

# Dhia Garbaya

Paris | dhia.garbaya@eleves.enpc.fr | in/dhia-garbaya

dhia680.github.io | github.com/dhia680

## Profile

Currently focusing on transfer learning, pre-training LLMs, and reasoning. Interested in how far we can push deep learning (scientific discovery?) and also how best to model learning: from experience (RL) or/and from imitation. I am seeking research internship opportunities (April 2026), ideally with a CIFRE PhD follow-up.

## Education

**ENS Paris-Saclay, MVA: Math, Vision, Learning** 2025 – 2026

- **Coursework:** Optimal transport · Convex optimization · Reinforcement learning · RMT · Probabilistic graphical models · Deep learning theory · Geometric deep learning · Graphs in ML · Vision · Robotics ...

**Ecole Nationale des Ponts et Chaussées (IP Paris), MS in Applied Math and ML** 2022 – 2026

- **Coursework:** Machine learning (JAX) · Deep learning (torch) · Optimization · Operations research (Julia) · Advanced algorithmics (C++ ) · Vision · Stochastic processes · Statistical physics · Stats in high dimensions · PDEs · Game theory...

**Esprit Prépa, Tunis, Prepa for Grandes Ecoles** 2020 – 2022

- **Coursework:** General, Linear Algebra/ Topology/ Analysis/ Theoretical Physics/ Eng sciences
- **Project:** Optical and thermodynamical optimization of solar cells.
- **Distinction:** Valedictorian

## Experience

**Visiting Student | Research Intern, EPFL, Switzerland** Feb 2025 - Now

- **Lab:** MLO
- **Supervisor:** Prof. Martin Jaggi
- Core member of **Swiss AI Initiative, Apertus LLM** team (training on ALPS cluster).
- Focus: pre-training (1k-10k GPUs), knowledge distillation, efficient training, model architecture.
- Exploring: Reasoning and over-thinking

**Research Intern, Technology Innovation Institute – Abu Dhabi** Jul 2024 – Jan 2025

- Focus: knowledge distillation, scalable optimization algorithms (1st and 2nd order)
- Large-scale transformer parametrization, param-free learning, pre-training (5k GPUs)
- Core contributor to **Falcon3** family of models.

## Publications

**FOG Architectures\*** : Towards Pure FP8 LLM Training at Scale · **NeurIPS'25** Preprint - V1  
EPFL & ETH Zurich - 2025

**Apertus\*** : Democratizing Open, Compliant, and Multilingual LLMs Technical Report  
Swiss AI team - 2025

**Falcon3 family of Open Models\*** : Showcasing transfer learning efficiency Official blogpost  
Falcon LLM team - 2024 (undisclosed report)

\* first-authorship / core-contribution (paper / release)

## Coding skills

---

**Languages:** Python, C++, Julia, R.

**Tools, frameworks:** Pytorch, JAX, Megatron-LM, NeMo, HF libraries, GIT, AWS, cluster computing, linux

## Academic Projects

---

### Solving PDEs with PINNs, ENPC+Airbus

github/dhia680/pinns-24

- Studied Physics-Informed-Neural-Networks for solving Helmholtz PDEs.

### Deep RL for optimizing traffic using autonomous vehicles (exploratory)

github/.../HighwayEnv

- Inspired from CIRCLES project (Berkley, Rutgers, Ecole des Ponts..).
- Studied phantom traffic jam, drivers behaviour. Used existing codebase to train a policy with DQN, PPO.
- Tools: Pytorch, Sumo simulator.

### RNN-based NMT model

github/dhia680/NMT

- Trained an LSTM-based machine translation toy model (20M) and integrated it in a web interface
- Tools: Tensorflow, Pandas.

### Operations Research

2023-2024

- Optimizing an offshore wind electrical network for RTE, FR
- Optimizing a car manufacturing chain for Renault, FR
- Tools: Python, C++, Julia, Gurobi (MIP), S.Annealing

## Distinctions, Online certificates

---

- **French government excellence scholarship:** Among 7 national holders of this scholarship, 2022-2026.
- **NLP** with Python, Udemy.
- **Supervised Learning, RL**, deeplearning.ai.
- **In progress:**
  - Advanced RL in python, DQNs, Udemy.
  - Diffusion Models, Nvidia.
  - Accelerated computing with cuda (python and C++), Nvidia.

## Languages

---

- **Arabic:** Native
- **French:** Fluent
- **English** Fluent
- **German:** Intermediate

## Volunteering

---

- **Reviewer** for MenaML'25.
- **Project Manager** in the Junior Enterprise of Ecole des Ponts for 14 months.