Eric Alcaide Medicine & Physics, Machine Learning

Location: Barcelona, Spain | ericalcaide1@gmail.com | Tel: +34625429886 | GitHub: @hypnopump

Skills

Multi-Language Coding
Python (and scientific stack pytorch, triton, cython, jax, etc), Web, JavaScript, Bash, C++

Cloud & HPC
Algorithmic optimization, CPU (1000+ cores) and GPU parallelism (1000+ A/H1/200s), server and cluster computing

Machine Learning
LLMs, Geometric deep learning, computer vision, natural language, clustering, graphs, self-supervised learning, etc

Teamwork,
Public Speaking,
Fast Prototyping,
Problem Solving,
Strategic Thinking,
Operational Excellence,
Complex Problems

Education

Barcelona, Spain

Medicine Degree, *University of Barcelona*
Medical Degree. Multiple distinctions.

Barcelona, Spain

Physics Degree, *University of Barcelona*
Theoretical Physics Mention

Professional / Research Experience

2024 – present
Switzerland

Founding Engineer, *poolside*
LLMs. (Pre)Training, Inference, ML operations. Distributed workloads on thousands of CPU/GPUs. Algoritgms and Execution. From core research to product.

2021 – 2024
London,
United Kingdom

Machine Learning & Translational Scientist, *CHARM Therapeutics*
From Bits to Molecules, and Everything in between: data ingestion pipeline (1000s of CPUs), geometric deep learning research, model training (100s of GPUs), evaluation, drug target research, virtual screening, etc. First employee, architect of DragonFold

2021 – present

Open Source Researcher, *EleutherAI, OpenBioML*
Research at the intersection of Natural Language Processing, Structural Biology and High Performance Computational Methods. Author of RWKV, an RNN Model for the Transformer Era (up to 14B models). RWKV versions 4, 5, 6, and 7. Diverse international research collaboration for the promotion of Open Source AI. Lead efforts of teams of 25+ researchers

2020 – 2021

Machine Learning Researcher, *VIR Biotechnology*
Machine Learning for Structural Biology. Geometric Deep Learning and Natural Language Processing techniques for organic molecules, proteins and monoclonal Antibodies (mAbs).

2019 – 2021

Private Machine Learning Tutor
Personalized advice to Masters' students from different backgrounds (from Computational Linguistics to Biomedical Engineering) on Text Classification, Image processing and Information representation. Advised how to carry out Masters' Thesis-level projects.

2019

Non-Profit Health Hackathon Mentor, *TV3 - La Marató*
Advised on healthcare and science X algorithms, including genetic clustering and molecular dynamics for protein conformational changes. Won several prizes.

Peer Reviewed Papers

2024

Eagle and Finch: RWKV with Matrix-Valued States and Dynamic Recurrence,
CoLM, Conference on Language Modelling
Next generation RWKV 5 & 6. Best scalable RNNs for their time. 0.1B - 14B models

2023

RWKV: Reinventing RNNs for the Transformer Era, *EMNLP*
Rearrangement of typical RNNs compute graph to allow for GPT training. Trained multilingual Large Language Models from 0.1B to 14B parameters, with chat interface and Open Source release. Future developments sponsored by the Linux Foundation.

2023

UMD-fit: Generating Realistic Ligand Conformations for Distance-Based Deep Docking Models, *NeurIPS 2023 Generative AI In Biology Workshop*
Addressed chemical inaccuracies in Deep Learning Molecular Docking models.

2023

Advancing structural biology through breakthroughs in AI,
Current Opinion in Structural Biology
Major recent advances driven by technology and applications to novel therapeutics.

2022

InterDocker: End-to-End Cross-Attentive and Geometric Transformers for Efficient Iterative Protein Docking, *LMRL-Learning Meaningful Representations of Life, NeurIPS2022*

2022

Relevance of myocardial injury biomarkers to the prognosis of COVID-19 patients,
Revista Española de Cardiología
COVID19 related revision of predictive power of myocardial injury biomarkers (NT-proBNP and hs-TnT) regarding Mechanical Ventilation and Deatch Events.

2021

MP-NeRF: Massively Parallel Natural Extension of Reference Frame,
Journal of Computational Chemistry
Parallelized the Natural Extension of Reference Frame for folding polymers (proteins, RNA, etc) from internal angles, 1000x faster. Usage in MD simulations and ML training.

Courses

2020 – 2020
Barcelona, Spain

HPC-based Computational Biomedicine, *Barcelona Supercomputing Centre*
Impact and Hands-on experience of applied supercomputing to biomedical problems (molecular simulations, genomic analysis, tissue modelling, etc.)

2018 – 2019

Deep Learning, Natural Language Processing and AI for Medicine, *Coursera*
Contents include: foundations of Deep Learning, project management, Computer Vision, sequential data, Natual Language Processing, AI in healthcare, etc

2017 – 2017

Artificial Intelligence Micromasters Program, *Columbia University*
CSMM.101x: Artificial Intelligence (AI) - (through edx.org). Average qualification: 8.1 / 10
Search methods, games, ML introduction, CSPs, NLP, robotics introduction, etc.

Side Projects

2020 – present

Open Source Contributions
Cutting-edge Open Source Software packages (Pytorch Geometric, Fastformers, etc)

2018 – present

Open Source projects
Projects and modules for scientific computing which recieved a high degree of community acceptance:

- **2023: Uni-Mol** ↗ : Accuracy improvements for ML-based molecular docking.
- **2021: AlphaFold2 open replication** ↗ **:** Main contributor to the Open Source effort for the replication (and improvement) of the AlphaFold2 architecture (state of the art, deep learning engine for protein structure prediction).
- **2021: E(n) Equivariant GNN, Geometric Vector Perceptron:** ↗ Graph Neural Network architectures on invariant representations in 3D or arbitrary dimensions.
- **2019: AlphaFold1 imitation: MiniFold**: ↗ Predict protein foldings from raw sequences
- **2018: Keras-WRN:** ↗ Wide Residual Networks for image recognition in Keras.

2017

Deep Learning - Can Computers Learn?
Research project focused on the AI and Deep Learning field, subfields and the state of the art techniques. Evolutionary Strategies for architecture optimization in Neural Networks.

Languages

Spanish	**English**	**German**	**Mandarin**
Native	C2 level	B1 level	HSK 2

Awards

2019 **ESADE - Accenture HealthHackathon Winnner Award,** *ESADE, Barcelona*

2019 **AlphaFold v1 Replication Contest Award, Nvidia Titan RTX,** *Nvidia*

2017 **Hackathon UPC Winner Award,** *HackUPC, Barcelona*