

Anais STORP

☎ +33 6 38 55 40 06 • ✉ anais.storp@polytechnique.edu • in anais-storp

Education

École Polytechnique & HEC Paris

M.Sc. in Data Science & Artificial Intelligence for Business

Palaiseau & Jouy-en-Josas

Sep 2025 – Present

- Completed 8+ courses in machine learning, deep learning, database management and Python/R programming
- Worked in 4-person teams to analyse 5+ real-world datasets, presenting results to faculty and peers through data visualisation

Sorbonne Université

B.Sc. in Chemistry, Data Science & AI

Paris

Sep 2022 – Jun 2025

- **Major in Chemistry:** Carried out 20+ laboratory experiments using spectroscopic analysis and titration
- **Minor in Data Science & AI:** Developed 3+ database systems using SQL and Python; built visualisations for 10+ datasets

Harvard University (edX)

Certificate in Data Science with Python, Machine Learning & AI

Online

Jan 2024 – Apr 2024

- Implemented machine learning algorithms (KNN, Naive Bayes) on 3+ datasets, reaching 85%+ classification accuracy
- Completed 5+ programming projects applying regression, classification and supervised learning with cross-validation

Professional Experience

Safran Tech

Research Intern – Machine Learning for Materials Science

Paris

Apr 2026 – Aug 2026

- Developed and benchmarked machine learning models to predict the creep rupture life of nickel-based single-crystal superalloys from chemical composition, temperature and stress
- Designed and compared several learned embedding architectures in PyTorch (MLP, self-attention) against an XGBoost baseline, evaluated with R^2 , RMSE and RMSLE
- Implemented uncertainty quantification and Rashomon-set analysis to assess prediction reliability in a small-data regime
- Structured a reproducible codebase (Python, Git/GitLab) and presented results to supervisors and domain experts

Projects

ENS Data Challenge (team of 2)

QRT Asset Allocation Performance Forecasting – top 7% of ~1250 teams

2026

- Ranked in the **top 7% of ~1250 teams** on a binary classification task: predicting the sign of an asset's next-day return from 20-day price and signed-volume history (527k training rows)
- Engineered 400+ features across six successive packs, then pruned to a compact 25-feature set that transferred better ood
- Designed a rolling-block time-series cross-validation to prevent temporal leakage and a magnitude-weighted XGBoost ensemble blended in rank space with linear challengers (logistic regression, elastic net) under a rare-switch arbitration scheme

Database Management Project (team of 2)

Connected Component Finder – Distributed Graph Algorithms in Spark

2026

- Implemented the CCF connected-components algorithm (iterative MapReduce) in Apache Spark across four variants – Python and Scala, each with the RDD and DataFrame APIs
- Benchmarked scalability on synthetic graphs and real-world SNAP graphs, validating correctness against the reference paper

Personal Project

GENFORM – Data-Driven Ergonomic Analysis for Adjustable Fitness Equipment

2026

- Built a reusable, modular Python library translating ergonomic and biomechanical models from scientific papers into code
- Engineered a clean, reproducible pipeline running on ANSUR II, the US Army anthropometric reference dataset
- Built configurable sensitivity analyses and automated reports with visualisations to evaluate ergonomic design trade-offs

Skills & Interests

Programming: Python – Advanced • SQL – Advanced • Scala, HTML – Intermediate

Tools: LaTeX, Git/GitLab, Microsoft Office (Excel, Word, PowerPoint) – Intermediate

Languages: **French & German:** Native/Bilingual • **English:** C1 (CAE) • **Spanish:** B1 (CLES)

Interests: Cooking workshops, Skiing, Technic Lego building, Piano (5 years)