

ADIT SHRIMAL

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

PROFESSIONAL EXPERIENCE

Optum (UnitedHealth Group)

Aug 2023 - Present

Senior Data Scientist

- Developed an advanced predictive model using **XGBoost** to identify high-risk healthcare claims, leading to a significant improvement in **True Positive Rate (TPR)** from 52% to 71% on test data.
- Extracted sequential patterns from diagnosis and procedure codes using **BioBERT embeddings** generated on **AzureML**, and created comprehensive risk and seasonal features to enhance model accuracy.
- Implemented robust model and data drift monitoring using **Population Stability Index (PSI)** to ensure model performance and reliability.
- Consistently applied **SHAP** values to ensure transparency and provide explanations for model predictions, improving trust in the model's decisions.

Amazon Web Services

Nov 2022 - Jul 2023

Machine Learning Engineer - Internship

- Devised a comprehensive **data preprocessing pipeline** to prepare datasets for training the GRU4Rec model.
- Designed and implemented **Recurrent Neural Networks (RNNs)**, specifically the **GRU4Rec** model, using **PyTorch** and the **ReChorus** library, achieving a notable **NDCG@10** of 0.42.
- Applied **SHAP**, **KernelSHAP**, and **TimeSHAP** to enhance the interpretability of machine learning models and provide time-based explanations for recommendations.

Fractal Analytics

Nov 2021 - Jun 2022

Data Scientist / Engineer

- Architected data-driven solutions for client data warehouses using **GCP** (Google Cloud Platform) products (**Airflow**, **Dataflow**, **BigQuery**, **Cloud Storage**), reducing data processing time by 30%.
- Boosted conversion rate from 20% to 23.6%, an 18% lift, by implementing **RFM (Recency Frequency Monetary) segmentation** in marketing strategies.
- Implemented a predictive model using **logistic regression** to forecast the buying behavior of customers, which led to a 15% increase in cross-selling opportunities.
- Built an end-to-end **data encryption** strategy with **Google Tink** and **KMS**, reducing security incidents by 80%.

Bewakoof Brands Pvt. Ltd.

Feb 2020 - Nov 2021

Data Scientist

- Devised a **product ranking** auto-sort algorithm, boosting **CTR (click-through rate)** by 26% and **A2C (add-to-cart rate)** by 12%.
- Utilized the **Alternating Least Squares (ALS)** method in the design and implementation of a **personalized recommendation engine**, resulting in a 14% rise in RPTI (revenue-per-thousand impressions).
- Developed a **sentiment analysis** model for customer reviews using **NLP**, leading to a 20% decrease in customer complaints and a 15% increase in positive feedback.

Reliance Jio

Jul 2018 - Feb 2020

Data Engineer

- Engineered a streaming data pipeline using **Kafka**, **Flink** & **Elasticsearch**, handling ~150k/sec events in real-time with zero data lag and outperforming the existing pipeline that had a data lag of two hours.
- Built comprehensive **dashboards** for non-technical stakeholders using **Tableau** and **Kibana**, which enabled Product Managers and Leaders to easily monitor key metrics and **business KPIs**.

EDUCATION

M.S. in Data Science, University of San Francisco

Jul 2022 - Jun 2023

Relevant Coursework: Probability & Statistics, Statistical Modeling, Linear Regression, Machine Learning, A/B Testing, Distributed Computing (Spark), Relational Databases (SQL), NoSQL, MLOps, Data Engineering

B.E. in Computer Engineering, University of Mumbai

Jul 2015 - May 2018

TECHNICAL SKILLS

- Programming Languages:** Python, Java, Bash
- Databases:** MySQL, MongoDB, Elasticsearch
- Big Data & Cloud:** Kafka, Spark, Flink, HDFS, Airflow, Google Cloud Platform (GCP), Amazon Web Services (AWS)
- Machine Learning:** Linear Regression, Logistic Regression, Decision Trees, Random Forest, Kmeans, Gradient Boosting Machines (XGBoost), SVD, PCA, Collaborative Filtering (ALS), Natural Language Processing (NLP).
- MLOps:** MLFlow, Data Version Control (DVC), Git, CI/CD, Docker, Kubernetes, API design (Flask)
- Data Visualization:** Tools - Tableau, Metabase, Kibana; Libraries - Matplotlib, Seaborn