

Abdiel Rodríguez Lara

Machine Learning Engineer

abdielr10330@gmail.com | +1 249-680-0549 | Montréal, QC (relocating Aug 2026) | Open work permit

linkedin.com/in/abdiel-rodríguez-lara | github.com/Ericks0330 | ericks0330.github.io

Summary

Machine Learning Engineer with a research background in robust and federated learning and 2+ years shipping production software. Undergraduate thesis designed a multi-aggregation operator that keeps federated training accurate under model-poisoning attacks, benchmarked against 5 aggregators and 6 attacks under a controlled experimental protocol. Trains, fine-tunes, and evaluates deep learning and computer vision models in PyTorch and YOLO, and builds the Python services, benchmarks, and tracing that turn model behaviour into numbers a team can act on. Relocating to Montréal in August 2026 on an open work permit.

Skills

Machine Learning & Deep Learning: PyTorch, TensorFlow, scikit-learn, NumPy, Pandas, Matplotlib, CNNs, supervised and unsupervised learning, data preprocessing, feature engineering, transfer learning and fine-tuning, hyperparameter tuning, cross-validation, design of experiments (DOE), model evaluation and benchmarking.

Federated & Robust ML: Federated learning (FLEX, FLEXClash), FedAvg, robust aggregation (Krum/Multi-Krum, Trimmed Mean, Median), model-poisoning and adversarial ML, differential privacy, client sampling and participation strategies, robustness benchmarking under attack.

Computer Vision: YOLO (Ultralytics), OpenCV, Roboflow, dataset curation and annotation, real-time inference, false-positive reduction with temporal logic.

Applied LLM & Retrieval: RAG (embeddings, chunking, retrieval), vector databases (Qdrant, Supabase/pgvector), LangChain, LangGraph, Google ADK, MCP, LLM evaluation and observability (LangSmith), prompt engineering, multi-agent systems, Anthropic Claude API, OpenAI API.

Software Engineering & Infrastructure: Python, TypeScript, Java, C, C++, FastAPI, REST APIs, PostgreSQL, PostGIS, Redis, Docker, CI/CD (GitHub Actions), pytest, Git and code reviews, Agile/Scrum, Jupyter and Colab notebooks, DigitalOcean, Render.

Machine Learning Research & Projects

MARFL: A Robust Multi-Aggregation Operator for Federated Learning | Undergraduate Thesis, CUJAE, 2026

- Designed a robust aggregation operator that defends federated training against model-poisoning attacks, composing robust aggregators along a temporal axis — Multi-Krum ($m = 5$) through the first half of training, switching at round 13 to a spatial multi-stage pipeline (Krum-exclude = 2 $\rightarrow$ Trimmed Mean 0.1, $\beta = 0.6$).
- Benchmarked it against 5 aggregators (FedAvg, Median, Trimmed Mean, Multi-Krum, Central-DP) under 6 poisoning attacks, sustaining 87% test accuracy, and ran the statistical comparison that isolates which composition strategy drives the robustness gain.

Experimental setup: CNN on MNIST, cross-entropy loss, SGD; 20 clients with 10 sampled per round (50% participation), IID partition, 25 training rounds; evaluated on held-out test accuracy.

Technologies: Python, PyTorch, federated learning (FLEX/FLEXClash), adversarial ML, design of experiments, NumPy, Matplotlib.

Real-Time Drowsiness Detection System | Computer Vision, 2026

- Trained a YOLO (Ultralytics) detector for eye state (open/closed) and yawning on a custom Roboflow dataset, fine-tuned in Google Colab, to infer driver fatigue from live video.
- Built a real-time Python inference application with OpenCV frame-by-frame detection and a state machine (sensitivity thresholds, debounce on yawns, temporal validation on sleep events) that cuts false positives, with a CustomTkinter GUI showing live bounding boxes, event logging, and Matplotlib analytics.

Technologies: Python, YOLO (Ultralytics), OpenCV, Roboflow, Google Colab, CustomTkinter, Matplotlib. Repository: github.com/Ericks0330/Real-Time-Drowsiness-Detection-System

Federated vs. Centralized Learning — Comparative Study | AI Research Developer, 2025

- Implemented federated training workflows and benchmarked federated against centralized models across 10 UCI datasets, comparing predictive performance, privacy guarantees, and robustness under a common evaluation protocol.

Technologies: Python, FLEX, scikit-learn, UCI Machine Learning Repository. Repository: github.com/AdrianRodriguezJorge/test_FL

Professional Experience

AIKoders — AI/ML Engineer & Full Stack Developer (full-time) | Feb 2026 – Present | La Habana, Cuba

AI product studio delivering ML-powered systems for hospitality, logistics, and marketing clients.

- Built the evaluation and tracing layer for the production LLM stack (LangSmith), scoring cost, latency, and task success per run; iterating on retrieval, prompts, and model routing against those traces cut token spend 26% and p95 latency from 11s to 7s.
- Shipped Guest Concierge (guest-concierge.ai), a production retrieval system over a pgvector/Supabase knowledge base with tuned chunking and embedding strategies, escalating critical cases to humans — owner-handled messages down from ~340 to ~155 per month across 12 properties (~54%).
- Built EasyPosting, a publishing pipeline that checks the customer's paid-plan entitlement, conditionally enhances submitted photos via Cloudinary, and generates brand-aligned captions from retrieved client context before auto-publishing to Instagram and Facebook.
- Deployed voice and WhatsApp assistants for EasyCargo (LatAm shipping) over a vector knowledge base of customs regulations plus live shipment APIs — 72% of ~1,200 monthly inquiries resolved without human handoff.

Technologies: Python, FastAPI, LangSmith, RAG, pgvector, Supabase, Anthropic Claude API, n8n, Vapi, PostgreSQL, Docker, DigitalOcean.

Guajiritos — ML/AI Engineer, Retrieval & Automation (contract, 2 months) | Jan 2026 – Feb 2026 | La Habana, Cuba

- Raised answer accuracy on proprietary business data from 52% to 87% (+35 pp) on a 500-question internal benchmark, by redesigning chunking and embedding strategy over vector databases and enforcing structured outputs.
- Built the offline evaluation harness scoring cost, latency, and task success per run, then used it to drive retrieval and model-routing changes that cut cost per query from \$0.045 to \$0.036 (~20%).

Technologies: Python, LangChain, vector databases, RAG architectures, MCP, n8n, OpenAI API, REST APIs.

Centro de Ingeniería Genética y Biotecnología (CIGB) — Bioinformatics Software Engineer (freelance, ~10 h/week) | Sep 2025 – Present | La Habana, Cuba

- Designed a quantitative scoring model producing normalized confidence scores for Dengue-LRP1 protein-protein interactions, weighting evidence type, interaction frequency, and dataset characteristics into a reproducible framework.
- Modeled heterogeneous experimental datasets into a single analyzable schema and shipped a production analytics platform with search by protein name, gene symbol, UniProt ID, and cell type on top of the scoring engine.

Technologies: Python, algorithm design, biological dataset modeling, Angular 21, TypeScript, REST APIs. Live: data-base-lrp1-web-app.onrender.com | github.com/Erick0330/Score-algorithm

A11ySolutions — QA Engineer → Web Accessibility Auditor (part-time) | Jan 2025 – Nov 2025 | La Habana, Cuba

- Delivered 10 months of functional, regression, and cross-browser testing on production platforms including BillShark.com (Mark Cuban-backed), and led WCAG 2.2 accessibility audits for Crownpeak enterprise sites.

INTRUST Security — Full Stack Developer (contract) | Nov 2025 – Jan 2026 | La Habana, Cuba

- Shipped a vulnerability-management platform (Spring Boot + Angular) consolidating 2 security scanners into a unified real-time risk dashboard, with automated domain-scanning tooling and versioned PostgreSQL migrations (Liquibase).

ALASolutions — Full-Stack Web Developer (part-time) | Oct 2024 – May 2025 | La Habana, Cuba

- Improved backend performance of a real-time GPS fleet-monitoring system (Traccar-based) and built API integrations for a smart EV-charging platform, including ETECSA Transfermóvil payments.

ETECSA — Backend Developer (contract) | Aug 2024 – Nov 2024 | La Habana, Cuba

- Designed the database and backend for a digital-transformation assessment platform that scores companies across dimensions, perspectives, and areas from structured questionnaires, with a team of three developers.

Education

B.Sc. Software Engineering — Ciudad Universitaria José Antonio Echeverría (CUJAE) | Feb 2023 – Jun 2026 | La Habana, Cuba | Thesis: MARFL — A Robust Multi-Aggregation Operator for Federated Learning.

Certifications & Languages

Machine Learning & Deep Learning — DataCamp, 2026. [Supervised](#) and [unsupervised predictive models](#), data preprocessing, and [deep learning with PyTorch](#).

AI Engineering & LLM Orchestration — DataCamp, 2026. [Multi-agent architectures \(LangGraph\)](#), [LLM pipeline orchestration \(LangChain\)](#), [Google ADK](#).

Languages — Spanish: first language. English: Advanced (C2), [EF SET certified](#).