

L'Intelligence Artificielle Générative et Prédictive au laboratoire : Le dépistage de PTH élevée pathologique chez des patients dialysés.

Y. MISSIAK¹, M. BERNIER¹, GUINAULT², A. LECOUR¹, D. MILONGO², C. ROSSIGNOL¹, A. DELARCHE², O. MILIOT², P. DAHAN²., T. JAUDON¹

1 -INOVIE CBM TOULOUSE; 2 -CLINIQUE NEPHROLOGIE ST EXUPERY TOULOUSE

Introduction :

Le suivi à long terme des patients dialysés est un défi majeur pour les centres de dialyse. La surveillance de la Parathormone (PTH) est un enjeu majeur pour le suivi thérapeutique à long terme des patients dialysés. L'Intelligence Artificielle (IA) pourrait répondre au besoin d'un dépistage précoce de la PTH. Nous voulons savoir si en générant de nouveaux patients "artificiels" grâce à une IA Générative française **Mixtral** (créée par **Mistral AI** et similaire à **ChatGPT**) , nous pouvons améliorer la prédiction d'un taux pathologique de PTH chez des patients dialysés à partir de variables clinicobiologiques.

Matériel et Méthode :

314 patients inclus (année 2023), avec 40 variables clinico-biologiques (dont PTH, âge, sexe). Classés en groupe **PTH Elevée Oui / Non** selon les recommandations néphrologiques KDIGO (PTH Pathologique > 331.2 pg/ml). Le langage Python a été utilisé pour générer les IA Prédictives Arbres Décisionnels, Forêts Randomisées, Deep Learning et l'IA Générative **Mixtral 8x7b** (version Dolphin).

Sur les 314 patients de départ, 32 patients ont été mis de côté pour l'évaluation de la performance (Précision, Aire sous la courbe AUC, PTH Elevée OUI / NON). L'IA Générative Mixtral a été utilisée pour générer de nouveaux patients (avec une similarité de 95%) à partir des 282 patients restant. (*cf Figure 1*) Ces patients "artificiels / simulés" ont ensuite servi de données d'entraînement pour nos modèles d'IA Prédictifs. Pour comparer l'apport de Mixtral, les IA Prédictives ont aussi été évaluées seule (sans IA Générative Mixtral). Les performances de statistiques classiques traditionnelles régression linéaire et logistique ont aussi été mesurées.

Résultats :

28482 nouveaux patients "artificiels" ont été générés grâce à l'IA Générative **Mixtral**. Pour la cohorte avec IA Générative (n = 28482), les taux de bonnes réponses pour les Arbres Décisionnels, **Forêts Randomisées**, et Deep Learning étaient respectivement de 81%, **97%** et 78%. Les AUC respectifs étaient de 77%, **94%** et 75%. Pour la cohorte sans IA Générative (n = 282), les taux de bonnes réponses étaient respectivement de 75%, 88% et 56%. Les AUC respectifs étaient de 73%, 78% et 58%. De plus, on obtient une Valeur Positive Prédicive (VPP) pour les Forêts Randomisées avec IA Générative de 99%. (*cf Figure 2*)

Discussion et Conclusion :

On remarque une meilleure performance des modèles IA Prédictifs Arbres Décisionnels et Forêts Randomisées à type "arbre" par rapport aux modèles de Deep Learning. Grâce à l'IA Générative **Mixtral** couplée à l'IA Prédicative Forêts Randomisées, nous obtenons un AUC 94% (comparé à 78% sans Mixtral) avec une VPP de 99%.

En cas d'usage, si un dépistage par l'IA couplée revient positif (IA : PTH Elevée), un dosage systématique de la PTH serait proposé au service de dialyse. La PTH étant dosée habituellement tous les 4 mois pour les patients dialysés, l'IA permettrait ainsi de détecter précocement une valeur pathologique de PTH lors d'un bilan usuel mensuel

En conclusion, la génération de patients artificiels grâce à l'utilisation d'une IA Générative **Mixtral** a permis de considérablement augmenter la valeur prédictive de nos modèles. Ce travail montre aussi que les IA Génératives de type **ChatGPT** et dans notre cas, **Mixtral**, peuvent être capables d'une certaine compréhension des pathologies médicales.

Bibliographie

Apport de l'IA pour l'analyse des paramètres biologiques du patient dialysé en relation avec le taux de parathormone . Y.Missiak Y, Bernier M, Guinault D, Lecour L, Rossignol C, Delarche A, Milongo D, Milioto O, Dahan P, Jaudon T. 66ème JIB, Paris, 2023.

Correspondant : Dr. Yohann Missiak, LBM INOVIE CBM, Toulouse / Muret, France, yohann.missiak@inovie.fr

Conflits d'intérêts : Aucun

Figure 1 : Flowchart de l'étude

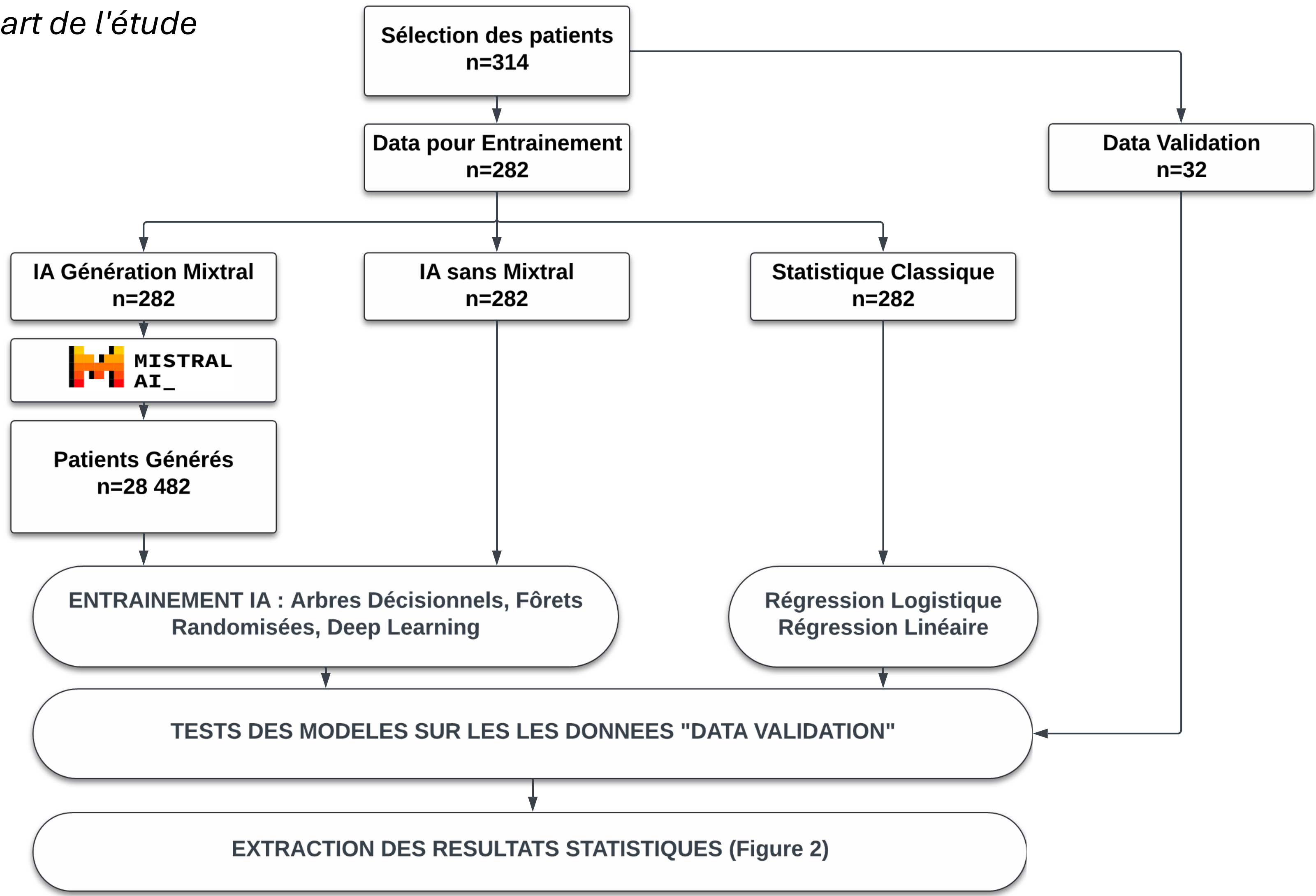


Figure 2 : Comparaison des performances des Modèles Intelligence artificielle développés en interne

Classification PTH Elevée OUI / NON		Taux de bonnes réponses "Précision"	AUC Aire sous la Courbe
Avec IA Générative Mixtral n = 28 482	Arbre Décisionnel	81 %	77 %
	Forêt Randomisée	97 %	94 %
	Deep Learning	78 %	75 %
Sans IA Générative Mixtral n = 282	Arbre Décisionnel	75 %	73 %
	Forêt Randomisée	88 %	78 %
	Deep Learning	56 %	58 %
Statistiques Traditionnelles Absence IA, n = 282	Régression Linéaire	59 %	58 %
	Régression Logistique	62 %	59 %