

Carolina Arriaga

Data Scientist | Vallejo, CA | caro.arriaga@gmail.com | caroarriaga.com

- Serial hackathon winner with experience in data analysis, modeling, and entrepreneurship.
- Recent graduate of Master's in Information and Data Science at UC Berkeley.
- Looking to join a mission-driven team with focus on environmental policies and sustainability.

Experience

Jan 2019 – Mar 2024

Founder, Data Scientist at Progressive Emergence

- Lead author of a working paper on electrification of an early-adopter fleet in Canada.
- Co-authored a paper for IEEE EDUCON 2020 on evaluating quality decision-making and applied sentiment analysis.

Apr 2023 – Dec 2023

Co-Founder, Data Scientist at [Facets](#)

- Implemented data analytics to web, and react native apps using Segment and Mixpanel.
- Conducted product fit surveys, focus groups, and data analysis, leading to major pivot decisions.
- Obtained 11k in funding through agile implementation in hackathons.

Jun 2022 – Aug 2022

Data Scientist, Zero Emission Vehicles Intern at [The ICCT](#)

- Modeled a fuel vs electricity cost savings calculator aimed for fleet managers.
- Deployed a web app with interactive visualizations to help fleet managers assess EV savings.

Education

Aug 2020 – Aug 2022

Master in Information and Data Science from University of California, Berkeley

- Leopard Spotting: A Deep Learning model to re-id leopards. Used roboflow, YOLOv5, combined with ResNet-18 CNN embedding model for re-id. We achieved a top-1 0.61 accuracy, on par with SOTA on other species.
- ShapSum: A framework to predict human judgement multi-dimensional quality scores for text summarization: Used a random forests model coefficients and shapley value analysis to understand

what NLP metrics evaluate a summary so its on par with human judgement.

- A study on people's memory after exposed to a controlled tweet: we found that people tend to remember false information with positive sentiment.
- Effectiveness of government interventions and Covid-19 deaths: Used a linear regression model, we found that in early pandemic, states with interventions saved more lives than states that didn't had them in effect.

Aug 2006 – May 2011

B.Sc. in Engineering Physics from Tec de Monterrey

Skills

Python R SQL HTML/CSS Pyspark Pytorch Scikit-learn Google cloud Databricks
ML modeling Statistics

Awards

Dec 2023

\$1k USD Winner - Honorable mention from DataPalooza

Nov 2023

\$10k USD Winner - 1st Place from AI and Identity - Okta Hackathon

Oct 2023

Winner - TinyMCE Challenge and Clarifai Challenge from API World + AI DevWorld Hackathon

Publications

Aug 2023

Case study: Electrification of an early-adopter fleet in Canada. by ICCT

Feb 2020

Teaching decision making in engineering using a multi-attribute approach and ship design by IEEE

Jul 2019

Hybrid twist tray ice maker by USPTO