

# Omar Pareja

Philadelphia, PA 19104

+1 301-385-4459 | pareja@engineering.upenn.edu | linkedin.com/in/omar-pareja- | github.com/Omar-Pareja | omarpareja.com

## EDUCATION

**University of Pennsylvania, School of Engineering and Applied Science**

Philadelphia, PA

Major: Artificial Intelligence, Bachelor of Science in Engineering (Expected May 2027) | GPA 3.5

- Key Coursework: Graph Neural Networks (graduate), Natural Language Processing, Computer Vision, Algorithms, Big Data Analytics

## PROFESSIONAL EXPERIENCE

### VideoAmp

New York, NY

*Machine Learning & Data Science Intern*

May 2026 – August 2026

- Built end-to-end conversion forecasting pipeline on 1.3B+ row ad datasets in Snowflake for guaranteed-outcome ad products.
- Trained an LSTM forecasting network and benchmarked it against SARIMA, Prophet, and LightGBM on MASE across horizons.
- Designed a config-driven forecasting framework for single-switch comparison across foot-traffic, pixel, and purchase data.
- Caught a measurement artifact where ~6x panel growth inflated conversion counts; built normalization to isolate true signal.
- Engineered data-quality guardrails (coverage tracking, dedup, partial-period handling) catching 100%-null identity columns.

### SwiftDoor

Philadelphia, PA

*Founder*

December 2025 – Present

- Built and run a white-label ticketing and payments platform for student organizations and live events (swiftdoor.co).
- Processed \$1.2M+ in ticket sales on a 5% fee model, with Stripe Connect checkout and automated organizer payouts.
- Shipped a drag-and-drop event-page builder, rule-based pricing with timed ticket releases, and QR door check-in.
- Built a 3D audience map that turns attendance data into a searchable network of groups and connections.

### Caffio

Philadelphia, PA

*Co-Founder*

September 2025 – October 2026

- Built a smart bottle that auto-dispenses caffeine to hold steady energy between user-set start and end times.
- Engineered ESP32-S3 BLE firmware, a React Native app, and a two-compartment Bateman pharmacokinetic dosing model.
- Validated the dosing model against peer-reviewed PK studies; added FDA-guideline guardrails and user-sensitivity limits.
- Ran a survey funnel of 2,000+ responses and 50+ beta users to segment caffeine habits and tune dosing recommendations.
- Filed a provisional patent; won the M&T Innovation Fund and Penn Wharton Innovation Fund Design Award.

### Ieng Group

New York, New York

*Machine Learning Intern*

July 2025 – September 2025

- Built and shipped an anomaly-detection service flagging power-grid failures across distributed sensors in near real time.
- Trained XGBoost models on large sensor datasets; tuned precision/recall around false-alarm vs. missed-failure cost.
- Engineered the feature pipeline (rolling-window aggregations, lag features) with retraining as new telemetry arrived.
- Validated detection thresholds with offline backtests and live A/B tests, balancing early warning against alert fatigue.
- Surfaced model scores and likely root causes in a triage dashboard, cutting mean time-to-resolution for grid engineers.

### Siren Analytics

Hybrid

*Software Engineering Intern*

June 2025 – July 2025

- Built and shipped full-stack features for a civic-information platform serving 10K+ users (React, FastAPI, PostgreSQL).
- Shipped a citizen record-search and document-upload module end-to-end, from Postgres schema to REST APIs to React UI.
- Owned backend data models and REST endpoints alongside engineering leads, from spec through production.
- Wrote unit and integration tests for new endpoints, reducing post-release regressions.

## PROJECTS

### VerifierGym

*Verifiable RL environment for training and evaluating LLM agents on Arabic web tasks*

- Built a sandboxed Arabic e-commerce site and 22 browser-agent tasks, each graded by deterministic checks on hidden app state, with a Verifiers/PRIME-RL interface for RL training.
- Proved every task solvable (scripted oracle scores 1) and non-trivial (no-op scores 0) in an 85-test suite; an adversarial review found 8 grading exploits, all fixed with regression tests.

### SecondBrain

*Self-updating knowledge engine that turns anything you read into spaced-repetition reviews and auto-grades recall*

- Built a full-stack retention engine (Next.js, FastAPI, Supabase/pgvector) that ingests articles, PDFs, and transcripts, extracts atomic concepts via LLM, auto-generates retrieval questions, and schedules reviews with an FSRS forgetting model.
- Engineered an LLM examiner that grades open-ended free-recall answers and re-fits per-user FSRS parameters from review logs, with a vector-similarity layer surfacing cross-domain links between concepts.

## LEADERSHIP & INVOLVEMENT

### OSOP Fraternity

Philadelphia, PA

*Finance Chair*

August 2023 – Present

- Managed \$150K annual budget; built automated tracking to monitor cash flow, flag delinquent accounts, and forecast spend.
- Designed end-to-end payment pipeline for dues collection, reimbursements, and vendor payouts in one dashboard.
- Delivered monthly reports to executive board with variance analysis and spend breakdowns to drive budgeting decisions.

### Wharton Undergraduate Finance & Technology

Philadelphia, PA

*External Executive*

August 2024 – Present

- Organized 10+ PM and FinTech events for 50+ members; built industry partnerships.
- Built onboarding templates and a member wiki adopted across the club.

## SKILLS, CERTIFICATIONS, LANGUAGES

**Skills:** Python, TypeScript, SQL, Java • PyTorch, XGBoost, LightGBM, scikit-learn, NumPy/Pandas • Time-series forecasting (LSTM, SARIMA, Prophet) • Snowflake, Spark, ETL • LLM APIs (OpenAI, Anthropic), RAG, agents, RL (PRIME-RL/Verifiers), evals • Next.js, React Native, FastAPI, Supabase/pgvector • AWS/GCP, Docker, Git

**Certifications:** Google Data Analytics • Technical Product Management (Cornell University)

**Languages:** English (fluent), French (fluent), Arabic (fluent), Spanish (intermediate)