Aymen Sadraoui, Ph.D.

Data Scientist & Data Engineer | Machine Learning, Deep Learning & Big Data Systems

Paris, France · aymensadraoui1919@gmail.com · +33 7 65 68 52 67 · Portfolio · LinkedIn

Professional Summary

Data Scientist and **Machine Learning Ph.D.** graduate specialized in building robust **predictive models** and scalable **data engineering pipelines** for high-dimensional, multimodal datasets. I bring extensive hands-on expertise in **deep learning**, **representation learning**, and **few-shot learning**, maintaining a rigorous evaluation-first methodology focused on **model calibration**, **out-of-distribution generalization**, and reproducibility. Proven track record of translating complex mathematical research into first-author publications at top-tier venues (**CVPR**, **ISBI**, **ICIP**) and engineering production-grade **ETL pipelines** for AI systems. I am seeking roles where I can architect end-to-end **data pipelines** and deploy scalable, impact-driven **AI/ML solutions**.

Technical Skills

Core Tech: Linux, Python, PyTorch, OpenCV, Git, CI/CD, Docker, MLflow, Streamlit, ETL/ELT Pipelines

Languages: English (Fluent), French (Fluent), Arabic (Native)

Professional Experience

Data Scientist & Machine Learning Engineer — **Feb 2026 — Present**
Deemea — *Paris, France*

- Architected and deployed end-to-end **automated data processing and ingestion pipelines** integrating directly with hospital PACS (DICOM) and cloud platforms for multi-centric clinical studies.
- Developed deep learning-based algorithms for **automated image quality control (QC)**, detecting motion artifacts, slice mismatches, and data corruption prior to clinical review.
- Engineered computer vision algorithms for **automated diagnosis, segmentation, and classification** of radiology scans (CT, X-ray, MRI) across orthopedics, oncology, and rheumatology modalities.
- Deployed scalable, containerized microservices (**Docker, Streamlit, PyTorch, CI/CD**) into production workflows, powering AI-driven pre-annotations and interactive 3D visualizations for clinical operators.
- Implemented robust **MLOps frameworks** for model calibration, and data drift detection to ensure high out-of-distribution generalization across multi-site hospital data.
- Partnered with cross-functional teams of radiologists, software engineers, and clinical managers to translate clinical research requirements into production-ready software features.

Ph.D. Researcher in AI & Computational Pathology — **Jan 2022 — Jan 2026**
CentraleSupélec (OPIS / CVN Lab) & Kremlin-Bicêtre Hospital — *Gif-sur-Yvette, France*

- **Thesis**: *AI-Driven Methods in Histopathology for the Diagnosis and Prognosis of Hepatocellular Carcinoma*.
- **Clinical-Technical Interface**: Served as the primary bridge between ML researchers and senior pathologists; translated qualitative histopathological criteria into quantifiable ML objectives and validated automated features against clinical consensus.
- **Advanced Deep Learning & Optimization**: Designed and unrolled proximal optimization algorithms (Unrolled Projected Gradient, UNEM) into neural networks for robust stain separation, color normalization, and transductive few-shot slide classification.
- **Terabyte-Scale Data Engineering**: Built scalable ETL preprocessing pipelines (PyTorch, OpenCV, Scikit-learn, OpenSlide, Optuna) to ingest, tile, and curate over 4+ TB of gigapixel whole slide images, handling high-throughput data processing and multi-source artifact removal.
- **Multi-Modal Predictive Modeling**: Developed deep learning models integrating histopathologic architecture quantification with clinical/prognostic clinical metadata to predict early liver cancer recurrence.
- **MLOps & Trustworthy AI**: Created rigorous benchmarking frameworks focused on out-of-distribution generalization, model calibration, uncertainty estimation, and interpretability for high-stakes medical deployment.

Achievements

- **1st Place Winner**: Lung Transplant Rejection Challenge — Organized by Hôpital Foch & Health Data Hub (2024)
- **National Rank 63 / 1,773 (Top 3%)**: National Engineering Entrance Exam, Tunisia (2018)

Education

Ph.D. in Machine Learning & Artificial Intelligence — **2022 — 2026**
CentraleSupélec, Université Paris-Saclay — *Gif-sur-Yvette, France*

Engineering Degree in Signals and Systems Processing — **2018 — 2021**
École Polytechnique de Tunisie — *La Marsa, Tunisia*

Publications

- **Sadraoui, A.**, *et al.* (2026). A Proximal Approach for Stain Separation and Normalization of Whole-Slide Histopathological Images. *IEEE 23rd International Symposium on Biomedical Imaging (ISBI)*.
- **Sadraoui, A.**, *et al.* (2025). UNEM: UNrolled Generalized EM for Transductive Few-Shot Learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- **Sadraoui, A.**, *et al.* (2024). Deep Learning Classification and Quantification of Pejorative and Nonpejorative Architectures in Resected Hepatocellular Carcinoma from Digital Histopathologic Images. *The American Journal of Pathology*
- **Sadraoui, A.**, *et al.* (2024). Unrolled Projected Gradient Algorithm for Stain Separation in Digital Histopathological Images. *IEEE International Conference on Image Processing (ICIP)*
- **Sadraoui, A.**, *et al.* (2024). A Transductive Few-Shot Learning Approach for Classification of Digital Histopathological Slides from Liver Cancer. *IEEE International Symposium on Biomedical Imaging (ISBI)*