Jérémie NDJOYE

06.17.87.43.91 | ndjoyej@gmail.com | LinkedIn | jeremiendjoye.fr

FORMATION

Université de Poitiers — **Poitiers, France**
Bachelor universitaire de technologie en Science des données – **Major de promotion** — 2024-2027

Lycée Calasanz — **Libreville, Gabon**
Baccalauréat général série C – **4e au classement national** — 2021-2024

EXPÉRIENCE PROFESSIONNELLE & ENGAGEMENT ASSOCIATIF

Institut National de la Santé et de la Recherche Médicale (INSERM) — **Poitiers, France**
Alternance – Assistant ingénieur — **Septembre 2025 – Présent**

- Modélisation statistique : mise en œuvre de modèles non linéaires à effets mixtes et de régressions linéaires pour estimer les paramètres pharmacocinétiques, analyser la variabilité interindividuelle et contribuer à l'optimisation des schémas posologiques.
- Automatisation de la validation de données : développement d'une application Python intégrant ces modèles statistiques et des contrôles qualité, réduisant le temps de traitement de plus de **20 minutes** à moins de **30 secondes** par analyse.
- Digitalisation des processus du laboratoire : développement d'un portail web centralisant les outils internes et automatisant la gestion des stocks de réactifs, ramenant le traitement des inventaires de **2 jours** à moins de **2 minutes**.

Data analyst | Notre Gabon (ONG jeunesse) – À distance — **Janvier 2026 – Présent**

- Conception et développement de Trinity, plateforme interne de pilotage des actions et de l'impact *(TypeScript, PostgreSQL)* pour 5 pôles et environ 20 utilisateurs : suivi de plus de 30 indicateurs et envoi automatisé d'e-mails.
- Nettoyage et analyse de données de questionnaires de satisfaction et rédaction de rapports à destination des pôles.

PROJETS

Scoring de crédit *(Python, scikit-learn, XGBoost, SHAP)*

- Modèle de scoring du risque de défaut sur 30 000 clients de carte bancaire : comparaison de la régression logistique, du Random Forest et de XGBoost ; gestion du déséquilibre des classes (22 % de défauts) par pondération et optimisation du seuil.
- Gain de 2,3 points d'AUC par rapport à la régression logistique (0,776 → 0,799 ; Gini 0,552 → 0,598) ; interprétation des variables déterminantes par SHAP.

Prévision de la volatilité : GARCH/EGARCH vs Machine Learning (Python, arch, XGBoost, PyTorch)

- Prévision de la volatilité du S&P 500 (estimateur de Parkinson) à J+1 et J+5 sur 12 559 séances (1977-2026) : comparaison de GARCH/EGARCH, d'un HAR enrichi, d'un XGBoost et d'un LSTM.
- Validation walk-forward sur 2 ans hors échantillon (504 séances) : XGBoost réduit la RMSE de 15 à 16 % par rapport au GARCH recalibré et de 4 à 7 % par rapport au HAR enrichi (écart significatif en QLIKE, test de Diebold-Mariano, $p < 0,05$) ; le LSTM bat le HAR à J+1 (-8 % de RMSE) sans gain significatif sur XGBoost.

COMPÉTENCES, LANGUES ET CENTRES D'INTÉRÊT

Compétences techniques :

- **Programmation :** Python (Pandas, NumPy, SciPy, scikit-learn, PyTorch), R, SAS, SQL, VBA.
- **Statistiques & modélisation :** Régression linéaire, logistique et régularisée, ACP, séries temporelles, modèles multivariés.
- **Machine Learning :** Classification (arbres, Random Forest, gradient boosting), clustering (K-means, CAH), validation croisée, métriques de scoring (AUC, Gini, KS).
- **Visualisation & outils :** Power BI, Tableau, Streamlit, Git, Docker, MkDocs.

Langues : **Français** (natif), **Anglais** (B2 – CLES).

Centres d'intérêt : Football (niveau amateur) | Course à pied | Boxe anglaise | Piano (niveau amateur) | Voyages.

DISTINCTIONS ACADÉMIQUES

Lauréat du Prix de l'Excellence de l'Étudiant Africain (PEEA), édition 2026 — **Paris, France**

- Distinction fondée sur les résultats académiques et l'engagement associatif.

Vainqueur du Concours national de Dataviz 2025 de l'INSEE — **Paris, France**

- Analyse statistique des données de l'INSEE sur la scolarisation, le chômage et l'insertion professionnelle des 15-29 ans dans les territoires à faible densité et conception d'une infographie mettant en évidence les disparités territoriales.