ADIT SHRIMAL

adit.p.shrimal@gmail.com | +1 (408) 203 2174 | San Francisco, CA, USA | www.aditshrimal.com | LinkedIn | Github

PROFESSIONAL EXPERIENCE

Amazon Web Services **Nov 2022 - Present**

Machine Learning Intern

- Developed a sequential recommendation model (**GRU4Rec**) using the Rechorus library.
- Utilized **SHAP** and **KernelSHAP** to explain model predictions and identify features with the highest impact on recommendations, improving **model interpretability**.
- Implemented **TimeSHAP** for time-based explanations of recommendations, further improving user engagement.

Fractal Analytics **Nov 2021 - Jun 2022**

Data Engineer

- Implemented **data-driven solutions** for enterprise data warehouses using **GCP** products (Composer, Dataflow, BigQuery, Cloud Storage), improving data processing efficiency by **30%**.
- Developed an end-to-end **data encryption** strategy with Google Tink and KMS, reducing security incidents by **80%**.
- Recognized with the '**Star Award 2022**' for exceptional problem-solving and technical skills.

Bewakoof Brands Pvt. Ltd. **Feb 2020 - Nov 2021**

Data Scientist

- Developed a product ranking auto-sort algorithm, boosting click-through rate by **26%** and add-to-cart rate by **12%**.
- Implemented personalized recommendations using ALS, which boosted revenue-per-thousand-impressions by **14%**
- Built a cost-efficient GCP data pipeline (**40%** cost reduction) with improved scalability and fault tolerance.
- Developed a system to send emails in bulk using Amazon SES with a **10x return-on-ads-spend** (ROAS) leveraging RFM segmentation for user base selection.
- Won '**Employee of the Month**' for October '20 and '**Employee of the Quarter**' for Q4 FY 20-21.

Reliance Jio **Jul 2018 - Feb 2020**

Data Engineer

- Constructed a **streaming** data pipeline using Kafka, Flink & Elasticsearch, handling millions of events in real-time and outperforming existing pipelines with a lag of over two hours.
- Implemented a data warehouse with HDFS, Spark, and Cassandra, enabling systematic computational analysis for Product Managers and Leaders.
- Automated **80%** data compression in Hadoop Cluster using Airflow, Spark, and MySQL, optimizing storage capacity to remain below **20%**.

ACADEMIC PROJECTS

Product Search Engine (GitHub Repo)

- Built a search engine pipeline using PySpark, MongoDB, and Airflow, fetching product data from the ASOS API, storing the data in Google Cloud Storage (GCS), and loading it into MongoDB.
- Implemented **BM25** using an inverted index for efficient, **constant-time** retrieval of relevant search results.

Entrepreneurship in Data Science - OnlyStats

- As part of a data science entrepreneurship course, collaborated with a team of five to conceptualize, design, and build an MVP for a sports analytics company. Secured **first place** in the course competition.
- Developed and optimized the technology infrastructure, implementing a Flask API and deploying the platform on Google Cloud App Engine for scalability, reliability, and security.
- Contributed to the front-end development of the platform, working on key React components.

EDUCATION

M.S. in Data Science, University of San Francisco **Jul 2022 - Present**

Relevant Coursework: Machine Learning, Linear Regression, A/B Testing, Distributed Computing (Spark), Relational Databases (SQL), NoSQL, MLOps, Data Acquisition

B.E. in Computer Engineering, University of Mumbai **Jul 2015 - May 2018**

Relevant Coursework: Data Structures & Algorithms, Computer Networks, Machine Learning

TECHNICAL SKILLS

- Programming Languages: Python, Java
- Databases: MySQL, MongoDB, Elasticsearch
- Big Data & Cloud: Kafka, Spark, Flink, HDFS, Airflow, Google Cloud Platform - Pub/Sub, BigQuery, Dataflow, Key Management Service, Cloud Functions, App Engine
- Machine Learning: Linear Regression, Logistic Regression, Decision Trees, Random Forest, Kmeans, Gradient Boosting Machines (XGBoost), SVD, PCA, Collaborative Filtering (ALS), Feature Engineering, and Model Evaluation
- MLOps: MLFlow, Data Version Control (DVC), CI/CD, Docker, Kubernetes, API design (Flask)