Alejandro Rodas

Berkeley, CA ● (510) 973-9134 ● alex9102017@berkeley.edu ●
Linkedin: https://www.linkedin.com/in/alejandro-rodas-/

EDUCATION

University of California, Berkeley, CA (CDSS) **Aug 2021-May 2025**

B.A. in Data Science with Domain Emphasis in Industrial Analytics

RELEVANT COURSEWORK: Structure and Interpretation of Computer Programs (**PROGRAM DESIGN, PYTHON**); Data Structures and Algorithms (**DATA STRUCTURES/GITHUB/GRAPHS**); Foundations of Data Science (**EDA, PANDAS, NUMPY**); Principles and Techniques of Data Science (**PIPELINES, SQL**); Probability for Data Science (**STATISTICAL MODELING**); Data Engineering (**ETL, MONGODB, OLAP/DATABASES**); Data Mining and Analytics (**ML ALGS**), Data Inference and Decisions (**ADVANCED STATISTICAL MODELING - GLMS, NONPARAM MODELS**)

SKILLS: Python, SQL, TensorFlow, PyTorch, Scikit-learn, Regression Analysis, A/B Testing, Statistical Inference, Bayesian Inference, Gibbs Sampling, NLP, Optimization Modeling, GLMs, AWS, GCP, Databricks, PySpark, Snowflake, MongoDB, Matplotlib, Seaborn, NumPy, GeoPandas, Sigma, ETL pipelines, Fluency in model building: RF's (Random Forests) NN, (Neural Networks)

EXPERIENCE

CDSS Discovery Research Program **Jan 2024 - May 2024**

Lead Researcher

- Led a team analyzing federal financial datasets, structuring unorganized records from 350+ centers all over the country.
- Designed and implemented A/B tests and trained supervised models (logistic regression, RF's) to evaluate performance.
- Applied model selection techniques, hyperparameter tuning, and evaluation metrics(AUC, precision/recall) to assess impact.
- Presented findings to policy and research teams unfamiliar with machine learning and data science terms, guiding decision-making through clear visuals. Relying on libraries like matplotlib seaborn and platforms like Sigma.

U.S. Health Department **May 2024 - Aug 2024**

Data Science Intern

- Built machine learning models (logistic regression, decision trees) to predict post-pandemic changes in service usage
- Conducted statistical analysis comparing pre- and post- pandemic data using hypothesis testing. Tuned models using cross-validation, grid search, and performance metrics (AUC, F1, precision/recall).
- Conducted sentiment analysis on user reviews using NLP tools (e.g., VADER, TextBlob) to quantify perception of centers.
- Created dashboards in BigQuery & Tableau to present risk insights to non-technical decision-makers. Collaborated with executive directors to release a federal report coming in late 2025.

UC Berkeley's Financial Aid & Scholarships **Mar 2022 - May 2025**

Data Science Intern

- Conducted statistical analysis using hypothesis testing, bootstrapping, and missing data imputation to identify trends and disparities in multi-year financial aid data.
- Developed and deployed ML models (random forests, decision trees) to predict student aid utilization and flag risk cases.
- Built interactive dashboards in AWSto communicate insights and support data-driven decision-making for financial officers.

HIGHLIGHTED PROJECTS

Research Project: Causal Inference, ML for U.S. Primary Elections **Mar 2025 - May 2025**

- Applied sentiment analysis to endorsement texts using NLP to extract polarity scores, used as features in supervised models.
- Built and evaluated logistic regression and Random Forest models to predict election outcomes; performed EDA on 2,600+ candidates across race, gender, and incumbency. Presented findings in a research paper.

Predicting Streaming Service Churn with Machine Learning **Oct 2024 - Nov 2024**

- Built machine learning models (NN's, Random Forests, Decision Trees) to predict churn rates with 85% accuracy.
- Used logistic regression to model churn probability; evaluated ROC curves and confusion matrices.

LEADERSHIP

Latin American Leadership Society at Berkeley **Jan 2022 - Jun 2022**

Forum Committee Vice President

- Organized events for 100+ attendees using data-driven tools to optimize scheduling and participation. Streamlined event logistics through technology platforms, enhancing communication while addressing complex organizational challenges.