

# IÑAKI ROSELLO

Data Scientist | Machine Learning Engineer

irosello@outlook.com.ar | linkedin.com/in/iñakirossellosignoris | github.com/ITRoselloSignoris

## SUMMARY

Junior Data Scientist and Computer Science student (UBA) with a strong foundation in MLOps and the full Machine Learning lifecycle. Experienced in building end-to-end pipelines using Python, SQL, and deploying scalable solutions via FastAPI and Docker. Native Spanish speaker with B2 First English certification, focused on delivering data-driven insights and measurable business impact.

## TECHNICAL SKILLS

**Languages & Data Engineering:** Proficient in **Python** (Pandas, NumPy) and **SQL** (PostgreSQL, Supabase) for complex data processing.

**Machine Learning & Modeling:** Expertise in end-to-end ML using **Scikit-learn**, **XGBoost**, and **Random Forest**, including feature engineering and hyperparameter tuning.

**Specialized Analytics:** Advanced application of **Time Series** (ARIMA, Prophet), **Clustering** (HDBSCAN, K-Means), and **Recommendation Systems** (FAISS, ANNOY, Collaborative Filtering).

**MLOps & Deployment:** Proven ability to deploy models using **Docker**, **FastAPI**, and **MLflow**, with automated **CI/CD** pipelines via GitHub Actions.

**Cloud & Infrastructure:** Experience managing environments on **Microsoft Azure**, Hugging Face Spaces, and Streamlit Cloud.

**Data Visualization & Tools:** Development of interactive dashboards and reports using **Plotly**, **Streamlit**, and **Gradio**.

## PROFESSIONAL EXPERIENCE

**GEINSA: Secondary Internship in Data Science** Aug 2023 – Dec 2023

- **Engineered unified datasets** from diverse raw data sources, reducing complexity for downstream regression analysis.
- **Optimized regression modeling** by performing deep EDA and data wrangling to uncover critical data patterns and distributions.

## PROJECTS

**EquineLead: Data-Driven Growth Engine for the Horse Industry** Feb 2026 – Mar 2026

NoCountry Job Simulation

- **Cascaded Architecture:** Designed two binary XGBoost classifiers: step 1 filters non-buyers; step 2 identifies high value clients, optimizing F2-score on the highest-value minority class.
- **Feature Engineering:** Built behavioral intent features from web navigation signals and applied SMOTE to address class imbalance, aligning the optimization metric with asymmetric business costs.
- **Interactive Deploy:** Built a Gradio app on Hugging Face Spaces enabling non-technical stakeholders to test the recommendation engine and lead scoring model in real time.
- **MLOps Pipeline:** Integrated DagsHub + MLflow for experiment tracking, Evidently for data drift monitoring, and a Streamlit executive dashboard with business KPIs.

**Customer Churn Prediction and Retention for FinTech**

Sep 2025 – Oct 2025

NoCountry Job Simulation

- **Model Optimization:** Trained and optimized XGBoost and Random Forest models using SMOTE for class balancing, achieving 76% Recall and 84% ROC AUC.
- **Explainability:** Identified 5 critical churn predictors using SHAP analysis, providing global and local interpretability.
- **Deployment:** Developed a FastAPI REST endpoint with <100ms latency on Hugging Face Spaces with automated logging to PostgreSQL.
- **MLOps Pipeline:** Implemented CI/CD with GitHub Actions and daily drift detection reports using Evidently.
- **Experiment Tracking:** Managed model versioning and performance metrics across multiple iterations using MLflow.

**Product Demand Prediction**

August 2025

AlixPartners ECI Competition

- **Forecasting:** Developed and compared **ARIMA** and **Prophet** models to extract insights from noisy demand patterns.
- **Elasticity Modeling:** Built a price optimization model using **log-log regressions** to simulate profit-maximizing price points based on demand elasticity.
- **Strategic Reporting:** Authored a technical report with data-backed sales strategies for a retail chain.

**Movie Recommendation Model**

Feb 2025 – Mar 2025

Personal Project

- **Hybrid Filtering:** Implemented Collaborative and Content-Based Filtering using the **Surprise** library and **SVD/PCA** for dimensionality reduction.
- **Efficient Search:** Integrated **FAISS** and **ANNOY** indexes to enable high-speed similarity search for real-time recommendations.
- **End-To-End ML:** Implemented a full-stack machine learning pipeline, deploying a FastAPI service with Docker on Microsoft Azure and a Gradio interface, despite being the most junior member.

**Fraud Detection and Prevention**

October 2024

Final Project for Edvai’s Data Science & MLOps Bootcamp

- **Detection System:** Engineered a **Random Forest** classifier and two clustering models (HDBSCAN, K-Means) to identify fraudulent behavior patterns.
- **Full-Stack ML:** Deployed a FastAPI service via Docker on Microsoft Azure, featuring an interactive Gradio web interface.

**EDUCATION**

|                                             |                     |
|---------------------------------------------|---------------------|
| <b>Degree in Computer Science</b>           | Aug 2024 - Present  |
| Universidad de Buenos Aires (UBA)           |                     |
| <b>Secondary Technical School</b>           | Mar 2017 – Dec 2024 |
| <i>Electronic Technician Specialization</i> |                     |
| Escuela Técnica Roberto Rocca (ETRR)        |                     |

**CERTIFICATIONS**

- **Languages:** Spanish (Native), English (*B2 First - Cambridge*).
- **Data Science & MLOps:** Bootcamp Data Science & MLOps, Deep Learning 360 (Levels 1 & 2), Statistics for ML – *Edvai*
- **Programming:** PCAP: Programming Essentials in Python – *Cisco Networking Academy*
- **Other:** Intro to Deep Learning – *Kaggle*; Introduction to SQL – *Edvai*