

Pedro Miguel Domingos Marques

DATA SCIENTIST

Lisbon, Portugal · pedrom02.dev@gmail.com · (+351) 939 915 944 · LinkedIn · GitHub · pedromarques.dev

PROFILE

Data Scientist with end-to-end experience across the ML lifecycle: exploratory analysis, feature engineering, model selection, hyperparameter optimisation, evaluation, and production deployment. Comfortable with gradient-boosted trees, time-series forecasting, conformal prediction, and SHAP-based model interpretability. Statistical foundations complemented by software engineering rigour — production APIs, automated testing, and reproducible pipelines.

PROJECTS

Energy Forecast PT — Production ML System for Portuguese Energy Grid github.com/energy-forecast-pt ↗
Hourly electricity consumption forecasting across 5 Portuguese regions using gradient-boosted trees, with a fully reproducible end-to-end training pipeline.

- ▶ Achieved **MAPE 4.33% / RMSE 81 MW** on held-out test set across 175k+ hourly samples; **90.7% RMSE improvement** over Moving Average baseline.
- ▶ Production FastAPI service with batch predictions, SHAP explanations, conformal prediction intervals, drift detection, and model hot-reload.
- ▶ **649 tests, 85%+ coverage**; CI/CD with pytest/ruff/bandit/Trivy; multi-cloud deployment to AWS ECS, Azure Container Apps, and GCP Cloud Run.

Python · CatBoost · XGBoost · LightGBM · Optuna · FastAPI · Docker · DVC · GitHub Actions · Prometheus

Duelt — AI-Powered Household Task Management Mobile App github.com/duelt-showcase ↗
Solo full-stack mobile app helping couples manage shared household responsibilities. End-to-end ownership from architecture through backend, AI, testing, and CI/CD.

- ▶ Real-time couple sync via Firestore listeners with **offline-first mutation queue** (MMKV-backed) and conflict resolution.
- ▶ **Gemini 2.5 Flash** chat assistant with Zod-validated, template-constrained outputs; end-to-end encryption (Argon2id + AES); GDPR-compliant.
- ▶ **2,300+ tests** across 122 suites, 10 Maestro E2E flows, 5-job CI/CD pipeline with EAS Build.

React Native · Expo · TypeScript · Firebase · Gemini AI · Zustand · TanStack Query · Jest · Maestro

RushTalk — Low-Latency Voice Communication Platform github.com/Pedrom2002/rushtalk ↗
Production-grade Discord-style desktop client with a custom Rust audio engine achieving **sub-60ms mouth-to-ear latency**. Spans 12,000+ LOC across Rust, Go, TypeScript, and SQL.

- ▶ **Zero-allocation Rust audio pipeline**: noise suppression, AGC, VAD, Opus codec, EMA-adaptive jitter buffer, packet loss concealment, lock-free SPSC ring buffers.
- ▶ **Go backend with Clean Architecture**: 41 REST endpoints, WebSocket hub, 64-bit bitmask permissions, JWT RS256 auth, PostgreSQL + Redis.
- ▶ Tauri 2.0 desktop client (~50 MB footprint vs Electron's ~300 MB) with Svelte 5 frontend; Kubernetes deployment with HPA, cert-manager TLS, and Prometheus monitoring.

Rust · Go · Svelte 5 · TypeScript · Tauri 2.0 · LiveKit (WebRTC) · PostgreSQL · Redis · Docker · Kubernetes

TECHNICAL SKILLS

LANGUAGES	Python, TypeScript, JavaScript, SQL, C#, Rust, Go
ML & DATA SCIENCE	Scikit-learn, XGBoost, CatBoost, LightGBM, Optuna, Pandas, NumPy, SHAP, Conformal Prediction, Feature Engineering, Time-Series Forecasting
BACKEND & APIS	FastAPI, Flask, REST, WebSockets, PostgreSQL, Redis
FRONTEND & MOBILE	React, React Native, Next.js, Expo, Tailwind CSS
MLOPS & DEVOPS	Docker, Kubernetes, GitHub Actions, DVC, Pytest, Pre-commit hooks
CLOUD	AWS ECS, Azure Container Apps, GCP Cloud Run
TOOLS & VIZ	Git, Linux, Jupyter, Tableau, Matplotlib, Seaborn, Streamlit

EXPERIENCE

Software Development Intern Apr 2026 — Aug 2026

QUIC Portugal · Lisbon

Developing internal automation tools and custom software solutions to streamline workflows for the company's audiovisual event productions.

Store Assistant — Pingo Doce Sep — Dec 2023

Processed 200+ daily transactions; cross-departmental customer service.

Customer Service — H3 Oct 2021 — Apr 2022

Managed stock control for 150+ SKUs and onboarded new staff.

CERTIFICATIONS

- ♦ **Data Science Professional Certificate** — IBM · Coursera
- ♦ **Advanced Data Analytics Professional Certificate** — Google · Coursera
- ♦ **AI Engineering Professional Certificate** — IBM · Coursera
- ♦ **Advanced Statistics for Data Science** — Johns Hopkins · Coursera
- ♦ **Data Visualization with Tableau** — UC Davis · Coursera

EDUCATION

CTeSP in Mobile Devices Development 2024 — 2026

ISTEC Lisboa · Portugal

EQF Level 5 — Higher Professional Technical Course. Progressing to a Bachelor's degree in Data Science (post-laboral schedule, expected 2028).

BSc Clinical Physiology (attendance) — ESTeSL 2022

Languages: Portuguese (Native) · English (C1) · Spanish (B1)