

# Romain Delhommais

Data Science & ML · ENSAE & IPParis

📍 Paris, France (CET) | ✉️ [romain@delhommais.com](mailto:romain@delhommais.com) | 🌐 [delhommais.com/cv](https://delhommais.com/cv)

in [linkedin.com/in/romain-delhommais](https://linkedin.com/in/romain-delhommais) | 🐙 [github.com/luuumine](https://github.com/luuumine)

## Education

**Polytechnic Institute of Paris**, Master's Degree in Data Science Sept 2024 – May 2025

- M2 Data Science
- **Core courses:** Deep Learning, Computer Vision, Representation Learning for Computer Vision and Medical Imaging, Audio and music information retrieval, Optimal Transport, Law and Ethics of AI

**ENSAE Paris**, Engineering Degree in Statistics and Data Science Sep 2022 - Jun 2025

- Data Science, Statistics and Learning track
- **Core courses:** Bayesian Statistics, Infrastructures and Software Systems, Advanced Machine Learning, Algorithm Design and Analysis, Deep Learning: Models and Optimization, Parallel Programming, Reinforcement Learning

**Lycée Saint Louis**, (Undergraduate) Preparatory Classes MPSI-MP Sep 2020 - Jul 2022

- Intensive mathematics, physics and computer science program

## Experience

**Data Scientist Intern**, Apivia Courtage – Tours, France Jun 2025 - Dec 2025

- Built a churn prediction pipeline with XGBoost using time-travel snapshots and rolling time-series CV.
- Engineered subscription-level features (tenure, contract changes, interaction history), addressed class imbalance, and automated ROC-/PR-AUC reporting with reproducible, config-driven runs and versioned artifacts (plots and JSON metrics).
- Ran an exploratory personas study (UMAP + k-means/HDBSCAN) to probe behavioral patterns. Assessed cluster quality/stability and documented results.

**Digital Marketing Services Intern**, HILTI Group – Boulogne-Billancourt, France Jun 2024 - Sep 2024

- Conducted multivariate analyses of customer behavior using Python (pandas, sklearn), including feature engineering and statistical testing (e.g. significance levels, p-values).
- Performed a statistical health check of the customer behavior analysis pipeline, validating data quality and consistency at each step with descriptive metrics and diagnostic tests.

**Digital Marketing Services Intern**, HILTI Group – Boulogne-Billancourt, France Jun 2023 – Jul 2023

- Cross-matched sales and marketing data to quantify campaign impact and build a reusable analytical tool.

## Research Projects

**Natural Language Processing**, L'Oréal Jan 2025 - Mar 2025

- Trained a BERT sentiment model and compressed it via LoRA, teacher-student distillation, and 8-bit quantization, cutting size/latency on custom F1 loss.

**Natural Language Processing**, INRIA Saclay Oct 2023 - May 2024

- Trained sub-100M-param Language Models on English Wikipedia via Transformers on AWS GPUs.

## Skills

### Programming

- Languages: Python, SQL, Bash, JavaScript/TypeScript
- ML/Data: scikit-learn, PyTorch, pandas, XGBoost
- Tooling: Docker, Linux, Git, uv, LaTeX, DBeaver
- Other: HTML, CSS, R, OCaml

### Languages

- French: C2 level - Native
- English: C1+ level - Fluent (Linguaskill score: 180+)
- Spanish: A2 level - Beginner
- German: A2 level - Beginner