plus établies sur les données tabulaires.

#### 3.2.3.1 Régression logistique

La régression logistique a été utilisée comme référence linéaire interprétable. Elle modélise la probabilité d'appartenance à la classe positive par une fonction sigmoïde appliquée à une combinaison linéaire des variables d'entrée, comme le résume la figure 3.1 :

$$P(y = 1 \mid \mathbf{x}) = \sigma(\mathbf{w}^\top \mathbf{x} + b) = \frac{1}{1 + e^{-(\mathbf{w}^\top \mathbf{x} + b)}}.$$

où $\mathbf{w}$ désigne le vecteur de poids, $b$ le biais, et $\sigma$ la fonction sigmoïde. Pour atténuer l'effet du déséquilibre des classes, les poids attribués à chaque classe dans la fonction de perte ont été ajustés de manière inversement proportionnelle à la fréquence de la classe dans l'ensemble d'entraînement. Ainsi, les erreurs commises sur la classe minoritaire sont pénalisées davantage. Ce modèle sert de point de comparaison simple, peu paramétrique et directement interprétable.

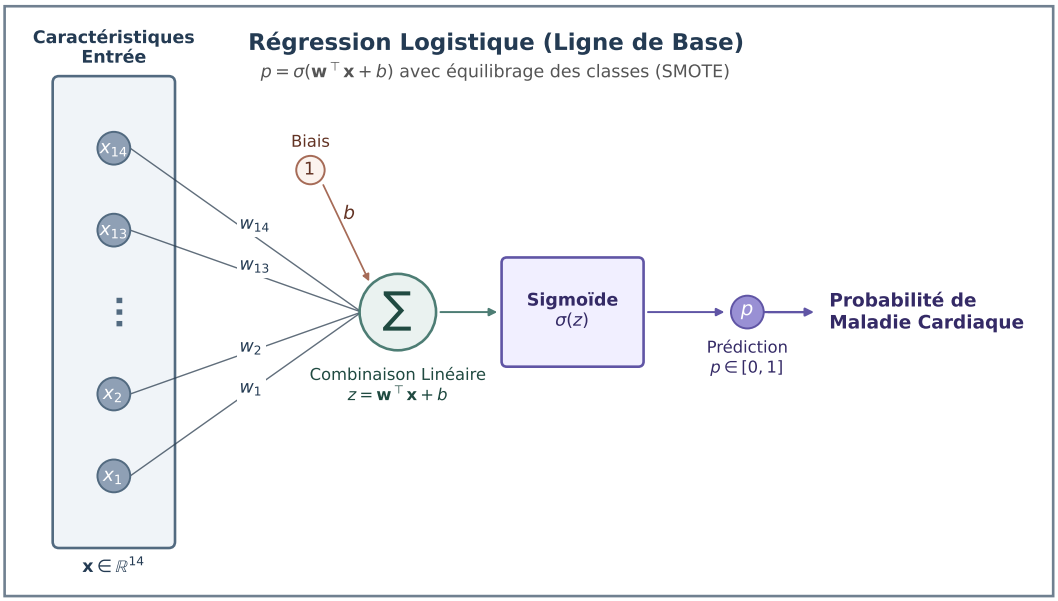



**Fig 3.1 :** Schéma de la régression logistique utilisée comme modèle de référence

#### 3.2.3.2 XGBoost

Nous avons également utilisé XGBoost (*Extreme Gradient Boosting*) (Chen et Guestrin 2016), un algorithme d'apprentissage par ensemble qui construit séquentiellement des arbres de décision. La figure 3.2 illustre ce principe d'addition progressive d'arbres correctifs. Dans la configuration adoptée, chaque nouvel arbre corrige les erreurs résiduelles de l'ensemble précédent par descente de gradient. Le modèle a été configuré avec 300 arbres, une profondeur maximale de 5 niveaux par arbre, et un taux d'apprentissage de 0,05. Le poids relatif de la classe positive a été fixé au ratio du nombre d'observations négatives sur le nombre d'observations positives, afin de compenser le déséquilibre. Ce choix fournit une ligne de base exigeante, car XGBoost est généralement très compétitif sur les données tabulaires.

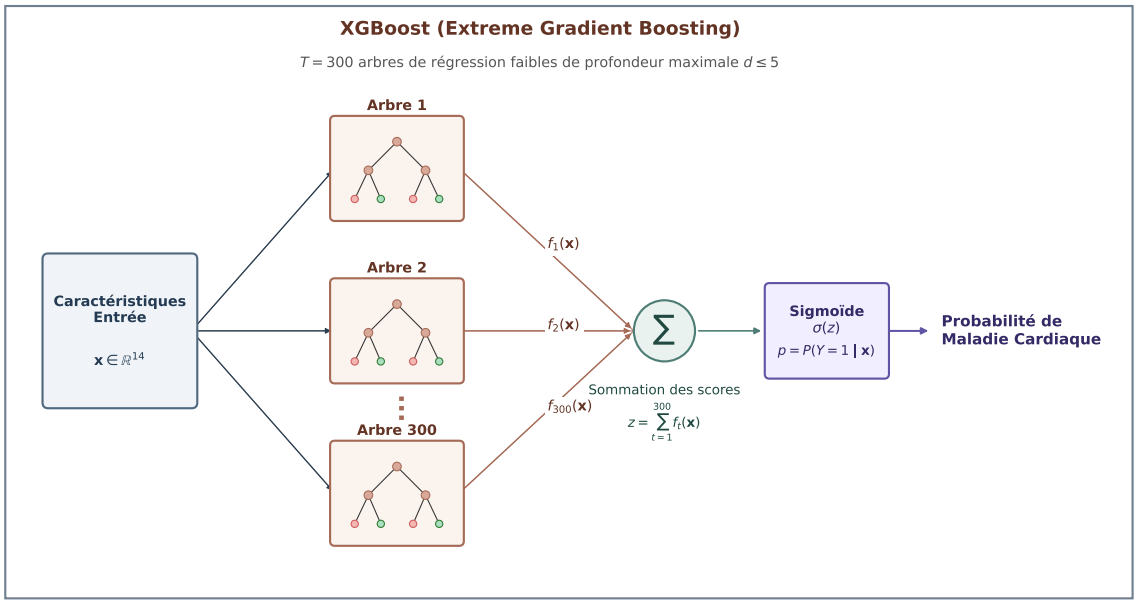



**Fig 3.2 :** Schéma de fonctionnement du modèle XGBoost utilisé

### 3.2.4 Architectures d'apprentissage profond

Nous avons comparé trois architectures d'apprentissage profond adaptées aux données tabulaires. Pour chacune, la description précise la configuration employée dans l'étude, notamment les blocs constitutifs, les choix de régularisation et la couche de sortie.

#### 3.2.4.1 Perceptron multicouche (MLP)

Nous avons d'abord utilisé un perceptron multicouche à propagation avant, composé de couches entièrement connectées. La figure 3.3 illustre l'enchaînement des couches entièrement connectées, de la normalisation, de l'activation, du Dropout et de la tête de sortie.

Dans cette configuration, chaque couche cachée $l \in \{1, \dots, L\}$ enchaîne quatre opérations dans l'ordre suivant :

$$\mathbf{h}_l = \text{Dropout}(\text{GELU}(\text{BN}(\mathbf{W}_l \mathbf{h}_{l-1} + \mathbf{b}_l))) .$$

où $\mathbf{W}_l$ et $\mathbf{b}_l$ désignent les paramètres de la transformation linéaire, BN la normalisation par lots (*Batch Normalization*) (Ioffe et Szegedy 2015), GELU la fonction d'activation, et Dropout la régularisation stochastique.

La fonction d'activation utilisée est GELU (*Gaussian Error Linear Unit*), intro-

duite par Hendrycks et Gimpel (2016), définie par :

$$\mathrm{GELU}(x) = x \cdot \Phi(x),$$

où $\Phi(x)$ désigne la fonction de répartition de la loi normale centrée réduite. Contrairement à ReLU, qui annule brutalement les valeurs négatives, GELU pondère les entrées par leur probabilité d'être positives sous une distribution gaussienne, ce qui favorise un apprentissage plus fluide.

La normalisation par lots est utilisée pour stabiliser et accélérer l'entraînement en réduisant le décalage interne des covariables. Le Dropout désactive aléatoirement une fraction des neurones à chaque itération, ce qui limite le surapprentissage en empêchant le réseau de développer des co-adaptations entre neurones.

Le MLP comprend aussi un mode en entonnoir (*funnel*). Lorsqu'il est activé, la dimension de chaque couche cachée est réduite de moitié par rapport à la précédente, avec un plancher de 32 neurones. Cette configuration compresse progressivement la représentation. La couche finale est une transformation linéaire vers un unique logit de sortie, sans fonction d'activation terminale.

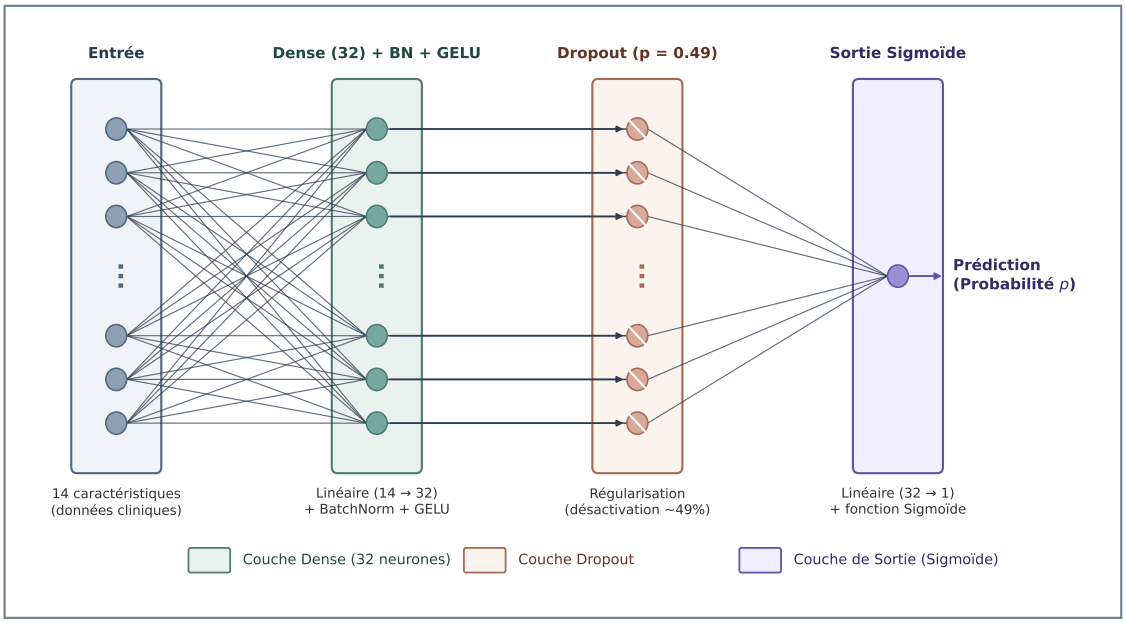



**Fig 3.3 :** Architecture utilisée du perceptron multicouche (MLP)

#### 3.2.4.2 Réseau résiduel tabulaire (ResNetTab)

Nous avons ensuite utilisé ResNetTab, une adaptation des réseaux résiduels aux données tabulaires, inspirée des travaux de Gorishniy et al. (2021). L'architecture utilisée se compose de trois parties.

La figure 3.4 met en évidence la projection initiale, les blocs résiduels successifs et la tête de classification finale.

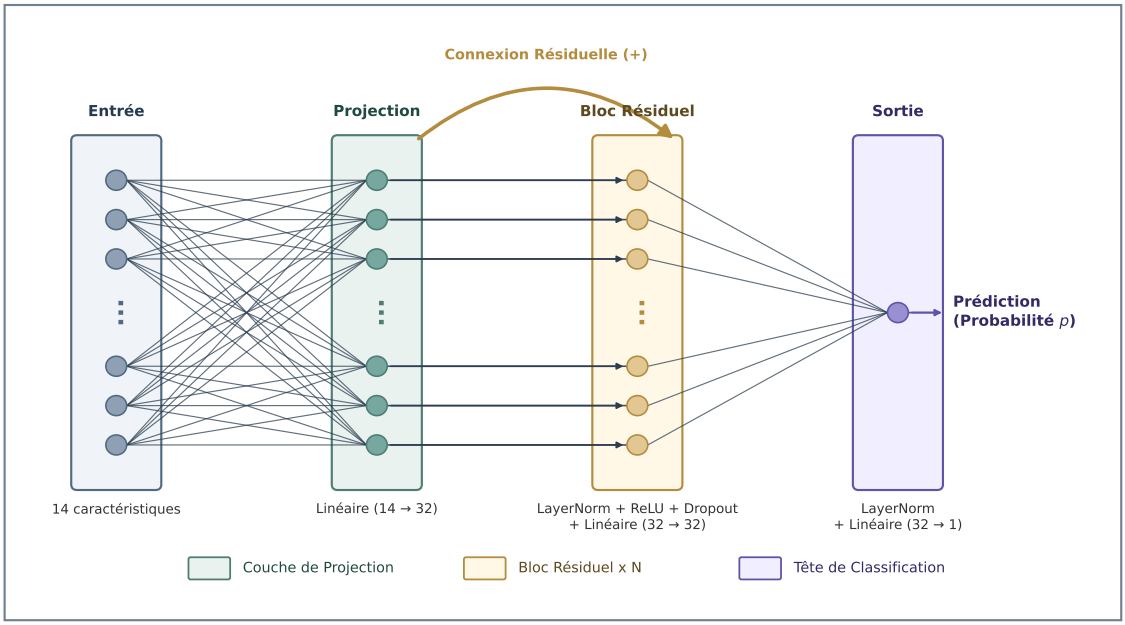



**Fig 3.4 :** Architecture retenue du réseau résiduel tabulaire ResNetTab

La première partie est une couche de projection linéaire qui transforme le vecteur d'entrée $\mathbf{x} \in \mathbb{R}^{14}$ vers un espace latent de dimension $d$ :

$$\mathbf{h}_0 = \mathbf{W}_{\text{proj}}\mathbf{x} + \mathbf{b}_{\text{proj}}.$$

La seconde partie est une séquence de $B$ blocs résiduels. Chaque bloc $k$ applique une transformation $f_k$ composée de cinq opérations : normalisation par couche (*Layer Normalization*) (Ba, Kiros et G. E. Hinton 2016), transformation linéaire, activation ReLU, Dropout, puis seconde transformation linéaire. La sortie de chaque bloc est définie par :

$$\mathbf{h}_k = \mathbf{h}_{k-1} + f_k(\mathbf{h}_{k-1}).$$

Ces connexions résiduelles, introduites par He et al. (2016), permettent au gradient de circuler directement à travers les couches lors de la rétropropagation. Elles facilitent ainsi l'entraînement de réseaux profonds et atténuent le problème de la disparition du gradient. La normalisation par couche (*Layer Normalization*) est utilisée dans les blocs résiduels à la place de la normalisation par lots. Elle normalise les activations indépendamment de la taille du mini-lot, ce qui la rend plus stable pour des données tabulaires.

La troisième partie correspond à une normalisation par couche finale suivie

d'une tête de classification linéaire :

$$\hat{y} = \mathbf{w}_{\text{head}}^{\top} \, \text{LN}(\mathbf{h}_B) + b_{\text{head}}.$$

#### 3.2.4.3 TabNet

Enfin, nous avons utilisé TabNet, proposé par Arik et Pfister (2021). Ce modèle repose sur un mécanisme d'attention séquentielle qui sélectionne les variables utiles à la prédiction. Au lieu de traiter toutes les variables simultanément à chaque couche, TabNet procède en $N_{\text{steps}}$ étapes séquentielles. À chaque étape $i$, un mécanisme d'attention appris génère un masque $\mathbf{M}_i \in [0,1]^d$ qui détermine les variables pertinentes pour la décision à cette étape :

$$\mathbf{M}_i = \text{sparsemax}(\mathbf{P}_i \cdot h_i(\mathbf{a}_{i-1})) \, .$$

où $h_i$ désigne une transformation apprise et $\mathbf{P}_i$ un facteur de pénalisation des variables déjà sélectionnées aux étapes précédentes. Les informations extraites à chaque étape sont agrégées pour produire la prédiction finale. Le coefficient $\gamma$ régule la réutilisation des variables d'une étape à l'autre, tandis que le coefficient $\lambda_{\text{sparse}}$ impose une régularisation de parcimonie sur les masques d'attention, encourageant la sélection d'un petit nombre de variables pertinentes.

TabNet présente aussi un avantage notable en matière d'interprétabilité. Ses masques d'attention fournissent directement, pour chaque observation, une mesure de l'importance accordée à chaque variable, sans nécessiter de méthode d'explication post-hoc. Cette propriété est particulièrement pertinente pour analyser les prédictions dans un contexte médical.

La figure 3.5 montre le fonctionnement séquentiel de TabNet et le rôle des masques d'attention dans la sélection des variables.

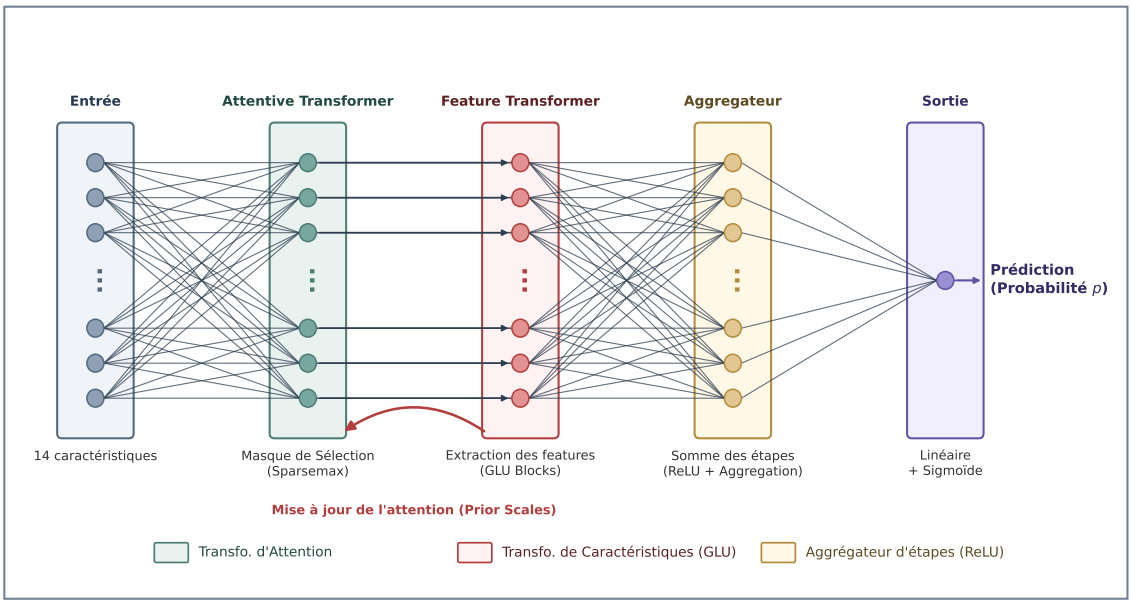



**Fig 3.5 :** Architecture adoptée de TabNet

### 3.2.5 Stratégie d'entraînement

La même stratégie générale d'entraînement a été appliquée aux modèles profonds afin de rendre la comparaison entre architectures aussi équitable que possible.

#### 3.2.5.1 Fonction de perte

Pour le MLP et ResNetTab, nous avons utilisé une entropie croisée binaire, avec possibilité de pondérer la classe positive lorsque la distribution d'entraînement reste déséquilibrée. Pour une observation $(\mathbf{x}, y)$, elle est définie par :

$$\mathcal{L} = -[w_+ \cdot y \cdot \log(\hat{p}) + (1 - y) \cdot \log(1 - \hat{p})].$$

où $\hat{p} = \sigma(\hat{y})$ est la probabilité prédite obtenue par application de la sigmoïde au logit $\hat{y}$, et $w_+$ est le poids de la classe positive, fixé au ratio $N_-/N_+$ entre le nombre d'observations négatives et positives dans l'ensemble d'entraînement effectivement utilisé. Lorsque SMOTE produit un ensemble parfaitement équilibré, ce ratio devient proche de 1 ; la formulation pondérée reste néanmoins utile pour expliciter le traitement d'un éventuel déséquilibre résiduel.

En pratique, le calcul combine la sigmoïde et l'entropie croisée en une seule opération numériquement stable (log-sum-exp), évitant les problèmes de sous-dépassement et de dépassement de capacité.

#### 3.2.5.2 Lissage des étiquettes

Un lissage des étiquettes (*label smoothing*) avec un coefficient $\epsilon = 0{,}05$ a été appliqué lors de l'entraînement des modèles MLP et ResNetTab. Les étiquettes cibles binaires sont transformées selon :

$$\tilde{y} = y \cdot (1 - 2\epsilon) + \epsilon.$$

Les étiquettes $\{0, 1\}$ deviennent ainsi $\{0{,}05\,;\, 0{,}95\}$. Cette technique, proposée initialement par Szegedy et al. (2016), réduit la confiance absolue du modèle dans ses prédictions, améliore la calibration des probabilités, et contribue à la régularisation. Le lissage est appliqué uniquement aux étiquettes d'entraînement ; les étiquettes de validation et de test restent binaires pour le calcul des métriques.

#### 3.2.5.3 Optimisation

Les entraînements utilisent l'optimiseur AdamW (Loshchilov et Hutter 2019), une variante d'Adam qui implémente la décroissance des poids de manière découplée de la mise à jour adaptative du taux d'apprentissage, produisant une régularisation L2 plus efficace que l'implémentation standard d'Adam. Le taux d'apprentissage initial et le coefficient de décroissance des poids sont déterminés par la procédure d'optimisation des hyperparamètres décrite en section 3.2.6.

Un planificateur adaptatif ajuste le taux d'apprentissage en cours d'entraînement. Lorsque l'AUROC sur l'ensemble de validation ne s'améliore pas pendant un nombre défini d'époques consécutives (fixé au tiers de la patience d'arrêt précoce, avec un minimum de 3), le taux d'apprentissage est multiplié par un facteur de 0,5, avec un plancher à $10^{-6}$.

Un écrêtage du gradient (*gradient clipping*) avec une norme L2 maximale de 1,0 a été appliqué à chaque itération pour prévenir l'explosion du gradient.

#### 3.2.5.4 Arrêt précoce (early stopping)

L'entraînement de chaque modèle final est soumis à un arrêt précoce fondé sur l'AUROC de validation. Le nombre maximal d'époques est fixé à 200 et la patience à 25 époques. Après chaque époque, si l'AUROC de validation dépasse le meilleur score observé, les poids du modèle sont sauvegardés et le compteur est réinitialisé. Si aucune amélioration n'est observée pendant 25 époques consécutives, l'entraînement est interrompu et les poids du meilleur modèle sont restaurés. Ce mécanisme empêche le modèle de mémoriser les données d'entraînement au-delà du point de

généralisation optimale.

### 3.2.6 Optimisation des hyperparamètres

La sélection des hyperparamètres a été effectuée avec le framework *Optuna*, en utilisant l'algorithme TPE (*Tree-structured Parzen Estimator*) comme stratégie d'échantillonnage. Cet algorithme modélise la distribution conditionnelle des hyperparamètres en fonction de leur performance passée pour orienter la recherche vers les régions les plus prometteuses de l'espace, ce qui est nettement plus efficace qu'une recherche par grille ou aléatoire.

Pour chaque architecture, 60 essais ont été réalisés. À chaque essai, un jeu d'hyperparamètres est échantillonné, le modèle correspondant est entraîné pendant un maximum de 100 époques avec une patience de 20, et l'AUROC de validation sert de métrique d'objectif à maximiser. Les espaces de recherche explorés sont détaillés ci-dessous.

Pour le MLP : dimension des couches cachées $\in \{32, 64, 128, 256\}$, nombre de couches $\in [1, 3]$, taux de Dropout $\in [0{,}1\,;\,0{,}5]$, taux d'apprentissage $\in [5\times10^{-5}\,;\,5\times10^{-3}]$ (échelle log), décroissance des poids $\in [10^{-6}\,;\,10^{-3}]$ (échelle log), taille du mini-lot $\in \{64, 128, 256\}$, et activation du mode entonnoir (oui/non).

Pour ResNetTab : dimension de l'espace latent $\in \{32, 64, 128, 256\}$, nombre de blocs résiduels $\in [1, 4]$, taux de Dropout $\in [0{,}05\,;\,0{,}5]$, taux d'apprentissage $\in [5\times10^{-5}\,;\,5\times10^{-3}]$ (échelle log), décroissance des poids $\in [10^{-6}\,;\,10^{-3}]$ (échelle log), et taille du mini-lot $\in \{64, 128, 256\}$. Ces choix contrôlent directement la largeur de la représentation latente, la profondeur résiduelle et le niveau de régularisation.

Pour TabNet : dimension de décision et d'attention $n_d = n_a \in \{8, 16, 32, 64\}$, nombre d'étapes séquentielles $\in [2, 8]$, taux d'apprentissage $\in [5\times10^{-4}\,;\,10^{-2}]$ (échelle log), coefficient $\gamma \in [1{,}0\,;\,2{,}0]$, régularisation de parcimonie $\lambda_{\text{sparse}} \in [10^{-6}\,;\,10^{-3}]$ (échelle log), momentum $\in [0{,}01\,;\,0{,}4]$, et taille du mini-lot $\in \{64, 128, 256, 512\}$. Cet espace de recherche couvre les principaux paramètres de l'architecture TabNet utilisée.

Après identification des meilleurs hyperparamètres, les modèles finaux sont réentraînés avec ces configurations sur l'ensemble d'entraînement complet, avec un budget de 200 époques et une patience de 25 époques.

### 3.2.7 Seuils de décision

Les modèles entraînés produisent en sortie une probabilité continue $\hat{p} \in [0, 1]$, convertie en décision binaire par comparaison à un seuil $\tau$. Le seuil par défaut de

0,5 n'est pas nécessairement optimal, en particulier en présence de déséquilibre de classes.

Pour les trois modèles d'apprentissage profond, le seuil optimal a été déterminé sur l'ensemble de validation par maximisation de l'indice $J$ de Youden :

$$J(\tau) = \text{Sensibilité}(\tau) + \text{Spécificité}(\tau) - 1 = \text{TPR}(\tau) - \text{FPR}(\tau).$$

Le seuil retenu est celui qui maximise cet indice, offrant le meilleur compromis entre la détection des cas positifs et l'identification correcte des cas négatifs. Géométriquement, ce seuil correspond au point de la courbe ROC le plus éloigné de la diagonale du hasard. Pour les deux modèles de référence, le seuil standard de 0,5 a été conservé afin de les maintenir dans une configuration de référence classique. Cette différence de traitement des seuils est prise en compte dans l'interprétation comparative des résultats.

### 3.2.8 Métriques d'évaluation

L'évaluation repose sur un ensemble de métriques complémentaires. Chacune capture un aspect différent de la performance, et leur combinaison permet une caractérisation complète du comportement des modèles. Toutes les métriques sont calculées sur l'ensemble de test, qui n'a jamais été exposé aux modèles lors de l'entraînement ou de la sélection des hyperparamètres.

Soit TP, TN, FP et FN les nombres de vrais positifs, vrais négatifs, faux positifs et faux négatifs respectivement.

**Sensibilité (rappel).** La sensibilité mesure la proportion des cas positifs réels correctement détectés :

$$\text{Sensibilité} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

En contexte clinique, une sensibilité élevée minimise le nombre de cas manqués (faux négatifs), ce qui est essentiel pour une pathologie dont les conséquences d'un diagnostic manqué sont graves.

**Spécificité.** La spécificité mesure la proportion des cas négatifs réels correctement identifiés :

$$\text{Spécificité} = \frac{\text{TN}}{\text{TN} + \text{FP}}.$$

Une spécificité élevée limite les fausses alertes et les examens complémentaires inutiles.

**Précision.** La précision (ou valeur prédictive positive) mesure la proportion de prédictions positives qui sont effectivement correctes :

$$\text{Précision} = \frac{\text{TP}}{\text{TP} + \text{FP}}.$$

**Score F1 (macro).** Le score F1 est la moyenne harmonique de la précision et du rappel. La variante macro calcule le F1 pour chaque classe séparément, puis en fait la moyenne arithmétique, accordant un poids égal aux deux classes indépendamment de leur fréquence :

$$F_1 = \frac{2 \cdot \text{Précision} \cdot \text{Rappel}}{\text{Précision} + \text{Rappel}} \qquad F_1^{\text{macro}} = \frac{1}{2}\left(F_1^{(0)} + F_1^{(1)}\right).$$

**AUROC.** L'aire sous la courbe ROC (*Area Under the Receiver Operating Characteristic curve*) mesure la capacité globale de discrimination du modèle. La courbe ROC trace le taux de vrais positifs (TPR) en fonction du taux de faux positifs (FPR) pour tous les seuils possibles :

$$\text{AUROC} = \int_0^1 \text{TPR}\Big(\text{FPR}^{-1}(t)\Big)\, dt.$$

Cette métrique s'interprète comme la probabilité qu'un échantillon positif choisi aléatoirement reçoive un score plus élevé qu'un échantillon négatif choisi aléatoirement. Une valeur de 0,5 correspond à un classifieur aléatoire et 1,0 à une discrimination parfaite. L'AUROC est indépendante du seuil de décision.

**AUPRC.** L'aire sous la courbe précision-rappel (*Area Under the Precision-Recall Curve*) est définie par :

$$\text{AUPRC} = \int_0^1 \text{Précision}\Big(\text{Rappel}^{-1}(r)\Big)\, dr.$$

Contrairement à l'AUROC, qui peut surestimer la performance lorsque la classe négative est très majoritaire (car elle prend en compte les vrais négatifs via le FPR), l'AUPRC se concentre exclusivement sur la capacité du modèle à identifier les cas positifs. Elle constitue une métrique plus exigeante et plus informative en présence de déséquilibre de classes.

### 3.2.9 Méthodes d'interprétabilité

#### 3.2.9.1 Valeurs SHAP (MLP et ResNetTab)

Pour les modèles MLP et ResNetTab, l'interprétabilité a été obtenue par les valeurs SHAP (*SHapley Additive exPlanations*), proposées par Lundberg et Lee (2017). Fondées sur la théorie des jeux coopératifs de Shapley, les valeurs SHAP attribuent à chaque variable une contribution marginale à la prédiction. La valeur SHAP $\phi_j$ d'une variable $j$ pour une observation $\mathbf{x}$ est définie comme la moyenne pondérée des contributions marginales de cette variable sur tous les sous-ensembles possibles des autres variables :

$$\phi_j = \sum_{S \subseteq \{1,\dots,d\} \setminus \{j\}} \frac{|S|!\,(d - |S| - 1)!}{d!} \left[ f(S \cup \{j\}) - f(S) \right].$$

L'estimation a été réalisée par un *KernelExplainer*, méthode d'approximation agnostique au modèle, avec un échantillon de référence de 50 observations représentatives obtenues par k-means sur l'ensemble d'entraînement. Les valeurs ont été calculées sur 150 observations de test avec 80 échantillons de permutation. La moyenne des valeurs absolues SHAP fournit un classement global de l'importance des variables.

#### 3.2.9.2 Masques d'attention (TabNet)

Pour TabNet, l'interprétabilité est intrinsèque à l'architecture. Les masques d'attention générés à chaque étape séquentielle attribuent un poids normalisé à chaque variable d'entrée. En moyennant ces poids sur l'ensemble des observations de test, on obtient une mesure globale de l'importance de chaque variable telle que perçue par le modèle, sans recours à une méthode d'explication externe.

### 3.2.10 Configuration expérimentale

#### 3.2.10.1 Environnement

L'ensemble des expériences a été conduit sur Google Colaboratory, équipé d'un GPU NVIDIA Tesla T4 (16 Go de mémoire, architecture Turing). Le framework d'apprentissage profond utilisé est PyTorch. Les modèles de référence reposent sur scikit-learn (régression logistique) et la bibliothèque XGBoost. L'architecture TabNet est implémentée par pytorch-tabnet. L'optimisation des hyperparamètres utilise Optuna, le suréchantillonnage SMOTE utilise la bibliothèque imbalanced-learn, et l'analyse d'interprétabilité utilise la bibliothèque SHAP.

Le prototype de démonstration présenté plus loin a été développé avec Streamlit et déployé sur Streamlit Community Cloud, afin de permettre l'exploration des prédictions, la visualisation des métriques et l'analyse de l'interprétabilité pour des observations individuelles.

#### 3.2.10.2 Reproductibilité

La reproductibilité est assurée par la fixation d'une graine pseudo-aléatoire unique (valeur 42) pour l'ensemble des sources de stochasticité : générateurs aléatoires de Python, NumPy et PyTorch, échantillonneur TPE d'Optuna, et algorithme SMOTE. La partition des données utilise cette même graine pour la stratification. L'exécution répétée du code produit ainsi des résultats identiques sous un environnement logiciel constant.

## 3.3 Résultats et analyse

La méthodologie précédente a décrit le cadre expérimental adopté : les architectures de modèles, le protocole de prétraitement, la stratégie de gestion du déséquilibre de classes, ainsi que la procédure d'optimisation des hyperparamètres. La présente section rapporte les résultats obtenus par les cinq modèles sur l'ensemble de test et en propose une analyse comparative.

### 3.3.1 Comparaison globale des performances

Le tableau 3.3 présente les résultats obtenus par les cinq modèles sur l'ensemble de test (848 observations : 719 négatives et 129 positives). Les seuils de décision des modèles d'apprentissage profond ont été optimisés sur l'ensemble de validation par maximisation de l'indice $J$ de Youden, tandis que les deux modèles de référence utilisent le seuil standard de 0,5.

**Tab 3.3 :** Performance des modèles sur l'ensemble de test. Les deux premières lignes correspondent aux modèles de référence (baselines), placés en haut du tableau. Les meilleurs scores de chaque métrique sont indiqués en gras.

| Modèle | AUROC | AUPRC | Sensibilité | Spécificité | F1 macro | Seuil |
|---|---|---|---|---|---|---|
| Régression logistique | **0,692** | **0,281** | 0,581 | 0,683 | 0,562 | 0,500 |
| XGBoost | 0,627 | 0,231 | 0,225 | **0,868** | 0,547 | 0,500 |
| MLP | **0,692** | 0,260 | 0,457 | 0,784 | **0,589** | 0,568 |
| TabNet | 0,675 | 0,254 | 0,419 | 0,789 | 0,578 | 0,566 |
| ResNetTab | 0,683 | 0,268 | **0,643** | 0,668 | 0,570 | 0,510 |

La figure 3.6 offre une vue synthétique de ces résultats sous forme de diagramme en barres.

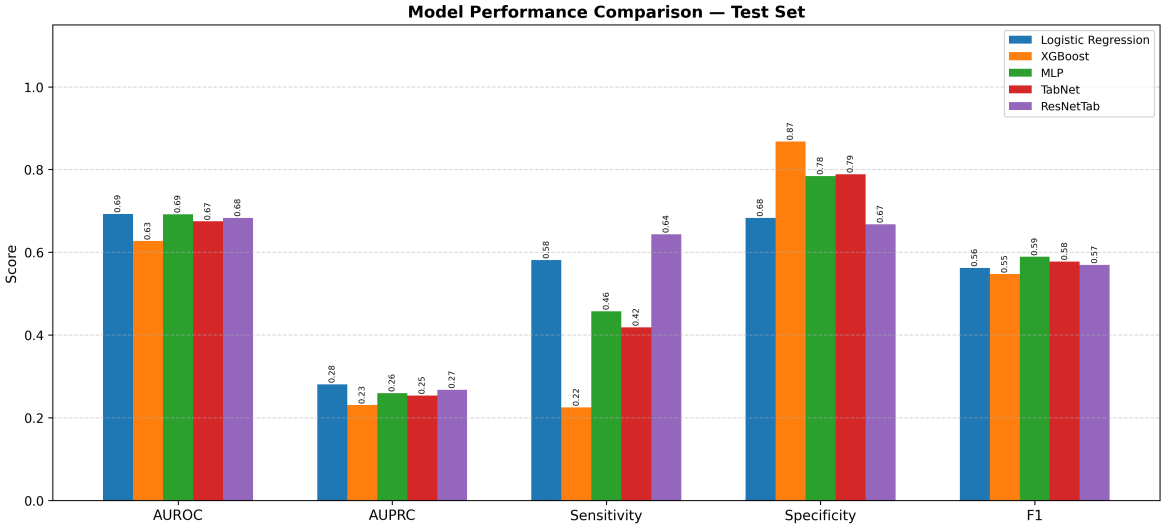



**Fig 3.6 :** Comparaison globale des performances des modèles sur l'ensemble de test

En termes d'AUROC, la régression logistique et le MLP obtiennent les scores les plus élevés (0,692), suivis de près par ResNetTab (0,683) et TabNet (0,675). XGBoost affiche le score le plus bas (0,627). Le MLP égale donc la régression logistique, et les écarts de ResNetTab et TabNet par rapport à celle-ci restent faibles (respectivement 0,009 et 0,017 points).

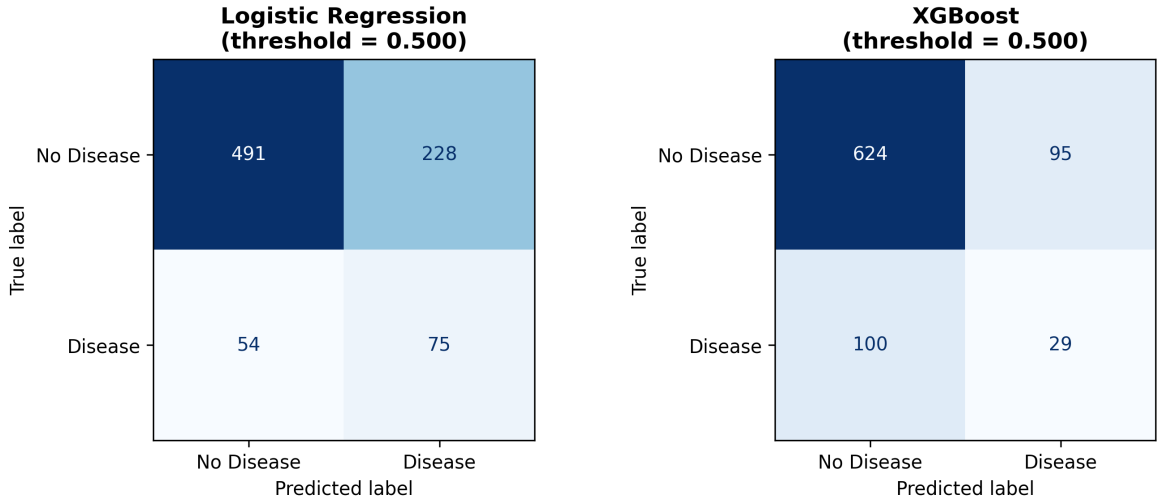
L'AUPRC, métrique plus exigeante en présence de déséquilibre de classes, confirme la tendance générale. La régression logistique obtient la valeur la plus élevée (0,281), suivie de ResNetTab (0,268) et du MLP (0,260). L'ensemble des valeurs d'AUPRC reste modeste, ce qui reflète la difficulté de la tâche de détection dans un jeu de données où la classe positive ne représente que 15,2 % des observations.

Le score F1 macro, qui accorde un poids égal aux deux classes, place le MLP

en tête (0,589), devant TabNet (0,578), ResNetTab (0,570), la régression logistique (0,562) et XGBoost (0,547). Ce classement diffère de celui de l'AUROC, car le F1 macro dépend du seuil de décision retenu et sanctionne les déséquilibres entre sensibilité et spécificité.

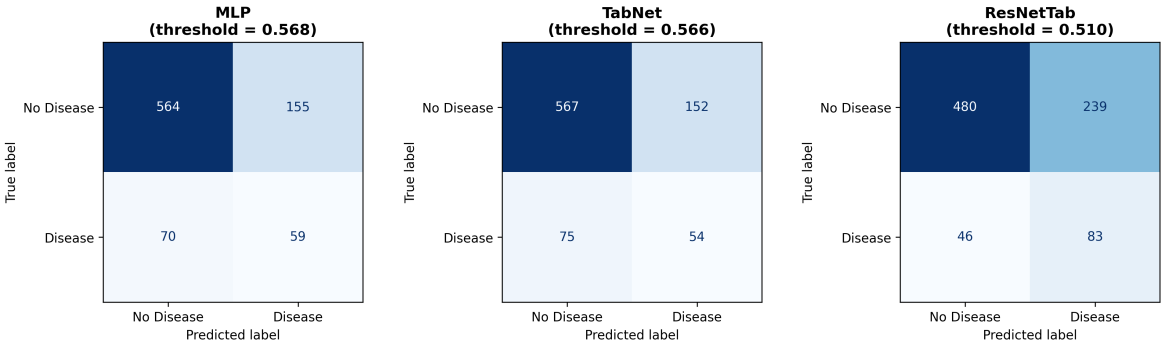
### 3.3.2 Analyse des matrices de confusion

Les matrices de confusion permettent d'examiner la répartition des prédictions de chaque modèle en termes de vrais positifs (TP), faux positifs (FP), vrais négatifs (TN) et faux négatifs (FN). La figure 3.7 regroupe les deux modèles de référence, avec la régression logistique en sous-figure 3.7a et XGBoost en sous-figure 3.7b. La figure 3.8 regroupe les architectures d'apprentissage profond, avec le MLP en sous-figure 3.8a, TabNet en sous-figure 3.8b et ResNetTab en sous-figure 3.8c.




(**a**) Régression logistique (**b**) XGBoost

**Fig 3.7 :** Matrices de confusion des modèles de référence sur l'ensemble de test




(**a**) MLP (**b**) TabNet (**c**) ResNetTab

**Fig 3.8 :** Matrices de confusion des architectures d'apprentissage profond sur l'ensemble de test

La comparaison des matrices révèle des profils de décision distincts.

ResNetTab affiche la sensibilité la plus élevée (0,643), détectant 83 des 129 cas positifs, mais au prix de 239 faux positifs sur 719 cas négatifs. Son seuil optimal, très proche de 0,5 (0,510), traduit une stratégie de classification qui favorise la détection des cas positifs.

La régression logistique obtient la deuxième sensibilité (0,581), avec 75 vrais positifs et 228 faux positifs. Ce résultat est notable car il est obtenu avec le seuil par défaut de 0,5 et sans optimisation des hyperparamètres par recherche bayésienne.

Le MLP et TabNet adoptent des profils intermédiaires avec des sensibilités de 0,457 et 0,419 respectivement, et des spécificités proches (0,784 et 0,789). Leurs seuils optimaux, relevés autour de 0,57, indiquent que l'optimisation de Youden a conduit à des seuils plus conservateurs qui privilégient la spécificité.

XGBoost présente le profil le plus conservateur : une spécificité élevée (0,868) mais une sensibilité de seulement 0,225, ne détectant que 29 des 129 cas positifs. Les 100 faux négatifs représentent 77,5 % de l'ensemble des cas positifs.

En contexte clinique, où un diagnostic manqué peut avoir des conséquences graves, la sensibilité de ResNetTab (0,643) présente un avantage, mais elle s'accompagne d'un taux de faux positifs de 33,2 % (239/719) qui entraînerait des examens complémentaires inutiles. Le choix du modèle dépend donc du compromis accepté entre faux négatifs et faux positifs, qui relève d'une décision clinique.

### 3.3.3 Courbes ROC et précision-rappel

La figure 3.9 présente les courbes ROC (panneau gauche) et les courbes précision-rappel (panneau droit) des cinq modèles.

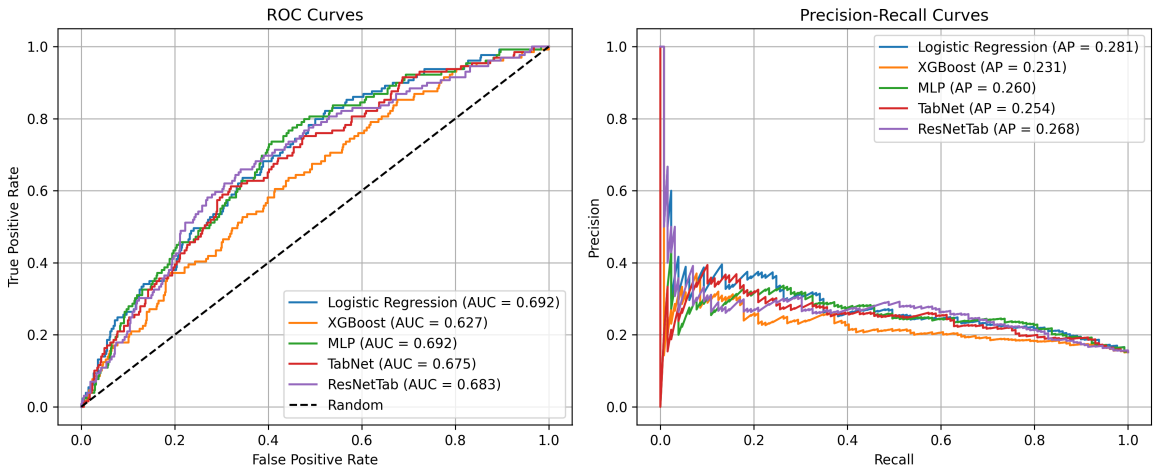



**Fig 3.9 :** Courbes ROC et précision-rappel des cinq modèles sur l'ensemble de test

Les courbes ROC montrent que les cinq modèles se situent nettement au-dessus de la diagonale du hasard, confirmant une capacité discriminante réelle. Les courbes de la régression logistique, du MLP et de ResNetTab sont proches et se chevauchent dans une grande partie de l'espace ROC, ce qui est cohérent avec les valeurs d'AUROC similaires rapportées dans le tableau 3.3. La courbe de XGBoost se distingue par un profil plus bas, particulièrement dans la région des faibles taux de faux positifs.

Les courbes précision-rappel offrent une perspective complémentaire. Toutes les courbes chutent rapidement à mesure que le rappel augmente, ce qui est caractéristique d'un problème fortement déséquilibré. La régression logistique maintient une précision légèrement supérieure dans la zone de rappel intermédiaire (0,3 à 0,6), ce qui se traduit par son AUPRC plus élevé (0,281). Les valeurs d'AUPRC confirment que la tâche reste difficile pour l'ensemble des modèles : même le meilleur modèle n'atteint que 0,281, soit environ le double de la prévalence de la classe positive (0,152), ce qui indique une capacité de discrimination limitée mais non négligeable.

### 3.3.4 Dynamique d'entraînement

#### 3.3.4.1 Courbes de perte

La figure 3.10 présente l'évolution des pertes d'entraînement et de validation au fil des époques pour le MLP et ResNetTab.

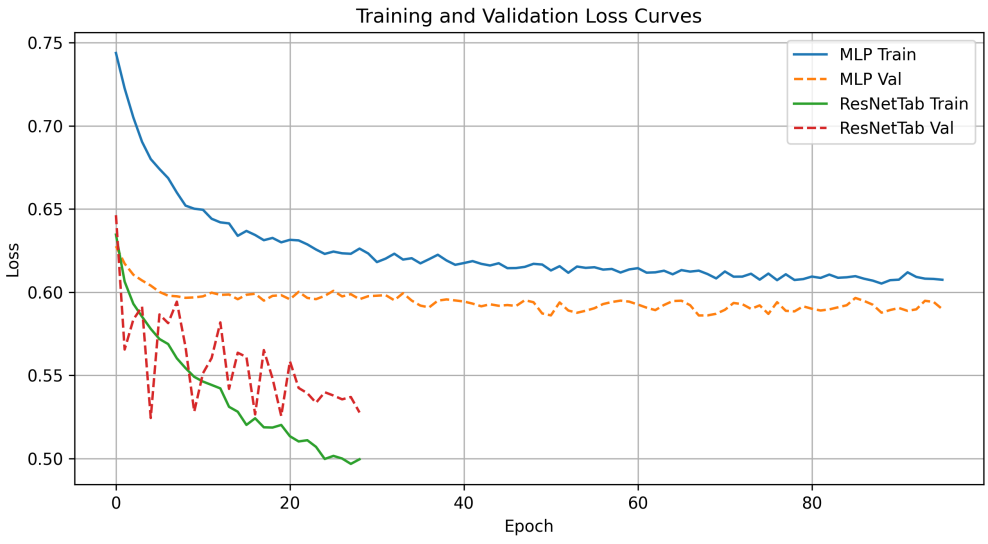



**Fig 3.10 :** Courbes de perte d'entraînement et de validation pour le MLP et ResNetTab

Le MLP démarre avec une perte d'entraînement élevée (environ 0,75) qui dé-

croît progressivement sur une centaine d'époques pour se stabiliser autour de 0,61. Sa perte de validation se stabilise rapidement autour de 0,60, et l'écart entre les deux courbes reste faible tout au long de l'entraînement. Ce profil indique un apprentissage régulier avec un surapprentissage limité.

ResNetTab présente une dynamique différente. Sa perte d'entraînement décroît plus rapidement et atteint des valeurs plus basses (environ 0,50), tandis que sa perte de validation reste plus élevée et plus variable. L'écart entre les deux courbes est plus marqué que pour le MLP, ce qui indique un degré de surapprentissage plus élevé malgré les mécanismes de régularisation employés (Dropout, décroissance des poids, arrêt précoce). L'arrêt précoce a interrompu l'entraînement de ResNetTab autour de l'époque 30, bien avant le budget maximal de 200 époques.

#### 3.3.4.2 Convergence de la recherche d'hyperparamètres

La figure 3.11 présente l'évolution du meilleur AUROC de validation au fil des 60 essais Optuna pour chaque modèle d'apprentissage profond.

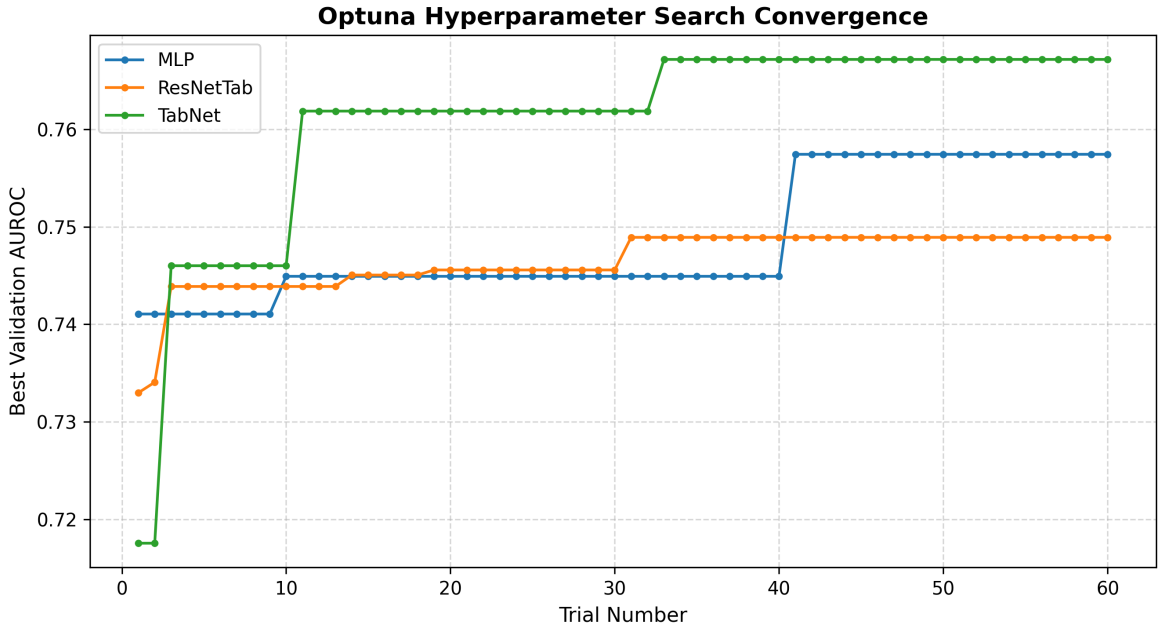



**Fig 3.11 :** Convergence de la recherche d'hyperparamètres Optuna pour les modèles d'apprentissage profond

Les trois modèles atteignent de bonnes configurations dès les premiers essais, avec des AUROC de validation supérieurs à 0,74 après 5 essais seulement. TabNet converge le plus rapidement, atteignant un plateau autour de 0,761 dès l'essai 10, avec une amélioration marginale vers 0,767 autour de l'essai 32. Le MLP atteint son meilleur score (environ 0,758) à l'essai 40. ResNetTab converge plus progressivement pour atteindre environ 0,749.

Les AUROC de validation obtenus lors de la recherche d'hyperparamètres (0,749 à 0,767) sont sensiblement supérieurs aux AUROC de test rapportés dans le tableau 3.3 (0,675 à 0,692). Cet écart est attendu, car l'optimisation des hyperparamètres sélectionne les configurations qui performent le mieux sur la validation, introduisant un biais d'optimisme. L'écart observé (environ 0,07 points) reste dans des marges habituelles et ne remet pas en cause la validité de la procédure.

### 3.3.5 Calibration des probabilités

La figure 3.12 présente les courbes de calibration (diagramme de fiabilité) des cinq modèles sur l'ensemble de test.

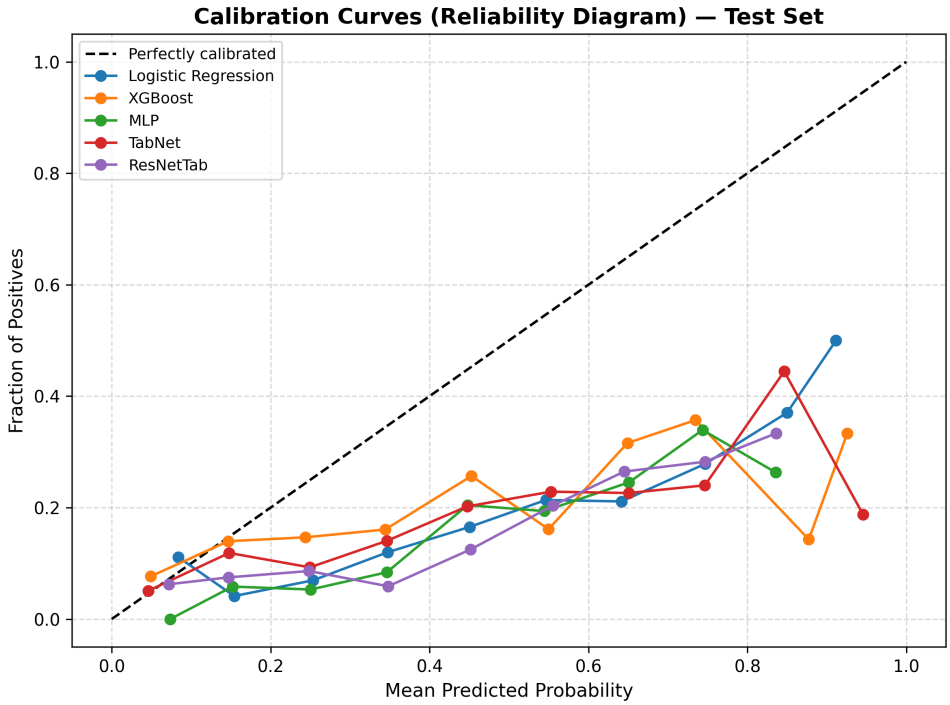



**Fig 3.12 :** Courbes de calibration des probabilités sur l'ensemble de test

L'ensemble des modèles présente une surestimation systématique des probabilités dans les intervalles supérieurs : lorsqu'un modèle attribue une probabilité de 0,7 à 0,9, la fréquence réelle de cas positifs est nettement inférieure (typiquement 0,25 à 0,45). Ce phénomène est en partie attendu compte tenu du fort déséquilibre du jeu de données. La prévalence de 15,2 % impose un plafond naturel aux fréquences observées dans chaque intervalle de probabilité, et les mécanismes de compensation du déséquilibre, notamment SMOTE et les pondérations de classe lorsqu'elles sont actives, peuvent produire des scores décalés vers le haut par rapport à la prévalence réelle.

La régression logistique affiche le profil le plus proche de la diagonale dans les intervalles supérieurs, ce qui est cohérent avec sa formulation probabiliste native. Les modèles d'apprentissage profond présentent des probabilités plus expressives mais moins directement calibrées, ce qui ouvre une piste d'amélioration naturelle par calibration post-hoc (par exemple, Platt scaling ou calibration isotonique) avant un éventuel déploiement clinique.

### 3.3.6 Interprétabilité et importance des variables

La figure 3.13 présente l'importance des variables pour les trois modèles d'apprentissage profond : les valeurs SHAP moyennes pour le MLP et ResNetTab, et les poids d'attention moyens pour TabNet.

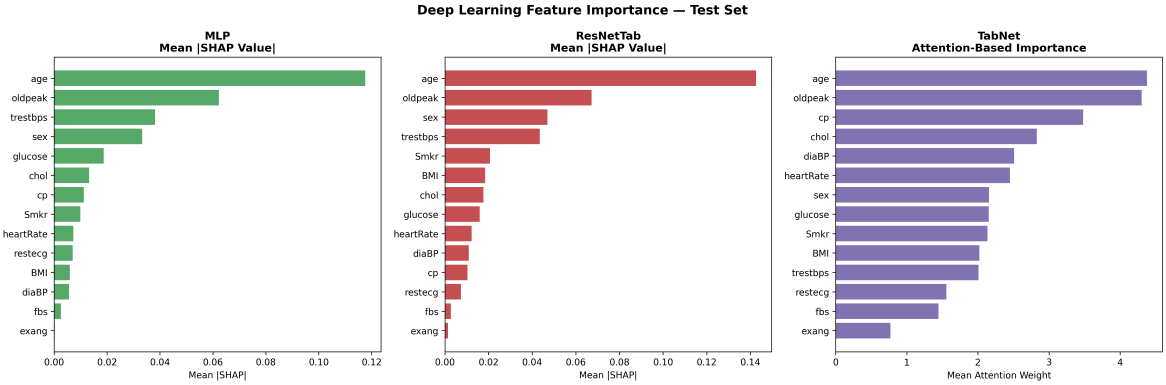



**Fig 3.13 :** Importance des variables pour les modèles d'apprentissage profond sur l'ensemble de test

Les trois modèles s'accordent sur les variables les plus influentes. L'âge domine nettement le classement dans les trois cas : c'est la variable qui contribue le plus aux prédictions du MLP (SHAP moyen $\approx 0{,}12$), de ResNetTab (SHAP moyen $\approx 0{,}14$) et de TabNet (poids d'attention le plus élevé). Ce résultat est cohérent avec les connaissances médicales, l'âge étant le facteur de risque cardiovasculaire le plus établi.

La dépression du segment ST (`oldpeak`) se place systématiquement au deuxième rang. Cette variable, qui mesure une anomalie électrocardiographique induite par l'effort, est un marqueur clinique reconnu d'ischémie myocardique.

La pression artérielle systolique (`trestbps`) et le sexe complètent le groupe des quatre variables les plus importantes pour le MLP et ResNetTab. TabNet accorde en revanche un poids plus élevé au type de douleur thoracique (`cp`) et au cholestérol, et répartit l'attention de manière plus uniforme entre les variables restantes. Cette différence peut s'expliquer par le mécanisme de sélection séquentielle de TabNet,

qui procède par étapes et peut capturer des interactions entre variables différentes de celles exploitées par les architectures entièrement connectées.

Les variables situées en bas du classement sont également cohérentes entre les modèles. La glycémie à jeun (`fbs`) et l'angine de poitrine à l'effort (`exang`) contribuent peu aux prédictions. Pour `exang`, cette faible importance peut paraître surprenante d'un point de vue clinique, mais elle peut s'expliquer par une redondance partielle avec `oldpeak`, les deux variables étant liées à l'épreuve d'effort.

Dans l'ensemble, les classements obtenus sont globalement cohérents entre les trois modèles et conformes à la littérature médicale sur les facteurs de risque cardiovasculaire, ce qui conforte la validité des prédictions.

## 3.4 Discussion

Plusieurs constats se dégagent de cette étude comparative.

Premièrement, les modèles d'apprentissage profond se montrent compétitifs dans un cadre tabulaire contraint. Le MLP atteint le même meilleur AUROC que la régression logistique (0,692), tandis que ResNetTab et TabNet restent proches malgré la taille modérée du jeu de données. Ce résultat est cohérent avec les observations de Grinsztajn, Oyallon et Varoquaux (2022) et Shwartz-Ziv et Armon (2022), selon lesquelles les écarts entre approches profondes et classiques dépendent fortement du volume, de la richesse et de la structure des variables tabulaires.

Deuxièmement, les modèles d'apprentissage profond apportent des avantages opérationnels distincts. Le MLP obtient le meilleur score F1 macro (0,589), ce qui indique le meilleur équilibre global entre les deux classes. ResNetTab atteint la sensibilité la plus élevée (0,643), ce qui en fait le modèle le plus adapté à un scénario de dépistage où minimiser les cas manqués est prioritaire. TabNet apporte une interprétabilité intrinsèque grâce à ses masques d'attention, sans nécessiter de méthode post-hoc.

Troisièmement, la performance relativement faible de XGBoost est inattendue, cet algorithme étant généralement compétitif sur les données tabulaires. Sa très basse sensibilité (0,225) avec le seuil par défaut de 0,5 suggère que ses probabilités sont mal calibrées pour ce seuil et qu'une optimisation du seuil de décision, comme celle appliquée aux modèles d'apprentissage profond, aurait pu améliorer sensiblement son profil sensibilité-spécificité. La comparaison entre les modèles de référence et les modèles d'apprentissage profond est donc partiellement asymétrique sur ce point.

Quatrièmement, les valeurs d'AUPRC, toutes inférieures à 0,30, doivent être interprétées à la lumière de la faible prévalence de la classe positive. Elles montrent

que la tâche reste exigeante avec les variables disponibles, mais aussi que les modèles captent un signal utile au-delà du hasard. L'intégration de données plus riches, telles que l'imagerie, les biomarqueurs ou des mesures longitudinales, pourrait particulièrement renforcer les architectures profondes, conçues pour exploiter des représentations complexes.

Enfin, la concordance de l'analyse d'interprétabilité entre les trois architectures d'apprentissage profond et avec les connaissances médicales établies constitue un résultat important. Elle confirme que ces modèles exploitent des relations cliniquement pertinentes et renforce leur intérêt comme outils d'analyse, au-delà de la seule comparaison de scores globaux.

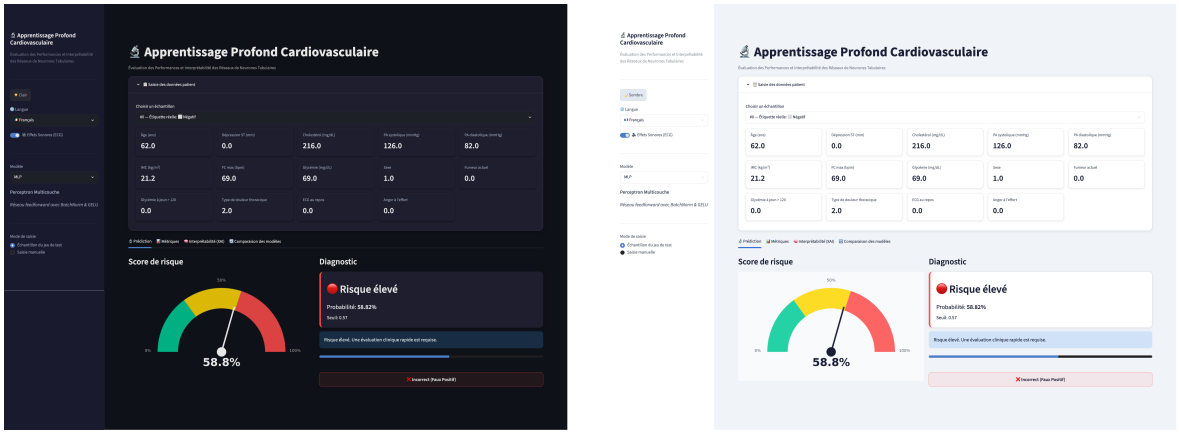
Sur le plan applicatif, ces résultats justifient la présentation, dans la section suivante, d'un prototype démonstratif destiné à illustrer l'utilisation possible des modèles et des explications associées. Ce prototype ne constitue toutefois ni un dispositif médical validé ni un outil d'aide à la décision clinique.

## 3.5 Prototype d'application interactive

En complément de l'étude expérimentale, une application web interactive a été développée avec *Streamlit* (Streamlit 2026) afin de présenter les modèles d'apprentissage profond dans un cadre d'utilisation simple. Elle permet de saisir les caractéristiques cliniques d'un patient, de choisir l'un des trois modèles profonds étudiés (MLP, ResNetTab ou TabNet), puis d'obtenir une prédiction accompagnée d'éléments d'explicabilité. L'application constitue un *prototype de démonstration et d'analyse*, **non** un outil clinique validé. Une version de démonstration est disponible en ligne[1].

Les figures suivantes présentent les principales captures de l'application, depuis l'écran initial avec le modèle MLP sélectionné par défaut jusqu'aux métriques, aux prédictions, à la comparaison globale des modèles et aux explications (XAI).

[1] https://lounissi-boudis-memoire.streamlit.app/




(**a**) Mode sombre (**b**) Mode clair

**Fig 3.14 :** Vue générale initiale de l'application avec le modèle MLP sélectionné par défaut, illustrant la prédiction effectuée pour un échantillon de patient. L'application permet également de changer de langue (arabe, français, anglais) et de basculer entre les thèmes clair et sombre.

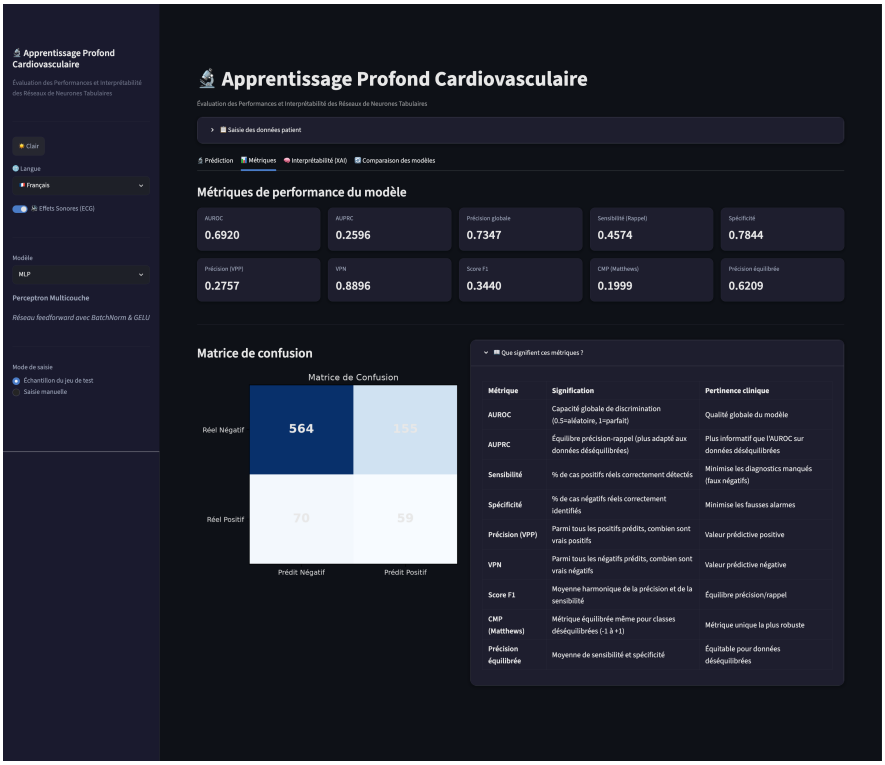



**Fig 3.15 :** Onglet des métriques du modèle MLP, incluant les indicateurs de performance, la matrice de confusion et le rapport de classification

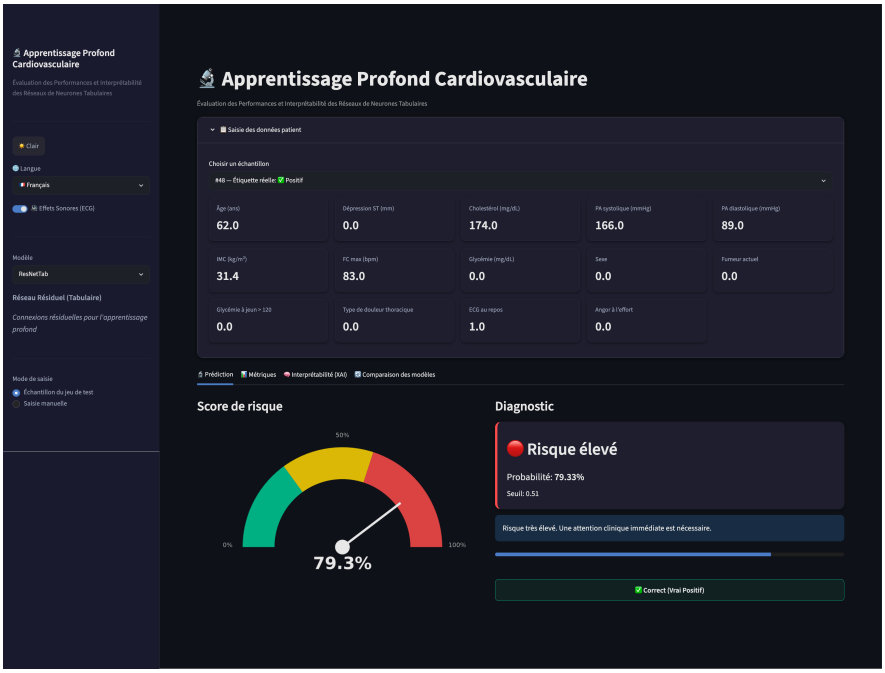



**Fig 3.16 :** Prédiction d'un échantillon patient avec le modèle ResNetTab, affichant la probabilité estimée et le niveau de risque associé

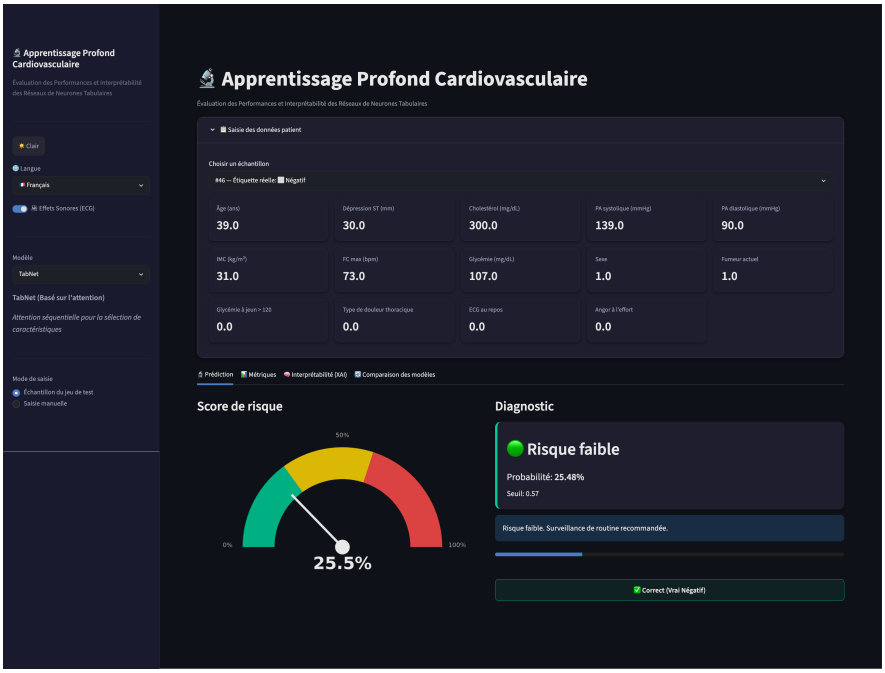



**Fig 3.17 :** Prédiction d'un échantillon patient avec le modèle TabNet, avec estimation du risque et synthèse de décision

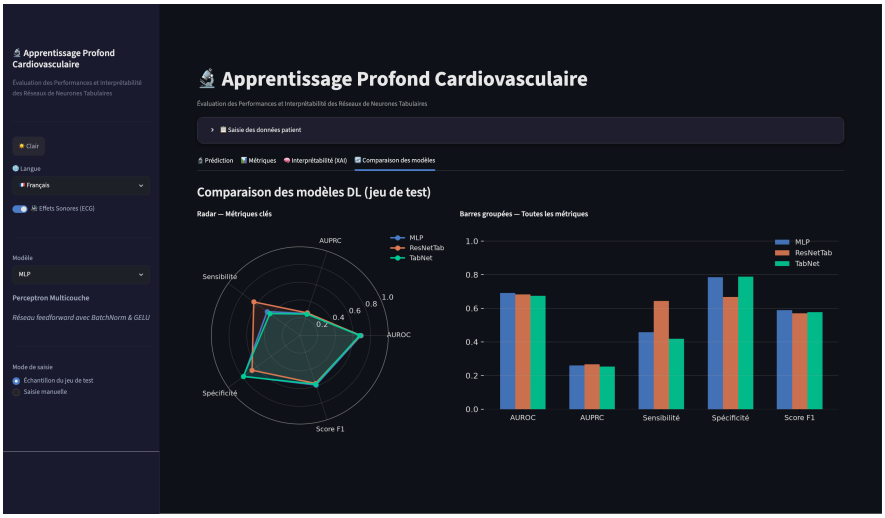



**Fig 3.18 :** Comparaison globale des modèles profonds à l'aide de graphiques radar et de diagrammes en barres des métriques de performance

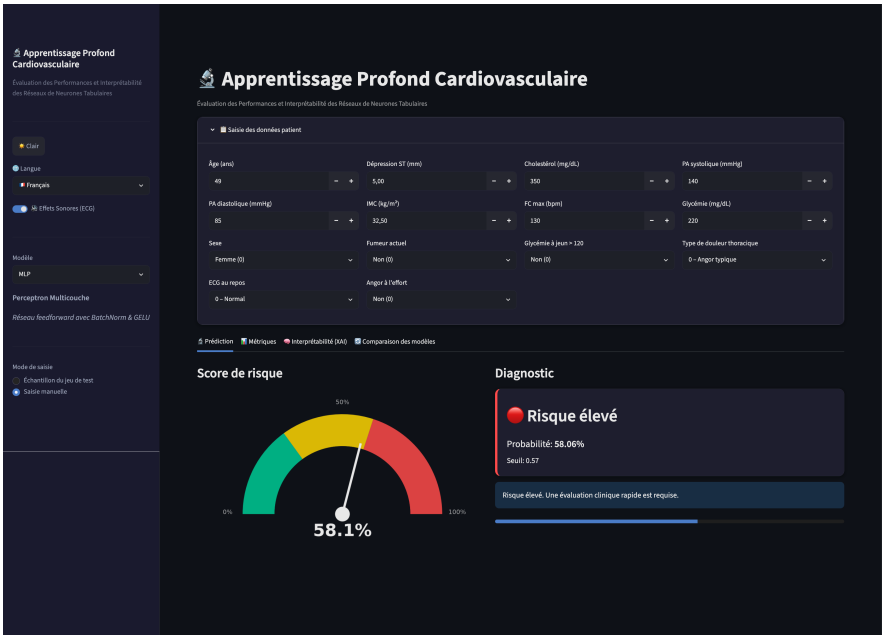



**Fig 3.19 :** Prédiction issue d'une saisie manuelle des caractéristiques cliniques du patient

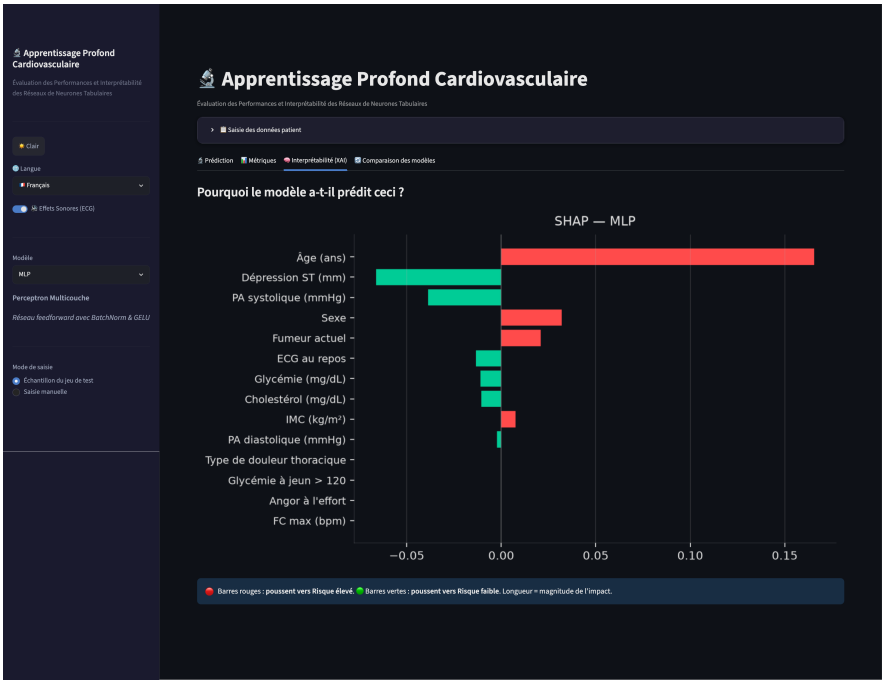



**Fig 3.20 :** Interprétabilité du modèle MLP par importance locale des variables (avec SHAP)

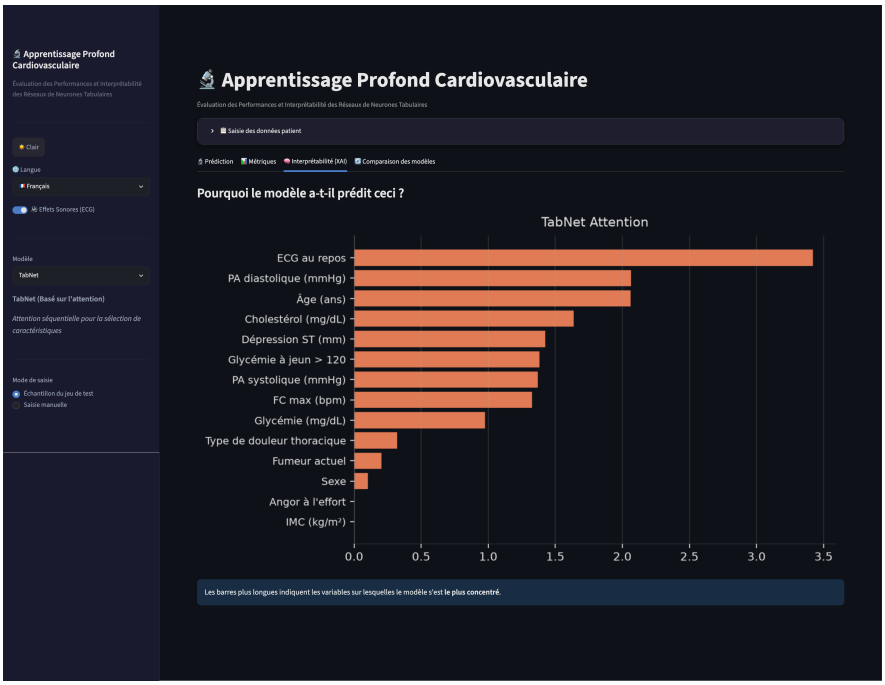



**Fig 3.21 :** Interprétabilité du modèle TabNet à partir des importances de variables issues du mécanisme d'attention

## 3.6 Conclusion

Ce chapitre a présenté l'ensemble du protocole expérimental mis en place pour la prédiction des maladies cardiovasculaires à partir de données cliniques tabulaires. Les données ont été préparées par écrêtage des valeurs extrêmes, partition stratifiée, normalisation et gestion du déséquilibre par SMOTE. Les modèles de référence ont ensuite été comparés à trois architectures profondes selon des métriques complémentaires, afin d'évaluer non seulement la discrimination globale, mais aussi le compromis entre sensibilité, spécificité et équilibre des classes.

Les résultats montrent que les modèles profonds sont compétitifs, sans dominer systématiquement les approches classiques. Le MLP atteint le meilleur score F1 macro et égale la régression logistique en AUROC, tandis que ResNetTab obtient la sensibilité la plus élevée, ce qui le rend particulièrement intéressant dans une logique de dépistage. TabNet présente quant à lui un intérêt spécifique grâce à son mécanisme d'attention, qui facilite l'interprétation des variables utilisées par le modèle. L'analyse d'importance des variables met en évidence des facteurs cohérents avec les connaissances médicales, notamment l'âge, la dépression du segment ST, la pression artérielle systolique et le sexe.

L'étude met également en évidence plusieurs limites importantes. Les valeurs d'AUPRC restent modestes en raison de la faible proportion de cas positifs, et les courbes de calibration montrent une tendance à surestimer les probabilités de risque. Ces résultats indiquent que les modèles doivent être améliorés et validés sur des données plus riches et plus larges avant toute utilisation clinique. Le prototype Streamlit développé dans ce travail constitue donc une démonstration pédagogique et analytique, et non un outil médical validé. Les résultats obtenus fournissent néanmoins une base solide pour poursuivre l'amélioration du système, notamment par la calibration post-hoc, l'intégration de nouvelles variables et une validation externe.

# Conclusion générale et perspectives

## Synthèse générale

Ce mémoire a étudié la prédiction des maladies cardiovasculaires à partir de données cliniques tabulaires. L'objectif principal était de comparer des modèles classiques et des modèles d'apprentissage profond dans un cadre expérimental contrôlé, en tenant compte des contraintes propres au domaine médical : déséquilibre des classes, importance de la sensibilité, besoin d'interprétabilité et prudence dans l'utilisation des probabilités prédites.

Le travail a d'abord présenté les notions médicales nécessaires à la compréhension des maladies cardiovasculaires, de leurs facteurs de risque et de leur prévention. Il a ensuite introduit les méthodes d'apprentissage supervisé, les modèles de classification et les architectures profondes utilisées pour traiter des données tabulaires. Enfin, une étude expérimentale a été menée sur le jeu de données composite *Heart_Disease_2024*, associé au BIG-Dataset (Tajnim 2024 ; Haque et al. 2024), comprenant 4 238 observations, 14 variables cliniques et démographiques, et une cible binaire indiquant la présence ou l'absence d'une maladie cardiovasculaire.

La méthodologie adoptée a couvert les étapes essentielles d'une chaîne de prédiction : traitement des valeurs aberrantes, partition stratifiée des données, normalisation des variables continues, gestion du déséquilibre des classes par SMOTE, optimisation des hyperparamètres avec Optuna et évaluation finale sur un ensemble de test indépendant. Cinq modèles ont été comparés : régression logistique, XGBoost, MLP, TabNet et ResNetTab.

## Bilan des résultats

Les résultats apportent une réponse nuancée à la problématique. Sur le jeu de données étudié, le MLP atteint le même niveau de discrimination que la régression logistique (AUROC de 0,692), tout en offrant le meilleur équilibre global entre les classes. ResNetTab se distingue par la sensibilité la plus élevée, ce qui le rend

intéressant pour un contexte de dépistage où les faux négatifs doivent être limités.

Ces résultats montrent que les modèles profonds ne remplacent pas systématiquement les modèles classiques, mais qu'ils peuvent apporter des profils complémentaires. Leur intérêt apparaît surtout lorsque l'analyse ne se limite pas à un score global et prend en compte la sensibilité, l'équilibre des classes et l'explicabilité des prédictions. L'analyse des variables influentes reste cohérente avec les connaissances cliniques, notamment pour l'âge, `oldpeak`, la pression artérielle systolique et le sexe. La calibration des probabilités demeure toutefois nécessaire avant toute utilisation pratique.

## Limites du travail

Ce travail reste limité par la nature du jeu de données. Les variables disponibles sont peu nombreuses et ne couvrent pas certains éléments importants, comme les biomarqueurs avancés, l'imagerie médicale, les traitements, les antécédents familiaux détaillés ou les données longitudinales. Le signal prédictif exploitable reste donc partiel.

L'évaluation repose aussi sur une seule base de données. Même si un ensemble de test indépendant a été utilisé, aucune validation externe n'a été réalisée sur une autre population. Les résultats doivent donc être confirmés sur d'autres cohortes avant d'envisager une utilisation clinique.

Enfin, l'explicabilité doit être interprétée avec prudence. Les analyses XAI aident à comprendre les facteurs associés aux décisions des modèles, mais elles ne prouvent pas de relation causale et ne remplacent pas l'expertise médicale.

## Perspectives

Les travaux futurs devraient d'abord porter sur la validation externe des modèles, idéalement sur plusieurs cohortes cliniques. Cela permettrait d'évaluer leur robustesse, leur stabilité et leur capacité à généraliser à d'autres profils de patients.

Une autre perspective consiste à enrichir les données utilisées. L'intégration de mesures biologiques plus détaillées, de données longitudinales, d'imagerie ou de signaux physiologiques pourrait améliorer la prédiction du risque cardiovasculaire et mieux exploiter les capacités des architectures profondes.

Enfin, l'amélioration de la calibration, le choix de seuils adaptés au contexte clinique et le renforcement des explications individuelles constituent des étapes im-

portantes. À terme, un outil d'aide à la décision pourrait être envisagé, à condition qu'il reste transparent, validé sur des données externes et utilisé comme support au jugement médical, non comme substitut.

# Références bibliographiques


Arik, Sercan O. et Tomas Pfister (2021). « TabNet : Attentive Interpretable Tabular Learning ». In : *Proceedings of the AAAI Conference on Artificial Intelligence*. T. 35. 8, p. 6679-6687. doi : 10.1609/aaai.v35i8.16826.

Arrieta, Alejandro Barredo et al. (2020). « Explainable Artificial Intelligence (XAI) : Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI ». In : *Information Fusion* 58, p. 82-115. doi : 10.1016/j.inffus.2019.12.012.

Ba, Jimmy Lei, Jamie Ryan Kiros et Geoffrey E. Hinton (2016). *Layer Normalization*. doi : 10.48550/arXiv.1607.06450. arXiv : 1607.06450 [stat.ML].

Benahmed, Selim, Laurence Beutler et Boris Chabloz (2015). « L'éducation thérapeutique post-infarctus du myocarde : le défi de la prévention tertiaire ». Mémoire de bachelor. Haute École Arc Santé. url : https://sonar.ch/global/documents/313493 (visité le 13/06/2026).

Bishop, Christopher M. (2006). *Pattern Recognition and Machine Learning*. New York : Springer. isbn : 978-0-387-31073-2. url : https://link.springer.com/book/9780387310732.

Breiman, Leo (2001). « Random Forests ». In : *Machine Learning* 45.1, p. 5-32. doi : 10.1023/A:1010933404324.

Cascino, Thomas et Michael J. Shea (déc. 2023). *Introduction au diagnostic des maladies cardiovasculaires*. Manuels MSD pour le grand public. url : https://www.msdmanuals.com/fr/accueil/troubles-cardiaques-et-vasculaires/diagnostic-des-maladies-cardiovasculaires/introduction-au-diagnostic-des-maladies-cardiovasculaires (visité le 13/06/2026).

Centre hospitalier universitaire de Rouen (2026). *Circulation du sang dans le cœur*. url : https://www.chu-rouen.fr/wp-content/uploads/sites/66/elementor/thumbs/Circulation-du-sang-dans-le-coeur-q9jgod1fegn6ua50ksomj7qp0qs45wgu9usrgag23k.png (visité le 13/06/2026).

Chawla, Nitesh V. et al. (2002). « SMOTE : Synthetic Minority Over-Sampling Technique ». In : *Journal of Artificial Intelligence Research* 16, p. 321-357. doi : 10.1613/jair.953.

Chen, Tianqi et Carlos Guestrin (2016). « XGBoost : A Scalable Tree Boosting System ». In : *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 785-794. doi : 10.1145/2939672.2939785.

Chirurgie Vasculaire Lyon (2026). *Hypertension artérielle*. url : https://www.chirurgie-vasculaire-lyon.com/sites/default/files/inline-images/hypertension-arterielle.jpg (visité le 13/06/2026).

Chollet, François (2021). *Deep Learning with Python*. 2e éd. Manning. isbn : 978-1-61729-686-4.

Cover, Thomas M. et Peter E. Hart (1967). « Nearest Neighbor Pattern Classification ». In : *IEEE Transactions on Information Theory* 13.1, p. 21-27. doi : 10.1109/TIT.1967.1053964.

D'Agostino, Ralph B. et al. (2008). « General Cardiovascular Risk Profile for Use in Primary Care : The Framingham Heart Study ». In : *Circulation* 117.6, p. 743-753. doi : 10.1161/CIRCULATIONAHA.107.699579.

Datta, S. et I. K. Gang (2021). « Robust cardiovascular disease prediction using logistic regression ». In : *Journal of Management and Engineering Integration* 14, p. 26-36. url : https://doi.org/10.62704/10057/24766.

Docteurclic (2026). *Angine de poitrine*. url : https://www.docteurclic.com/galerie-photos/image_2911_400.jpg (visité le 13/06/2026).

Doshi-Velez, Finale et Been Kim (2017). *Towards A Rigorous Science of Interpretable Machine Learning*. arXiv : 1702.08608 [stat.ML].

Dreamstime (2026). *Athérosclérose et artériosclérose*. url : https://thumbs.dreamstime.com/b/ath%C3%A9roscl%C3%A9rose-et-art%C3%A9rioscl%C3%A9rose-31746909.jpg (visité le 13/06/2026).

Ester, Martin et al. (1996). « A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise ». In : *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. AAAI Press, p. 226-231.

Esteva, Andre et al. (2019). « A Guide to Deep Learning in Healthcare ». In : *Nature Medicine* 25.1, p. 24-29. doi : 10.1038/s41591-018-0316-z.

Freepik (2026). *Ensemble de deux battements de cœur en ligne ECG*. url : https://img.freepik.com/vecteurs-libre/ensemble-deux-battements-coeur-ligne-ecg_78370-1007.jpg?semt=ais_hybrid&w=740&q=80 (visité le 13/06/2026).

Gaut (2026). *Cours de physiologie : le cœur*. url : http://f2.quomodo.com/5C852034/uploads/8489/physiologie (visité le 13/06/2026).

Géron, Aurélien (2022). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow : Concepts, Tools, and Techniques to Build Intelligent Systems*. 3e éd. O'Reilly Media. isbn : 978-1-098-12597-4.

Goodfellow, Ian, Yoshua Bengio et Aaron Courville (2016). *Deep Learning*. Cambridge, MA : MIT Press. isbn : 978-0-262-03561-3. url : https://www.deeplearningbook.org/.

Google Images (2026). *Cœur normal et cardiomyopathie dilatée*. url : https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSErHCgJH6I3HyrZT5cmE_6Hfh93nzZVfEqxA&s (visité le 13/06/2026).

Gorishniy, Yury et al. (2021). « Revisiting Deep Learning Models for Tabular Data ». In : *Advances in Neural Information Processing Systems*. T. 34, p. 18932-18943.

Grinsztajn, Léo, Edouard Oyallon et Gaël Varoquaux (2022). « Why Do Tree-Based Models Still Outperform Deep Learning on Tabular Data ? » In : *Advances in Neural Information Processing Systems*. T. 35, p. 507-520.

Guyon, Isabelle et André Elisseeff (2003). « An Introduction to Variable and Feature Selection ». In : *Journal of Machine Learning Research* 3, p. 1157-1182. doi : 10.1162/153244303322753616. url : https://jmlr.csail.mit.edu/papers/volume3/guyon03a/guyon03a.pdf.

Haque, Mahfuzul et al. (2024). « Multi-class Heart Disease Detection, Classification, and Prediction Using Machine Learning Models ». In : *2024 27th International Conference on Computer and Information Technology (ICCIT)*. IEEE, p. 1375-1380. doi : 10.1109/ICCIT64611.2024.11022576.

Hastie, Trevor, Robert Tibshirani et Jerome Friedman (2009). *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*. 2e éd. New York : Springer. doi : 10.1007/978-0-387-84858-7.

He, Kaiming et al. (2016). « Deep Residual Learning for Image Recognition ». In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 770-778. doi : 10.1109/CVPR.2016.90.

Hendrycks, Dan et Kevin Gimpel (2016). *Gaussian Error Linear Units (GELUs)*. doi : 10.48550/arXiv.1606.08415. arXiv : 1606.08415 [cs.LG].

Ioffe, Sergey et Christian Szegedy (2015). « Batch Normalization : Accelerating Deep Network Training by Reducing Internal Covariate Shift ». In : *Proceedings of the 32nd International Conference on Machine Learning*. T. 37. Proceedings of Machine Learning Research. PMLR, p. 448-456.

Janin, Émilie (24 sept. 2015). « Participation du pharmacien d'officine à l'éducation thérapeutique des patients après un programme de réadaptation cardiaque : pérennisation de l'observance médicamenteuse, du respect des règles hygiéno-diététiques et du maintien de l'activité physique ». Thèse de doctorat en pharmacie. Université de Lorraine, Faculté de pharmacie. url : https://docnum.univ-lorraine.fr/public/BUPHA_T_2015_JANIN_EMILIE.pdf (visité le 13/06/2026).

Jolliffe, Ian T. (2002). *Principal Component Analysis*. 2e éd. New York : Springer. doi : 10.1007/b98835.

Lachance, Kim (2015). « Facteurs de risque d'insuffisance rénale chronique chez les greffés cardiaques : du phénotype aux tests pharmacogénomiques ». Thèse de doct. Université de Montréal. url : https://library-archives.canada.ca/eng/services/services-libraries/theses/Pages/item.aspx?idNumber=1163632985 (visité le 13/06/2026).

Laroche, Jean-Pierre et al. (2005). « Médecine vasculaire ». In : *Concours Médical* 127.35, p. 3-4.

LeCun, Yann, Yoshua Bengio et Geoffrey Hinton (2015). « Deep Learning ». In : *Nature* 521.7553, p. 436-444. doi : 10.1038/nature14539.

Loshchilov, Ilya et Frank Hutter (2019). « Decoupled Weight Decay Regularization ». In : *International Conference on Learning Representations*. url : https://openreview.net/forum?id=Bkg6RiCqY7.

Lundberg, Scott M. et Su-In Lee (2017). « A Unified Approach to Interpreting Model Predictions ». In : *Advances in Neural Information Processing Systems*. T. 30.

Maaten, Laurens van der et Geoffrey Hinton (2008). « Visualizing Data using t-SNE ». In : *Journal of Machine Learning Research* 9.86, p. 2579-2605. url : https://www.jmlr.org/papers/v9/vandermaaten08a.html.

MacQueen, James (1967). « Some Methods for Classification and Analysis of Multivariate Observations ». In : *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. T. 1. University of California Press, p. 281-297.

Montgomery, Douglas C., Elizabeth A. Peck et G. Geoffrey Vining (2012). *Introduction to Linear Regression Analysis*. 5e éd. Hoboken, NJ : Wiley. isbn : 978-0-470-54281-1.

Ornstein, Steven M. et al. (5 oct. 2004). « A Multimethod Quality Improvement Intervention to Improve Preventive Cardiovascular Care : A Cluster Randomized Trial ». In : *Annals of Internal Medicine* 141.7, p. 523-532. doi : 10.7326/0003-4819-141-7-200410050-00008.

Rajkomar, Alvin, Jeffrey Dean et Isaac Kohane (2019). « Machine Learning in Medicine ». In : *The New England Journal of Medicine* 380.14, p. 1347-1358. doi : 10.1056/NEJMra1814259.

Ray, Marie-Céline (26 avr. 2017). *AVC : accident vasculaire cérébral*. url : https://www.futura-sciences.com/sante/dossiers/medecine-avc-accident-vasculaire-cerebral-947/ (visité le 13/06/2026).

Ribeiro, Marco Tulio, Sameer Singh et Carlos Guestrin (2016). « "Why Should I Trust You ?" : Explaining the Predictions of Any Classifier ». In : *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 1135-1144. doi : 10.1145/2939672.2939778.

Rudin, Cynthia (2019). « Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead ». In : *Nature Machine Intelligence* 1.5, p. 206-215. doi : 10.1038/s42256-019-0048-x.

Santé sur le Net (2026). *Cardiologie : généralités sur le cœur, rythme cardiaque, structure anatomique*. url : https://www.sante-sur-le-net.com/maladies/cardiologie/generalites-coeur/ (visité le 13/06/2026).

Sarker, Iqbal H. (2021). « Deep Learning : A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions ». In : *SN Computer Science* 2.6, p. 420. doi : 10.1007/s42979-021-00815-1.

Shwartz-Ziv, Ravid et Amitai Armon (2022). « Tabular Data : Deep Learning Is Not All You Need ». In : *Information Fusion* 81, p. 84-90. doi : 10.1016/j.inffus.2021.11.011.

Streamlit (2026). *Streamlit*. url : https://streamlit.io/ (visité le 16/06/2026).

Szegedy, Christian et al. (2016). « Rethinking the Inception Architecture for Computer Vision ». In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 2818-2826. doi : 10.1109/CVPR.2016.308.

Tajnim, Mahfuzul Haque (2024). *Heart_Disease_2024*. url : https://www.kaggle.com/datasets/mahfuzulhaquetajnim/heart-disease-2024.

Thompson, Andrea D. et Michael J. Shea (août 2024). *Présentation des symptômes des maladies cardiovasculaires*. Manuels MSD pour le grand public. url : https://www.msdmanuals.com/fr/accueil/troubles-cardiaques-et-vasculaires/sympt%C3%B4mes-des-maladies-cardiovasculaires/vue-d-ensemble-des-sympt%C3%B4mes-des-maladies-cardiovasculaires (visité le 13/06/2026).

Vapnik, Vladimir N. (1995). *The Nature of Statistical Learning Theory*. New York : Springer. doi : 10.1007/978-1-4757-2440-0.

— (1998). *Statistical Learning Theory*. New York : Wiley. isbn : 978-0-471-03003-4.

Vaswani, Ashish et al. (2017). « Attention Is All You Need ». In : *Advances in Neural Information Processing Systems*. T. 30, p. 5998-6008. arXiv : 1706.03762.

Weng, Stephen F. et al. (2017). « Can Machine-Learning Improve Cardiovascular Risk Prediction Using Routine Clinical Data ? » In : *PLOS ONE* 12.4, e0174944. doi : 10.1371/journal.pone.0174944.

World Health Organization (2025). *Cardiovascular Diseases (CVDs)*. url : https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (visité le 13/06/2026).

Zhou, Chunjie et al. (2024). « A Comprehensive Review of Deep Learning-Based Models for Heart Disease Prediction ». In : *Artificial Intelligence Review* 57.9, p. 263. doi : 10.1007/s10462-024-10899-9.

الجمهورية الجزائرية الديمقراطية الشعبية

Republique Algérienne Démocratique Populaire
وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur de la Recherche Scientifique

Université 20 août 1955 Skikda

جامعة 20 أوت 1955 – سكيكدة

Faculte des Sciences
Departement d'informatique

كلية العلوم
قسم الإعلام الآلي

الرقم: .... / ق/ إ/ أ/ ك.ع2026/

## Autorisation de Dépôt de Mémoire de Master

Je soussigné : ...Dr.....Selma....Hadiby.....

Certifie que l'étudiant(e) : ....Lounissi....Bouchis.....

Spécialité : ....Intelligence Artificielles (IA).....

Ayant soutenu le projet intitulé : prédiction des maladies cardiovasculaires a' l'aide du DL

A apporté les corrections nécessaires sur son manuscrit de Master.

Signature de l'encadreur