Léo BURGUND

leo.burgund@gmail.com | Paris, France | +33 6 84 93 75 58 | github.com/leob000

Master 2 Math/AI student with research experience at INRIA and former medical student, seeking a 5-6 month AI/ML internship (Apr-Sep 2026).

Experience

AI Research intern - Attention Growth for Transformers (report link) Apr. 2025 – Aug. 2025

INRIA Saclay, LISN, TAU Team

- Enabled Transformers to grow capacity when useful, improving flexibility without retraining from scratch (closed-form, on-the-fly attention-dimension growth via functional gradients; head-selection criterion).
- Built and trained a custom Growing Vision Transformer on CIFAR-10/100 and Imagenette (PyTorch; Slurm GPU cluster; experiment tracking with Weights & Biases).
- Research manuscript in preparation.

Medical extern 2018 – 2021

Assistance Publique - Hôpitaux de Paris

- Assisted medical residents in patient care and provided surgical assistance in the operating room.

Education

Master 1 & 2: Mathematics & AI 2024 – 2026

Paris-Saclay University (Mathematical Institute of Orsay), joint program with CentraleSupélec Paris

- Focus: Statistics/Optimization/Probability, ML/DL, Computer vision, NLP (applications in Python and R)
- Including courses from the MVA track santé (ENS Paris-Saclay) program

Double-Licence in Applied Mathematics and Economics 2021 – 2024

Paris-Nanterre University

General Training Diploma in Medical Sciences (first 3 years of M.D.) 2016 – 2019

Faculty of Medicine, Sorbonne University

- Including one year at the University of Copenhagen through the Erasmus program

Projects

Fine-tuning, Pruning and Quantization of DINOv3 ViT-S/16 for Histology (github link) 2025

- Cancer-patch detection (PatchCamelyon dataset): AUROC 0.980, Sens@95%Spec 0.920 with full fine-tuning; compared results to LoRA and head-only training.
- Implemented architecture-preserving compression (attention-head & MLP pruning + per-layer SVD): 7% fewer params and GFLOPs for 0.025 AUROC trade-off.
- Quantization to bfloat16, halving model memory footprint with no significant performance loss.
- Easily reproducible pipeline (Local and Slurm GPU cluster), tracking with Weights&Biases.

Selected Master 1 & 2 coursework projects 2025

- Spotify track popularity prediction using audio features and metadata (link).
- Electricity Demand Forecasting using “classical” machine learning methods (link) & using Deep learning methods (link).
- Interactive web application for visual question answering for the Clevr dataset (NLP/Vision interaction) (link).

Skills

- Programming: Python (PyTorch, Pytest, Scikit-learn, Numpy, Pandas, Matplotlib, OpenCV, Streamlit), Bash, R, SQL, C
- Tools: Git, HPC use (via Slurm), Weights&Biases, Hugging Face, Timm, Linux, LaTeX, Typst, Neovim
- Languages: French (native), English (Fluent, TOEIC 955/990)