

Baptiste Gerbouin

06 41 19 06 50 | bgerbouin@gmail.com | [LinkedIn](#) | [github](#) | [portfolio](#)

FORMATION

Université de Bordeaux

Date prévue d'obtention: Septembre 2025

Master 2 : CMI ISI - Cours Master en Ingénierie, Ingénierie de la Statistique et de l'informatique

Université Libre de Bruxelles - Erasmus

Date d'obtention: Juin 2023

Licence 3 : CMI ISI

COMPÉTENCES

Programmation : Python (Pandas, Numpy, Scipy, Scikit Learn, Plotly, Dash, Pytorch, TensorFlow, Matplotlib) / R / C / SQL(PostgreSQL) / NoSQL (MongoDB) / Git

Mathématiques : Probabilités / Statistiques (et en grande dimension) / Optimisation / Algèbre linéaire

Informatique : Big Data / Système de Gestion de Bases de Données (SGBD) / Machine Learning / Deep Learning / Apprentissage supervisé - non supervisé / Réseaux de neurones / Fine Tuning

Soft-Skills : Autonome, sérieux, rigoureux, travail d'équipe.

Hard-Skills : Azure / Docker

EXPÉRIENCES PROFESSIONNELLES

Stage Data Scientist – SUEZ

Été 2025

- Amélioration des modèles hybrides via un feature engineering ciblé.
- Développement d'une base scalable pour l'intégration de données environnementales massives (ERA5-Land, SAFRAN, RADAR), plus de 11 To de données.
- Exploration statistique avancée (régression robustes, AUC) et visualisation interactive (Dash + Flask).
- Évaluation des modèles sur plusieurs horizons à l'aide du NSE.

EXPÉRIENCES UNIVERSITAIRES

Dashboard d'analyse des ressources en eau en France

Septembre 2024 à Février 2025

- Développement d'un tableau de bord avec Dash pour le suivi en temps réel des ressources en eau en France.
- Conception d'un pipeline automatisé en Python pour traiter les données via une API, avec stockage dans MongoDB.
- Amélioration de la précision des analyses en temps réel, aidant la prise de décision et la visualisation des données.

Défi IA - Prédiction de réachat de produits - CARREFOUR

Septembre 2024 - en cours

- Développement d'un modèle prédictif en machine learning pour recommander des produits sur carrefour.fr, en analysant environ 85 millions de lignes issues de données clients et produits.
- Optimisation des recommandations client via des méthodes de deep learning et de systèmes de recommandation avancés.
- Évaluation des performances du modèle avec la métrique Hit Rate @10, visant une amélioration significative du taux de réachat.

Création et Développement d'une Base de Données INSEE

Janvier 2024 à Avril 2024

- Conception d'une base de données relationnelle avec PostgreSQL pour gérer des données massives de l'INSEE.
- Automatisation des processus via des scripts Python et optimisation des requêtes avec des triggers et index.
- Réduction du temps de réponse des requêtes de 30%, améliorant l'efficacité des analyses démographiques.

Monte Carlo Tree Search (Machine Learning)

Septembre 2023 à Janvier 2024

- Développement d'un modèle d'apprentissage par renforcement appliqué au Morpion pour optimiser la prise de décision.
- Conception d'algorithmes d'exploration/exploitation pour améliorer les stratégies de jeu via la méthode MCTS.
- Réduction du temps de réponse des requêtes de 30%, améliorant l'efficacité des analyses démographiques.

INFORMATIONS COMPLÉMENTAIRES

Permis B / Français, Anglais, Espagnol

Bénévolat : Courses solidaires pour les associations 'Les blouses roses' et 'Chiens guides d'aveugles'