

# SHARANYA SHARMA she/her/hers

DATA ENGINEER

SF/Goleta, CA | (925) 699-2120 | sharanyasharma21@gmail.com

sharanyasharma.com | linkedin.com/in/sharanyasharma21 | github.com/sharanyasharma21

## TECHNICAL SKILLS

---

### Languages & Tools:

Python, SQL, Tableau, DBT, Dagster, AWS (S3), JavaScript, Spark, Hadoop, Kafka, Google Cloud, Docker, Kubernetes

### Data & ML Expertise:

Regression, Machine Learning, Data Structures & Algorithms, Time Series Analysis, Regularization & Optimization, Bayesian Modeling, Stochastic Processes, ETL Pipeline Design, Generative AI/LLM, Big Data, A/B Testing, MCP

## EXPERIENCE

---

### Scalable Commerce, Santa Barbara - CA

Data Analyst / Engineer

July 2024 - present

- Created data pipelines using **DBT**, **Dagster**, **Amazon Redshift**, and **Microsoft Azure**, as well as **AWS S3** orchestration, ensuring reliable ETL workflows and optimized query performance for high-volume datasets
- Integrated ML through retention prediction models, **Model Context Protocol (MCP)**, and the **DBT Semantic Layer**, to automate complex data querying and enable self-serve analytics, reducing manual SQL workloads and improving data accessibility for non-technical users.
- Leveraged **SQL**, **Python**, and **Google Apps Script** to build automated analytics for customer retention and acquisition, enabling stakeholders to make data-driven decisions on product improvements and features
- Built interactive dashboards in **Tableau** and **Google Analytics** to surface **A/B testing metrics**, delivering insights to stakeholders on customer behavior, funnel performance, and ROI

### Stanford Linear Accelerator Laboratory (SLAC), Palo Alto - CA

Machine Learning Engineer

June 2024 - December 2024

- Developed CNN, ResNet, and Vision Transformer models to detect Bragg peaks in X-ray diffraction images, and determine protein crystal orientation through prediction of detector and beam tilts
- Applied computational physics, Fourier optics, and diffraction analysis to reconstruct electron density maps for protein crystallography, for potential drug discovery research

### Appfolio, Santa Barbara - CA

Generative AI Engineer, Capstone at UCSB

Jan 2024 - May 2024

- Developed an LLM & UI to revolutionize data retrieval of numerous accounting documents for property managers
- Implemented Langchain and a code interpreter with few-shot prompt engineering, previous accuracy 84%
- Improved accuracy to **92%** using RAG, storing data in vector database with semantic similarity search to query
- Formulated evaluation model using fuzzy match algorithm to test accuracy. Working on reinforcement learning

### BayOne Solutions, Pleasanton - CA

Software Engineering Intern

June 2022 - August 2022

- Collaborated with R&D team by using MongoDB from MERN stack to optimize storage allocation of user data
- Enhanced the front end view of product *techscore* version 1 by redesigning the user interface using React/NodeJs

## TECHNICAL PROJECTS

---

**Google Developers Solution Challenge - 2<sup>nd</sup> place at UCSB:** Coded an app from scratch facilitating electric vehicle owners to rent out their private chargers for those in long-distance travel, where access to public charge points is limited.

**Spotify Music Recommender System:** Used low-rank matrix factorization with L2 regularization and k-means clustering techniques to develop an optimization algorithm for recommending songs with 85% accuracy.

**Election Prediction using Deep Neural Networks:** Created a DNN to predict active voter participation rate in the US and election outcome from the 2020 election, 85% accuracy.

**Stroke Diagnosis Probability Forecast:** Prepared data through Extract-Transform-Load (ETL). Employed k-NN, Ridge, and Lasso Regression models to compare prediction accuracies. Investigated Random Forests, XGBoost, and LightGBM for further comparison.

## EDUCATION & CERTIFICATIONS

---

University of California, Santa Barbara - Bachelor of Science, Statistics & Data Science

Technology Management Program, UCSB

AWS Cloud Practitioner & SAS Certified (in progress)

Google - Deep Learning with TensorFlow and Pytorch

Data Engineering Bootcamp - DataExpert.io by Zach Wilson