Dhia Garbaya

dhia.garbaya@eleves.enpc.fr

dhia680.github.io | github.com/dhia680

Profile

Currently focusing on *reasoning, distillation, and pre-training* of language models.

I am seeking a 5-6 month **research internship** (April 2026), followed by a PhD.
An ideal topic would be centered around *RL, representation learning*, and grounded in principled approaches.

Education

École Normale Supérieure (Saclay), MVA: MSc (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), MEng (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 · PDEs · Game theory...

Esprit Prépa, Tunis, Prépa (Math, Physics) 2020 – 2022

- **Coursework:** Algebra/ Topology/ Analysis/ Probability/ Theoretical Physics/ Algorithms...
- **Project:** Optical and thermodynamical optimization of solar cells.

Experience

Research Intern, EPFL – Switzerland Feb 2025 - Now

- Teams: **MLO** lab & **Swiss AI Initiative**
- Supervisor: Prof. Martin Jaggi
- Focus: efficient LLM pre-training (1k+4k GPUs), knowledge distillation, model architecture.
- Recently: over-thinking · exploration in RL (context: reasoning LMs).

Research Intern, TI Institute – UAE Jul 2024 – Jan 2025

- Teams: **AI theory** & **Falcon**
- Supervisor: Dr. MEA Seddik
- Focus: knowledge distillation, scalable optimization algorithms (1st and 2nd order)
- Large-scale parametrization (MuP), param-free learning, pre-training (5k GPUs)

Publications

FOG Architectures* : Toward Pure FP8 LLM Training at Scale · **NeurIPS'25** Paper v2
EPFL & ETH Zurich - 2025

Apertus* : Democratizing Open, Compliant, and Multilingual LLMs Technical Report (under review)
Swiss AI team - 2025

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

* first-authorship

Coding skills

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

Tools, frameworks: Pytorch, JAX, Megatron-LM, NeMo, HF libraries, GIT, AWS, Slurm.

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.

RNN-based NMT model github/dhia680/NMT

- Trained an LSTM-based machine translation tiny model and integrated it in a web interface

Operations Research 2023-2024

- Optimizing an offshore wind electrical network for RTE, FR
- Optimizing a car manifacturing chain for Renault, FR

RL for push recovery of a wheeled robot 2025

- PPO, Reward shaping, curriculum, inductive bias (supervised by the Willow team at INRIA)

Denoising score matching 2025

Distinctions, Online certificates

- **French government excellence scolarship**: Among 7 national holders of this scolarship, 2022-2026
- **Valedictorian** in both high school and Prépa
- **Completed certificates**: NLP with Python (Udemy) · Supervised Learning, RL (deeplearning.ai)

Languages

- **Arabic** (Native) · **English** (Fluent) · **French**: (Fluent) · **German**: (Intermediate)

Volunteering

- **Reviewer** for MenaML'25.
- **Project Manager** in the Junior Enterprise of Ecole des Ponts (2023 - 2024).