Daniel Varela Perez

Mexico City, MX | bedaniele0@gmail.com | +52 55 4189 3428
GitHub: bedaniele0
LinkedIn: daniel-varela-perez

Profile

Data Scientist (Applied / Production ML) with an end-to-end focus: problem framing, training, evaluation, and deployment (API/batch). Operational monitoring of drift and performance. Emphasis on **reproducibility**, **traceability**, **observability**, and runbook-driven operations. Availability: open to remote roles.

Key Impact

- Reproducible, production-ready E2E pipelines: versioned artifacts, batch/API execution, dashboards, and operational documentation.
- Model and data monitoring: drift (PSI/KS), performance, and post-deployment analysis with operational criteria.
- Traceability and operations: MLflow **tracking**, handoffs, runbooks, and checklists; automated testing with pytest.

Top Projects (3 main + 1 operational)

Fraud Detection System (E2E)

- Fraud classification with reproducible pipeline + API (FastAPI) + dashboard (Streamlit).
- Results: ROC-AUC 95.28% | Precision 93.62% | Recall 72.13%.
- Repo: `projects/fraud-detection`

Customer Churn Prediction (E2E, API + Dashboard + MLflow)

- Reproducible pipeline with FastAPI/Streamlit and MLflow; operating threshold via cost/benefit.
- Results: AUC 83.80% | Recall 92.51% | KS 54.29%.
- Repo: `projects/churn-prediction`

Credit Risk Scoring (E2E)

- Default scoring with evaluation and decision policy.
- Results: AUC 78.13% | KS 42.51% | Recall 87.04%. Brier 0.1349.
- Repo: `projects/credit-risk`

Walmart Demand Forecasting (E2E)

- Retail forecasting with temporal backtesting + API/dashboard.

- Results: MAE 0.6845 | RMSE 3.9554. MAPE 52.75% (batch N=28,000).
- Repo: `projects/demand-forecasting`

User Score Prediction (Batch + Monitoring + Runbook)

- E2E batch pipeline with drift (PSI) and performance (MAE/R2) monitoring + operational runbook.
- Repo: `projects/user-score`

Tech Stack

Python, SQL, pandas, NumPy, scikit-learn, LightGBM.
FastAPI, Streamlit, MLflow, pytest, Git, Docker.

Experience

ML Engineer / Applied Scientist (Production ML Portfolio)

2024 - Present - Built reproducible pipelines for training, evaluation, and serving (API/batch) with experiment and artifact traceability. - Implemented drift/performance monitoring and operational documentation (runbooks) for post-deployment support. - Improved robustness with automated tests (pytest) and standardized handoffs/checklists.

Operations and Technical Consulting | MacDaniels

2019 - Present - Handled 2–3 tickets per day across Mac and Windows environments, with ~2-hour average resolution time. - Structured diagnosis and root-cause analysis; documentation and standardization of procedures. - Used AI to accelerate diagnosis and responses. - Improved client communication; high satisfaction with only 1 negative comment due to communication.

Education and Languages

- Training: continuous learning in applied ML and ML systems (evidence on GitHub).
- Spanish: Native | English: Technical (reading/writing; for calls I use AI Phone with real-time translation).

Links

- Portfolio: `README.md`
- Portfolio EN: `README_EN.md`