ADIT SHRIMAL

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

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

PROFESSIONAL EXPERIENCE

Amazon Web Services   **Nov 2022 - Present**

Machine Learning Intern

- Built a sequential recommendation model (**GRU4Rec**) using the **Rechorus** library achieving a notable NDCG@10 score of 0.2555.
- Utilized **SHAP** and **KernelSHAP** to explain model predictions and identify features with the highest impact on recommendations, improving **model interpretability**.
- Integrated **TimeSHAP** to provide time-based explanations of recommendations, enabling users to understand better why specific items are being recommended at a given time.

Fractal Analytics   **Nov 2021 - Jun 2022**

Data Engineer

- Designed **data-driven solutions** for client enterprise data warehouses using **GCP** (Google Cloud Platform) products (Composer, Dataflow, BigQuery, Cloud Storage), increasing data processing efficiency by **30%**.
- Implemented an end-to-end **data encryption strategy** with Google Tink and KMS, reducing security incidents by **80%** and ensuring compliance with data protection regulations.
- Won **Star Award 2022** for delivering outstanding solutions and solving complex technical customer challenges in a short period at Fractal Analytics

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

Data Engineer

- Constructed a cost-efficient **GCP** data pipeline (**40%** cost reduction) with improved scalability and fault tolerance.
- Devised a **product ranking** auto-sort algorithm, boosting CTR (click-through rate) by **26%** and A2C (add-to-cart rate) by **12%**.
- Optimized personalized recommendations employing **ALS** (used implicit ratings), resulting in a significant **14%** increase in RPTI (revenue-per-thousand impressions).
- Developed a system to send bulk **emails** using Amazon SES with a **10x ROAS** (return-on-ads-spend) leveraging **RFM** (Recency-Frequency-Monetary) **segmentation** for user base selection.
- Recognized as **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

- Built 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.
- Constructed a **data warehouse** with HDFS, Spark, and Cassandra, enabling systematic computational analysis for Product Managers and Leaders.
- Automated raw **data compression** (**80%** reduction) in Hadoop Cluster using Airflow, Spark, and MySQL, ensuring storage capacity remained below **20%** and optimizing external storage.

TECHNICAL SKILLS

- Programming Languages: Python, Java
- Databases: MySQL, MongoDB, Cassandra, Elasticsearch
- Big Data & Cloud: Kafka, Spark, Flink, HDFS, Airflow, Google Cloud Platform - Pub/Sub, BigQuery, Dataflow, Key Management Service, Cloud Functions
- 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)