

JUAN HERNANDEZ VARGAS

Data Scientist | ML & Predictive Modeling Specialist | Fintech, Insurance & Biotech

+57 315 457 0159 | juandanihv@gmail.com | Bogotá, Colombia | <https://www.linkedin.com/in/juan-hernandez-vargas/>

GitHub: <https://github.com/Juanhv24>

PROFESSIONAL SUMMARY

Results-driven Data Scientist with a strong statistical background and a professional certification in Data Science. Expert in building end-to-end Machine Learning solutions (LightGBM, XGBoost, Random Forest) and automating complex processes using Python and SQL. Specialized in risk assessment, user behavior prediction, and biotech data analysis, with a proven track record of deploying interactive tools that translate data into measurable business strategies.

TECHNICAL SKILLS

- **Data Science & ML:** Gradient Boosting (LightGBM, XGBoost), Random Forest, Scikit-learn, Classification/Regression Models, Feature Engineering, Hypothesis Testing.
 - **Analysis & Processing:** Python (Pandas, NumPy), SQL (PostgreSQL), Data Cleaning, and ETL.
 - **Advanced Statistics:** Bootstrapping, Operational Risk Analysis, Probability of Loss, and Research Methodology.
 - **Visualization & Deployment:** Streamlit (Dashboard Creation), Plotly (Interactive Charts), Power BI (Functional Level), Matplotlib, and Seaborn.
 - **Version control:** Git, GitHub
-

KEY PROJECTS

Interconnect — Client Churn Prediction

- Developed an end-to-end solution to identify users at risk of cancellation, optimizing retention strategies.
- Trained and compared LightGBM, XGBoost, and Random Forest models, selecting LightGBM for its superior efficiency.
- Achieved an AUC-ROC of 0.9054 and a global accuracy of ~85% on the test set.
- Deployment: Created an interactive Streamlit dashboard featuring a real-time risk simulator and technical KPI visualizations (SHAP, Confusion Matrix).
- Insights: Identified that tenure under 2 months and monthly contracts are the primary risk factors.

OilyGiant — Operational Risk and Profitability Modeling

- Conducted an exhaustive risk assessment using linear regression to identify the most profitable locations for new business development.
- Implemented Bootstrapping techniques to simulate 1,000 scenarios, calculating 95% confidence intervals for profits and quantifying the probability of loss.
- Provided data-driven recommendations to maximize profit margins while maintaining operational risk below the 2.5% threshold.
- Impact: Validated decision-making processes through the statistical quantification of risk.

Beta Bank — Customer Churn and Behavior Analysis

- Developed a Random Forest machine learning model to predict customer churn, analyzing historical data from 10,000 bank clients.
- Addressed data imbalance using oversampling techniques, enhancing the model's ability to detect customer exit patterns.
- Achieved an F1 score of 0.61 and a ROC-AUC of 0.86, demonstrating high precision in distinguishing loyal customers from those at risk of leaving.

EDUCATION

- **Specialization in Statistical Analytics for Data Science** | Universidad De La Salle | In Progress (2026).
 - **Data Science Career Track** | TripleTen Bootcamp | Graduated (2026).
 - **B.S. in Biology** | Universidad de La Salle | Graduated (2023).
-

PROFESSIONAL EXPERIENCE

Operations Customer Expert | Teleperformance

2023-Present

- **Data-Driven Performance Management:** Lead personalized coaching sessions based on KPI analysis (CSAT, QA, AHT) to identify and bridge performance gaps.
- **Monitoring and Documentation:** Manage the documentation flow of sessions and deliverables per advisor, ensuring individual progress traceability and performance database integrity.
- **Root Cause Analysis:** Perform deep-dive analyses on metric deviations to propose action plans that optimize operational efficiency and customer satisfaction.

Museum Management and Planning Intern | Universidad de La Salle

2022-2023

- **Logistics & Tour Planning:** Orchestrated and planned educational tours, managing schedules and coordinating visitor logistics to ensure operational efficiency.
 - **Data Support:** Utilized Excel to maintain visitor logs and collection records, ensuring organized documentation of museum activities.
 - **Communication:** Served as a museum guide, translating complex scientific or historical information into accessible narratives for diverse audiences.
-

LANGUAGES

- **Spanish:** Native.
- **English:** B2 (Professional Working Proficiency).