

SKILLS

Languages & Software: Python (Pandas, Scikit-learn, PyTorch, Plotly), AWS (CLI, SageMaker), GCP (BigQuery, AI Platform), MLFlow, Databricks, OpenAI (Chat-GPT), R, MongoDB, SQL, Flask, Apache Kafka, Spark, Hadoop, Tableau, Looker Studio

Statistics and Machine Learning: NLP/LLMs, Classification & Clustering, Regressions, Neural Networks, Deep Learning, Hypothesis Testing, Optimization, A/B Testing, Probability & Causal Inference, Field Experiments

WORK EXPERIENCE

FiscalNote, Senior Data Scientist - NLP

09/2022 – Present

- Owner and responsible for the continuous integration/continuous delivery (CI/CD) of AI models in the Curate platform
- Tested, fine-tuned and integrated LLMs (BERT, GPT) into backend to deploy capitalizable client-facing features
- Leveraged data engineering tools (Docker, Mongo, MLFlow) to deploy topic models, sentiment scoring features, document classifying features, golden data auto-labelers, etc; in close collaboration with non-technical coworkers

Accenture, Data & AI

05/2021 – 09/2022

Meta, Data Scientist

- Responsible for model development and experimentation for marketing campaign measurement and optimization
- Collaborated with Meta research teams to understand the target audiences for lookalike and causal demand models

Google, Data Science Practitioner

- Developed forecasting models to predict product performance for Google Maps marketing metrics. Reduced previous project lead time by half while improving accuracy by 40% with feature engineering
- Built prediction models using MLP regressor, random forest regressor, FB Prophet models at over 90% R²
- Constructed an optimization model using a fitted cost function to allocate marketing budget across countries with respect to given constraints. Assisted with building the model-interface pipeline in GCP

Google, Analytics Consultant

- Employed data analytics solutions to improve the efficiency of Trust & Safety program training using Google PLX, Data Studio and SQL in an agile process
- Ensured validity of reported metrics for consumer-facing campaigns of 1M+ users by running data validation, auditing and hypothesis tests for Grow with Google

Happiness Alliance, Data Scientist

- Leverage regressions to help the non-profit Happiness Alliance publish economic insights focused on happiness

Self-Employed, Data Science Consultant

06/2020 – 05/2021

- Deployed a logistic regression in AWS to predict whether voltage drops originated onsite or from the power grid for Powerside Inc. Creative feature engineering led to 98% accuracy and a reduction of \$2,500/month in computing costs
- Increased donations and operation efficiency for Nehemiah Gateway USA by working on a marketing and fundraising analytics dossier including NLP text analysis on newsletters, social media visualizations and A/B testing for email blasts

ViviendasLeon, Data Scientist

01/2020 – 01/2021

- Improved farming outcomes by building a classifier to determine best combinations of pesticides and crops
- Maximized farming data insights by creating Tableau dashboards and teaching leadership/volunteers how to use the dashboards to answer ground operation questions

PROJECTS

U.S. Department of Defense: GameChanger Platform

- Collected requirements from contractors and detected topics for DoD policy documents in a semi-supervised setting
- Built topic models using embeddings trained on a set of different architectures (Word2Vec/BERT) and clustered topics using kMeans. Evaluated results using DoD Issuance Numbering Systems, as well as with qualitative analysis

Field Experiment on Children's Book "Ennie & Entropy"

- Designed and ran a multifactor field experiment on Facebook Ads to causally infer whether gender representation in online advertisements increases the click-through rate for my self-published STEM children's book "Ennie & Entropy"
- Prefaced the experiment with a power analysis and modeled product performance results with regressions in R

EDUCATION

MS: Master of Information and Data Science. **University of California, Berkeley**, CA (GPA 3.9)

Aug 2022

BA: Data Science. Minor: Business/Industrial Analytics. **University of California, Berkeley**, CA (GPA 3.9)

May 2021

AS: Engineering Science and Mathematics. **Bergen Community College**, Paramus, NJ (GPA 4.0)

May 2019