

Pau Hidalgo Pujol

Barcelona / Banyoles, Spain | +34 682 49 56 07 | pauhidalgopujol@gmail.com
github.com/pauhidalgoo | huggingface.co/pauhidalgoo | [LinkedIn](#)

Profile

BSc student in Artificial Intelligence at UPC with ETH Zürich exchange and research experience in geometric deep learning for protein graphs. Interested in mathematics for machine learning, graph representation learning and interpretable neural systems. Seeking advanced mathematical training in geometry, probability/statistics and learning theory toward a research internship and possible PhD.

Education

Exchange Student, Computer Science **Sep. 2025 – Feb. 2026**
ETH Zürich *Zürich, Switzerland*

- Master-level coursework (25 ECTS): Probabilistic AI, Neural Network Theory, Reliable and Trustworthy AI, Learning in Deep Artificial and Biological Neuronal Networks, Artificial Intelligence in Education.
- Key topics: Bayesian inference, VC dimension, approximation theory, convex relaxations, randomized smoothing.
- Grade average: 5.0/6.0. Artificial Intelligence in Education: 6.0/6.0.

Bachelor's Degree in Artificial Intelligence (240 ECTS) **Sep. 2022 – Jun. 2026**
Universitat Politècnica de Catalunya (UPC), Barcelona School of Informatics *Barcelona, Spain*

- Average mark: 8.83/10. Compulsory courses outside initial stage: 9.1/10. Bachelor's thesis: 10/10.
- Ranking (mobility): 13/192 among FIB exchange applicants, based on a weighted academic-record score across BSc/MSc applicants.
- Honors in Introduction to Machine Learning and Parallelism & Distributed Systems.
- Relevant mathematical/ML coursework: Statistics (9.6), Statistical Modelling (9.5), Introduction to ML (9.1), Advanced Data Analysis (9.7), RL/Unsupervised Learning (9.4), Optimization, CV, NLP, Algebra, Calculus.

Research Experience

Bachelor's Thesis - Human Attention for Large Language Model Training **Mar. 2026 – Jun. 2026**
Collaboration with Telefónica I+D and UPC *Grade: 10/10*

- Studied whether human eye-tracking signals can guide, evaluate, and interpret Large Language Models. Built controlled analyses separating gaze from lexical confounds such as word length and frequency, tested gaze-informed training and reward-model interventions and performed mechanistic diagnostics with attention-head analyses.
- Received a perfect grade of 10/10 and proposed for Honors distinction.

AI Research Intern **Feb. 2026 – Present**
Nostrum Biodiscovery *Barcelona, Spain*

- Researched geometric deep learning for protein structure modelling, focusing on graph construction, message passing, and efficient GearNet variants.
- First-author accepted GRaM @ ICLR 2026 blogpost, *Fewer Edges, Faster Protein Graph Learning*: introduced Angle Rewiring and a FiLM-style Efficient IEConv variant for protein graphs.
- Extended Efficient IEConv to AlphaFoldDB-based pretraining for GearNet variants; achieved state-of-the-art Fold3D performance in internal experiments.
- Effective-resistance-based graph rewiring experiments for over-squashing: implemented optimized resistance computations using sparse eigensolvers and Johnson-Lindenstrauss approximations, designed heuristics for directed rewiring, and evaluated Relational GCN variants on additional benchmarks.

Selected Research Output and Projects

Fewer Edges, Faster Protein Graph Learning **GRaM Blogposts @ ICLR 2026**

- Accepted blogpost; first author with Manel Gil-Sorribes, Alexis Molina, and Bertran Miquel-Oliver. Studies topology, geometric inductive biases, memory-efficient protein GNNs, and Pareto trade-offs across EC, GO, and Fold3D benchmarks.

Cucafera - Catalan LLM **Personal project**

- Built and trained a Catalan-focused 244M-parameter LLaMA-like language model from scratch on 5.5B tokens; curated datasets, implemented the model in PyTorch/Hugging Face, and released models/datasets publicly.

Other projects

- BitsXLamarató BSC challenge, 2nd place: high-resolution NO₂ estimation with U-Net. University projects in NLP, Deep RL, autoresearch, speech recognition, robotics, recommender systems, statistical analysis, and computer vision.

Awards and Achievements

- **HackUPC 2025 Winner:** Grafana challenge, interactive dashboard for Sustainable Development Goals.
- **DatathonFME 2023 & 2024 Winner:** clothing recommendations; Chicago housing price prediction.
- **AINAHack Winner:** LLM-powered assistant for the Catalan administration, with integrated web search.
- **Challenge IA i Auditoria:** time-series forecasting and autoencoders for inventory-impairment prediction.
- **Proves Cangur Mathematics Competition:** ranked 95th in 2018, 191st in 2017 and 171st in 2022 in Catalonia.
- **Premi Astrobanyoles 2022:** high-school research project on neural-network-based melody composition.
- PAU Recognition: scored above 9.5/10 in the general university entrance examination phase.

Additional Background

Summer Monitor and Teacher

Jul. 2021, Jun. 2023, Jun. 2024

Codelearn, Banyoles

- Taught programming fundamentals to children and led a week-long introductory machine-learning workshop.
- Professional Music Diploma in Cello
- 2020 – 2023
- Conservatori de Música de Girona
- Scientific-Technological Baccalaureate
- 2020 – 2022
- Institut Pere Alsius, Banyoles Average grade: 9.24/10.

Skills

Programming: Python, C, C++, SQL, R, JavaScript.

ML / AI: PyTorch, Hugging Face, TensorFlow, scikit-learn, OpenCV, TorchDrug; geometric deep learning, GNNs, LLMs, NLP, Reinforcement Learning, Computer Vision, Probabilistic ML.

Other: CUDA basics, AMPL, PDDL, CLIPS, Git, Linux.

Languages: Catalan and Spanish (native), English (C2, IELTS 8.5), French (school-level coursework), German (A1.1).