Alvaro Calafell

alvarocalafell@gmail.com | github.com/alvarocalafell | linkedin.com/in/alvarocalafell

PROFESSIONAL EXPERIENCE

Feb 24 - Present | **Stanford University** | **Stanford, CA**

AI/ML Researcher

- Developing HyperFast v2, a highly distributed meta-trained hypernetwork for instant tabular data classification, outperforming traditional ML methods and AutoML systems (NeurIPS 2023 publication).
- Scaled meta-learning capabilities by increasing training datasets from 300 to 4,000 using OpenML and Hugging Face APIs, resulting in a 14% improvement in model generalization across diverse data scenarios.
- Optimized model architecture using low-rank techniques, increasing performance by 3% while reducing parameter count by 10x, leading to a 30% reduction in inference time.
- Engineered efficient ML research pipelines using PyTorch and SLURM, reducing experiment cycle time by 10% and improving resource utilization across heterogeneous GPU clusters by 35%.

Apr 23 - Sep 23 | **Criteo** | **Barcelona, Spain**

Data Analyst Internship

- Developed a tailored audiences tool using HiveSQL and MySQL, resulting in a 17% increase in conversion rates and a 15% boost in campaign ROI.
- Created the CPC/CPM Benchmark Insight tool, increasing client retention by 8% and improving overall client satisfaction scores by 12%.
- Designed and executed A/B testing frameworks for targeting strategies, leading to a 13% improvement in ad spend efficiency and an 11% increase in conversion rates.
- Analyzed the impact of continuous vs. seasonal ad campaigns, identifying key performance indicators that optimized ad strategy, and presented actionable insights to stakeholders, driving evidence-based decision making.

Mar 21 - Jun 21 | **Conecta2 Energia** | **Barcelona, Spain**

Data Analyst Internship

- Optimized energy purchase strategies by developing a predictive model with <4% variability, resulting in a €200k cost reduction and 10% improvement in energy efficiency.
- Implemented an optimization model for new photovoltaic solar plants, reducing startup costs by 12% and increasing energy output by 8% in the first year of operation.

PROJECTS

Aug 2024 | **Vector Database API [GitHub]**

- Developed a scalable REST API for vector database management and k-NN search, showcasing proficiency in building AI-driven backend solutions.
- Implemented multiple indexing algorithms (KD-Tree, Ball Tree, Brute Force) from scratch, demonstrating in-depth understanding of search optimization techniques.
- Designed and implemented a Docker and Kubernetes deployment solution with Helm charts, exploring modern DevOps practices for AI application deployment.

Jul 2024 | **Stable Diffusion from Scratch [GitHub]**

- Built a complete Stable Diffusion model using PyTorch, gaining hands-on experience with state-of-the-art generative AI techniques.
- Designed a modular project structure with separate components (encoder, decoder, CLIP, attention, diffusion, pipeline), showcasing software engineering best practices in AI development.
- Experimented with various optimization techniques for inference, exploring the balance between generation speed and image quality.

EDUCATION

2022 - 2024 | **Ecole Polytechnique & HEC Paris** | **Paris, France**

MSc in Data Science, GPA – 3.8/4.0

2017 - 2021 | **Universitat Politecnica de Catalunya (Barcelona-Tech)** | **Barcelona, Spain**

BSc in Industrial Engineering, GPA - 6.95/10 (Top 10%)

SKILLS & INTERESTS

Languages: Spanish and Catalan (Native), **English** (Proficient, Proficiency C2), **German** (Fluent, Goethe Institut B2).

Interests: Former national-level competitive skier, avid traveler, and culinary enthusiast. Currently training for first half marathon.