

Amine NAKROU

Ingénieur Data Science & Actuariat

Assurance santé • Machine Learning • Data Engineering

CDI • Disponible septembre 2027 • France / Luxembourg

aminenakrou635@gmail.com
+33 7 44 28 00 10
Villeneuve-d'Ascq, France
linkedin.com/in/amine-nakrou
github.com/aminenakrou
aminenakrou.github.io/portfolio

Profil

Futur ingénieur **Polytech Lille (ISIA, grade Master)**, en alternance à la **Direction technique de Swiss Life** sur les contrats **santé collectifs**. J'allie **rigueur actuarielle** et **compétences d'ingénieur data** : fiabilisation de données de prestations à grande échelle, automatisation des contrôles et **modèles prédictifs interprétables** au service des décisions techniques.

Expérience professionnelle

Chargé d'études actuarielles (alternance) | **Swiss Life France** — Sept. 2026 – Août 2027 • Lille

Direction technique • Assurance non-vie • Contrats santé collectifs d'entreprise

- **Automatise le rapprochement des prestations santé** entre les **délégataires de gestion** et le système de gestion interne (**SAS 9.4**, **SQL**) ; analyse et explique les écarts auprès de l'équipe actuarielle.
- **Industrialise les rapports de contrôle** (synthèses, matrices **survenance** × **exercice**, contrôles de cohérence), désormais produits en un seul traitement.
- **Fiabilise les données de prestations** : normalisation des identifiants de contrat, détection et correction d'anomalies.
- **Développe un modèle de prédiction de la sinistralité** par contrat collectif (**XGBoost**, interprétation **SHAP**), alternative à la méthode **Chain-Ladder**, en appui de la tarification et du renouvellement (en cours).
- **Diffuse les bonnes pratiques Python** : guide d'installation et modèle de projet pour l'équipe.

Stage Data Engineering & BI | **GRDF – DSI** — Avr. – Août 2026 • Saint-Denis

- **Conçu une dizaine de tableaux de bord Power BI** de pilotage de la DSI, consolidant les indicateurs clés d'**une quinzaine d'outils** (Jira, GitLab, SonarQube, outils de déploiement, Kubernetes).
- **Développé les pipelines d'extraction** en **Python** (API REST / OData), automatisés en **GitLab CI/CD** et historisés.
- **Modélisé les données en étoile** (Power Query, Dataflows) et défini les indicateurs **DAX**.

Stage Développeur Data & Web | **HAMATEL** — Juin – Août 2025 • Maroc

- **Automatisé les rapports financiers** d'un groupe de grande distribution : **une semaine de travail ramenée à 2 heures**.
- **Développé une application Flask** de gestion des chantiers, des équipes et des interventions.

Projets

Tarification automobile – fréquence × coût moyen — Python • GLM • XGBoost • freMTPL2

- GLM **Poisson / Gamma** vs XGBoost sur **678 000 contrats** : prime pure, **Gini 0,32 → 0,37**.

Provisionnement – Chain-Ladder vs Machine Learning — Python • Mack • Cape Cod • XGBoost

- Backtest sur **343 triangles réels** : Chain-Ladder, Mack, Cape Cod vs XGBoost (erreur **6,6 → 4,5 %**).

Pipeline de données médaillon – compagnie aérienne — Python • DuckDB • Dagster • SQL

- Ingestion quotidienne, modèle **bronze / silver / gold**, historisation **SCD2**, orchestration **Dagster**.

Prédiction du churn client – de la donnée à l'application — scikit-learn • Streamlit • Docker

- Régression logistique, Random Forest, SVM (**AUC 0,85**) ; application **Streamlit** en ligne, **Docker**.

Formation

Diplôme d'ingénieur (grade Master) — 2024 – 2027
Polytech Lille – ISIA
Statistiques, ML, Data Mining, MLOps, data engineering.

Classes préparatoires MPSI / MP — 2022 – 2024
Lycée Moulay Idriss, Fès

Bac Sciences Mathématiques — 2022
Mention Très Bien – Maroc

Langues & intérêts

Français courant • **Anglais** C1 • **Arabe** maternelle • **Espagnol** A2

Football, voyages

Compétences

Actuariat / Assurance : santé collective, sinistralité, Chain-Ladder, survenance / exercice, contrôle des prestations, gestion déléguée
Machine Learning : régression, classification, Random Forest, XGBoost, SHAP, deep learning (TensorFlow / Keras, PyTorch), statistiques
Langages : Python, R, SQL, SAS, VBA, HTML / CSS / JavaScript
Librairies : pandas, NumPy, scikit-learn, Matplotlib, Plotly, Flask, Streamlit, R Markdown
Bases de données : SAS Enterprise Guide, Snowflake, DuckDB, PostgreSQL, MySQL, MongoDB
Data Eng. / MLOps : Spark, Airflow, Dagster, Docker, Kubernetes, MLflow, Git / GitHub, GitLab CI/CD, API REST / OData
Cloud / Systèmes : Azure, AWS, Linux
BI / Reporting : Power BI (DAX, Power Query, Dataflows), Tableau, Excel avancé
Outils : Jupyter, VS Code, Jira, SonarQube, SharePoint, LeanIX