

Mikiko Bazeley

SF Bay Area, CA | mmbazel@gmail.com | [linkedin.com/in/mikikobazeley](https://www.linkedin.com/in/mikikobazeley) | github.com/MMBazel | medium.com/@mmbazel

SUMMARY

Software Engineer with 7+ years shipping ML models & data-driven solutions for high-growth companies in a variety of industries (solar, construction, recruitment-tech, & health-tech). Focused on full-cycle product engineering, including data engineering, ML model development, deployment, serving, monitoring & evaluation. Speaker, consultant, & mentor for initiatives & organizations focused on increasing representation in data science & machine learning.

EXPERIENCE

FEATUREFORM – San Francisco, CA

MLOps Engineer - DevRel (Content + Open-Source)

Sep 2022 – Present

Summary: Building out MLOps Content & OSS at Featureform, a virtual feature store focused on empowering data science teams.

INTUIT – San Francisco, CA

Senior SWE, MLOps

Nov 2021 – Sep 2022

SWE, MLOps (Mailchimp)

May 2021 – Nov 2021

Summary: SWE focused on developing MLOps platform to efficiently scale model pipeline training & deployment, integration, testing & monitoring for the Data Science team.

- Major contributor to migration of Data Science project templating CLI tool & local model development to pure cloud model training & deployment with VertexAI, Github, and Jenkins, helping cut project & dev env setup time from weeks to minutes
- Selected for the first Global Mailchimp Engineering hiring committee as ML Ops & Data Engineering sole representative; created interviewing process & collaborated on developing system design interviews for ML Ops team (closed offers for 4 senior MLOps Engineers & 5 senior Data Engineers, out of 50+ candidates)
- Lead change management of major upstream templating command-line tool including updates to 50+ downstream repos, as well as managing merge conflicts, performing code reviews, & developing initial repo health checklist; Also shipped improvements to DS project templating tools including bug fixes & patches
- Won Mailchimp GPT-3 Hackathon & pitched idea to product; accepted as new product feature (generating persona based copy for holiday newsletters)
- Developed "Software Engineering Skills 101" learning path for Data Science team, including modules and workshops on testing, version control, & Python language fundamentals

SIDEWALK.AI – San Francisco, CA

Machine Learning Engineer

May 2019 – Jan 2021

Summary: First MLE at early-stage start-up focused on building AI-first enablement tools for residential real-estate.

- Developed ML roadmap, research & implemented proof-of-concept models from papers (in listing recommendations, user profile clustering) utilizing Python (Scikit-learn, Surprise, Dask, GeoPandas), AWS RDS, Jupyter
- Developed early iteration of ETL pipelines using Python (including SQLAlchemy, paramiko, Dask) with logging, data cataloguing, & job status tracking for processing 1000+ text files (up to 8MB in size)
- Designed data models & schema, created DDL & DML scripts for Postgres (then Amazon RDS)
- Researched, documented, & acquired external datasets through scraping (Python Requests, BeautifulSoup, Scrapy) as well as through APIs (Quandl, Kaggle, etc)

TELADOC – Mountain View, CA

Data Scientist II, Growth Marketing

Oct 2019 – Nov 2020

Summary: Collaborate closely with senior management to increase enrollment, activation & retention for members living with chronic & behavioral conditions (Diabetes, Hypertension, Pre-Diabetes, Weight Management).

- Provided experimentation support: Designed, set-up, & analyzed experiment results, including A/B tests (Python, Seaborn, SciPy, statsmodels, Metabase, SQLAlchemy)
- Lead projects in developing, documenting, scripting, & visualizing key metrics & analyses around marketing & growth
- Developed machine learning models (classification, regression, time series) & pipelines using Python (Pandas, Scikit-learn, Prophet, statsmodels), Redshift, S3. Leverage Metabase, Tableau & TabPy to engage stakeholders with analyses & machine learning model outputs around forecasting & prediction

WALKME – San Francisco, CA

Sr BI Analyst, Sales & Revenue Ops

Aug 2018 – Oct 2019

Summary: Responsible for leading & executing advanced analytics & data science initiatives for the revenue operations organization. Worked with key stakeholders to scope, develop, & present ML models.

- Developed sales & marketing pipeline demand model to set sales targets, evaluate pipeline funding options, & monitor key metrics (lead generation, lead coverage, opportunity conversion). Leveraged Python, Google sheets, Redshift, Tableau. Made strategic recommendations using model results to prevent potential risk of \$4M+ global revenue target miss.
- Developed pipeline allocation model & pipeline (with Python, Numpy) allocating lead generation targets from Demand Model to teams & individuals based on quota, start date & quota ramp. Model output ingested into external BI tool & CRM (Clari, Tableau) for real-time tracking of sales & actionable insight into opportunity velocity, leading to improvements in sales forecasting.
- Created propensity-to-purchase model using Salesforce data (opportunity, lead, & call records) as well as Python (simple-salesforce, NLTK, Spacy, Scikit-learn). Using features created around lead job titles, prospect account pain points, & historic demo call data, developed random forest model to predict propensity-to-first-purchase of account. Provided model interpretability using pdpbox, SHAP, & Scikit-learn.

VARIOUS COMPANIES – SF Bay Area, CA

Dec 2013 – Jul 2017

Autodesk (Data Scientist/Analyst), Sunrun (Analyst), Digimarc Guardian (BizOps), RecruitLoop (Growth Hacker), eSalon (Biz Ops)

Summary: Held various roles for a number of early- to mid-stage start-ups in bizops growth, and customer service.

MEDIA KIT & PLATFORMS

PODCAST APPEARANCES – [SuperDataScience](#), Ken's Nearest Neighbors ([Part 1](#), [Part 2](#)), [So You Want To Be a Data Scientist](#), [Artists of Data Science](#), etc.

PANELS + Q&A – [LesbiansWhoTech \(Not\) IRL](#), [General Assembly's Women in Data Science](#), [CareerKarma AMA](#)

SOCIAL MEDIA & Content – Platforms: [Youtube](#), [Substack](#), [Twitter](#), [Twitch](#) w/ clients: [Nvidia](#), [Bright Data Solutions](#) ([State of Data Science](#)), [LEARN Media](#)

EDUCATION

UNIVERSITY OF CALIFORNIA, SAN DIEGO – La Jolla, CA

Economics & Anthropology, B.A.

2013

SKILLS

Languages: Python [*PyTest, Great Expectations, scikit-learn, pandas, numpy, genism, spacy, OpenCV*], SQL [*Postgres, MySQL, Oracle*], C++, Solidity, Rust

Applications: Docker, Kubernetes, Airflow, GCP [*BigQuery, AI Platform, Dataflow*], Streamlit, Flask, Databricks/PySpark, FastAPI, Jenkins, Github Actions

Certificates: Full Stack Deep Learning, Deep Learning Specialization (Coursera – *deeplearning.ai*), Supply Chain Management (Micromasters – MITx), Database Programming: MySQL (CCSF), Data Science Bootcamp (Springboard), Machine Learning Engineering (Springboard)