EDWIN A. ESPINOZA

San Diego, CA (open to relocation) | (619) 953-7986 | e6espinoza@ucsd.edu
linkedin.com/in/aldair-ai | github.com/aldair-ai

Data Scientist with 5 years of experience turning ambiguous product and business questions into models, experiments and metrics across marketplace, retail and financial services, plus 2 years leading data science teams. M.S. Data Science candidate at UC San Diego (graduating March 2027), with research on contextual-bandit recommenders and author of **abaudit**, an open-source library for auditing A/B test validity. Owns problems end to end, framing them with stakeholders, defining evaluation metrics, extracting and validating data in SQL and Python, and shipping models to production.

EDUCATION

University of California, San Diego — *M.S. Data Science — Expected Mar 2027*

Massachusetts Institute of Technology (MITx) — *MicroMasters, Statistics and Data Science*

National University of Engineering, Peru — *B.S. Computer Science*

EXPERIENCE

Credit Bank of Peru (BCP) — *Lima, Peru | Jan 2023 – Sep 2025*
Largest bank in Peru; Credicorp group ($13.6B market cap).

Lead Data Scientist, Audit Analytics — *Nov 2023 – Sep 2025*

- **Led a 4-person data science team** delivering GenAI systems for audit: **LLM + RAG** document verification and anomaly detection over phone sales transcripts (Hugging Face Transformers, LangGraph agents, vector databases) **~ $250K in cost avoidance**
- Built a computer-vision compliance monitor (**YOLO, OpenCV**) replacing limited manual sampling across **320+ branches** with full-population coverage, saving **40,000+ audit hours** and **~$1M annually**

Senior Data Scientist — *Jan 2023 – Nov 2023*

- Developed an automated contract-validation solution (**Azure SDK, OpenCV, RegEx**) to extract and verify contractual clauses, increasing review capacity **100×**, reducing non-compliance risk exposure, and avoiding **~ $180K in annual penalties**
- Built an AI-powered document-processing pipeline (**Azure, OpenAI Whisper, Hugging Face NER**) to extract and validate sales guidelines and commercial conditions, increasing processing capacity by **~ 75×** and generating **$150K in annual cost savings**

AB InBev — Senior Data Scientist, Marketplace — *Lima, Peru | Jan 2022 – Jan 2023*
World's largest brewer ($117.7B market cap); B2B digital marketplace.

- Productionized an **XGBoost** targeting and scheduling model (**Python, Spark on AWS** — S3, EMR, Glue) for in-app and push notifications, lifting **click-through rate from 7% to 10%** and generating **+$210K in annual revenue**
- Designed and ran **A/B tests** on geo-targeted summer campaigns — identified high-propensity customers with a **K-Nearest Neighbors** model on geolocation purchase patterns and measured lift experimentally, delivering **~$100K/month** in incremental sales.
- Forecast weekly demand (**ARIMA**) for **30% of marketplace sales** to drive inventory planning and in-app product recommendations, reducing stockouts and raising conversion from **2% to 4%** for **$130K in annual revenue**

Pichincha Bank — Data Scientist, Risk — *Lima, Peru | Jan 2021 – Jan 2022*

- Built a **Random Forest** credit-risk model (**Python, SQL feature engineering, ensemble learning**) over large loan portfolios, reducing credit defaults by **5%** and recovering approximately **$750K in annual loan repayments**
- Developed an **ANN-based collections prioritization model** (**Keras, Databricks**) to identify high-risk overdue accounts and optimize recovery strategies, increasing collection rates by **4%** and driving **~ $1.2M in additional annual recoveries**

SELECTED PROJECTS & RESEARCH

abaudit — Open-Source Python Library for A/B Test Validity Auditing — *2026*
github.com/aldair-ai/abaudit | published on PyPI (pip install abaudit) | Python 3.9–3.12, 184 tests, 99% coverage.

- Built a library that answers the question a p-value cannot: **given a significant A/B result, how likely is the effect real?** Computes **Positive Predictive Value** via the Ioannidis Bayesian framework and returns a composite 0–1 bias score with self-contained HTML reports for stakeholders.
- Implements **post-experiment audits** (Bonferroni correction, statistical power, **Sample Ratio Mismatch**, quantified **optional-stopping/peeking bias**), **sample-size planning** that jointly optimizes power and trustworthy PPV, and runtime health checks for live experiments.

Comparative Evaluation of Contextual Bandit Algorithms for Movie Recommendations — *2026*
UC San Diego, Halıcıoğlu Data Science Institute (DSC 291, Prof. Yian Ma).

- Benchmarked **7 contextual bandit algorithms** (LinUCB, LogisticUCB, Thompson Sampling, Bootstrap TS, NeuralUCB, NeuralTS, ε-Greedy) on **MovieLens 10M** under offline replay — **1.16M interactions**, 24-dim context features, chronological train/validation/test split, and 25-item candidate sets mimicking a production **two-stage retrieval-then-ranking** pipeline.
- Diagnosed a structural confound between match rate and cumulative regret in offline replay and introduced **regret-per-matched-event** as a normalized comparison metric; LinUCB and NeuralUCB proved most efficient per interaction (regret per matched event of 0.35 and 0.36) while NeuralUCB achieved the highest reward rate (**0.683 vs. 0.626 base rate**).

SKILLS

Statistics & Experimentation: A/B testing, experiment design, statistical power, multiple-testing correction, sample ratio mismatch, causal inference, metric definition, forecasting, contextual bandits

Programming: Python, SQL, R, Scala | Visualization**:** Tableau, Power BI, Matplotlib

ML & AI: scikit-learn, XGBoost, LightGBM, PyTorch, TensorFlow, Keras, Transformers, LLMs, RAG, LangChain/LangGraph, recommender systems, ranking & retrieval, computer vision (YOLO, OpenCV)

Data & Infrastructure: BigQuery, Spark/PySpark, Databricks, Airflow, Dataflow, MySQL, Oracle, MongoDB, GCP, AWS, Azure, Git

Languages: English, Spanish, Portuguese