# Cristobal Lillo

+1 (786) 257-8741 | cal390@miami.edu | LinkedIn | GitHub

## Experience

**University of Miami** — Coral Gables, FL

*Research Assistant, Frost Institute for Data Science and Computing* — *Apr 2025 – Present*

- Collaborated to develop a new multi-model fusion framework for kidney cancer segmentation, combining predictions from diverse medical imaging models via self-attention.
- Engineered an ETL pipeline for DICOM MRI scans via TCIA's public API to support training and ROC-AUC validation of segmentation models.
- Containerized models and scaled training/inference across 4 L40 GPUs using PyTorch's Distributed Data Parallel (DDP) module.
- Formalizing attention-based late fusion segmentation architecture and its experimental performance in research paper; listed as third author on forthcoming publication.

*Teaching Assistant, Department of Computer Science* — *Aug 2024 – Present*

- Automated assignment collection, grading, and return for Computer Programming II by refactoring legacy Python 2.7 scripts to Python 3 using subprocess.
- Guided 500+ students through labs and office hours, teaching data structures, algorithms, and object-oriented programming in Java and Python.

**Examedi.cl** — Santiago, Chile

*Software Developer Intern* — *Jun 2024 – Aug 2024*

- Optimized practitioner travel scheduling using Django, OR-Tools, and Google Maps APIs with geospatial Redis caching.
- Developed Django REST API routes to enable seamless map integration in the mobile frontend.
- Wrote end-to-end automated tests with pytest to ensure robust routing and scheduling functionality.

**SmartUp** — Santiago, Chile

*Software Developer Intern* — *Dec 2024 – Jan 2025*

- Reduced AI agent response latency by 25% by migrating OpenAI API calls from completion to streaming packets, eliminating large packet losses.
- Built type-safe tRPC routes in Next.js to support streaming and ensure data and conversation concurrency.

## Projects

**tinyHPO** | *Python, tinygrad, tinyJIT* — *March 2025*

- Contributed hyperparameter optimization tools for RNNs, CNNs, and MLPs to the open-source tinygrad framework (30k stars).
- Created model-agnostic GridSearchCV and Bayesian optimization APIs compatible with all tinygrad backends.
- Reduced search times by implementing lazy evaluation with tinyJIT-based tensor operations.

**Graph-Based HRM** | *PyTorch, OR-Tools CP-SAT, Docker* — *July 2025 – Present*

- Adapted the HRM architecture for NP-hard graph-based Constraint Satisfaction Problems, tested on the TSP.
- Utilized CP-SAT for deterministic solution computation to evaluate feasibility and optimality of predictions.
- Built sparse edge embedding model for adjacency matrices, reducing embedding size whilst preserving graph structure.
- Combined supervised and reinforcement learning with problem-specific heuristics to improve training convergence and test accuracy.

## Education

**University of Miami** — Coral Gables, FL

*BS in Computer Science, BS in Mathematics, Minor in Computational Astrophysics* — *Expected May 2027*

- GPA: 3.86; Honors: Singer Scholar (Full Tuition Scholarship), Foote Fellow (Honors Program)
- Clubs: Math Union, Putnam Team, Alliance of Latin American Students, Club Football

## Skills

**Languages:** Python, C, Rust, Java, JavaScript/TypeScript, R, Bash/Zsh
**ML/AI:** PyTorch, JAX, tinygrad, CUDA, FlashAttention, Pandas, Polars, NumPy, GradCAM, Hugging Face Spaces
**Dev Tools:** Docker, AWS, GitHub Actions, SSH, PostgreSQL, Redis, Next.js, Vercel, tRPC, uv/pip, Cargo