Diego Ramirez Madriz

Chihuahua, Mexico | diegomadriz04@gmail.com | diegoramirezmadriz.dev | github.com/diegomadriz | LinkedIn

Education

Tecnológico de Monterrey — Chihuahua, Mexico
B.S. Computer Science – GPA 92/100 – Academic Talent Scholarship — *June 2026*

Universitat Politècnica de Catalunya (Exchange Semester) — Barcelona, Spain
Research Assignment on multi-agent LLM reasoning, graded 10/10 — *Sep 2025 – Jan 2026*

Experience

Technical Support Engineer Intern → Software Engineer Intern — Jun 2024 – Jun 2025
SYSCOM (B2B tech distributor) — *Chihuahua, Mexico*

- Built a production-ready React/Django REST application integrating AI-powered parking cameras via API, and handed it off with documentation ahead of its planned production deployment.
- Automated a manual data-extraction pipeline with Selenium, BeautifulSoup and pandas, cutting report generation time by 90%.
- Built a custom Django admin dashboard that let non-technical stakeholders configure system settings 50% faster.
- Hardened the backend with parameterized queries, input sanitization, CSRF protection and rate limiting, and wrote unit and integration tests with pytest and Jest.

Projects

Agentic Code-Delivery System | *TypeScript, Node.js, Git, Next.js, Redis* | diegoramirezmadriz.dev/system — 2026

- Built a code-delivery harness around the open-source Pi coding agent: a custom extension that limits it to bounded read, search and patch tools, isolated Git candidates, repository-owned deterministic checks, a separate read-only reviewer, and controller-owned pass/fix/escalate outcomes capped at two repair rounds.
- Bound checks, review and compare-and-swap promotion to one candidate identity, and delivered 111 agent-written changes to my portfolio repository: 62% passed review with no repairs, 91% within one (Sep–Oct 2026).
- Built the app's public AI endpoint, which answers only from an independently reviewed corpus, with HMAC-hashed client identifiers, daily-rotating salts and an atomic Redis rate limiter.

Multi-Agent LLM Reasoning Benchmark | *Python, llama.cpp, pandas* | GitHub — 2025

- Compared single-agent, Proposer→Critic→Refiner and Proposers+Judge orchestration of a local 4-bit Qwen2.5-7B on 200 GSM8K problems across three output caps, measuring tokens and seconds per correct answer.
- Proposer→Critic→Refiner reached 92.0% vs 79.5% for a single agent (57–61% relative error reduction across caps) at about 4× the tokens per correct answer. Splitting a small shared budget across judge candidates fell below the baseline.
- Packaged as a tested CLI (26 tests, CI) that reproduces the published results and checks prompt fidelity against the original experiment.

FutBot Tracker *(robot-soccer vision challenge)* | *Python, OpenCV, SAM 3* | GitHub — 2026

- Built the pipeline for a team competition entry: OpenCV color and shape proposals, SAM 3 mask refinement, class-aware tracking with persistent team identity, and possession estimation that separates visible-ball estimates from ownership carried through occlusion.

Fleet Maintenance & Inventory Analytics | *Python, pandas, SciPy, lifelines* | GitHub — 2026

- Built the analytics pipeline over 100+ mixer trucks and 250K+ maintenance, work-order and engine-hour rows: Croston-SBA intermittent-demand forecasting, Weibull and AFT reliability models, and sparse-data fallbacks.
- Produced scenario-based inventory and maintenance-priority outputs for human review, keeping analytical signals separate from validated procurement policy.

Skills

Languages: Python (pandas, NumPy, SciPy), TypeScript, JavaScript, Java, SQL, C++
AI & data: LLM agents and evaluation (Pi, Claude Code, Codex), llama.cpp, PyTorch, scikit-learn, OpenCV
Web & backend: React, Next.js, Django REST, Node.js, Spring Boot
Infrastructure: Git, GitHub Actions, Vercel, AWS EC2, GCP, PostgreSQL, MySQL, Redis
Spoken: Spanish (native), English (C1, TOEFL iBT 107)