**ADIT SHRIMAL**

[adit.p.shrimal@gmail.com](mailto:adit.p.shrimal@gmail.com) | +1 (408) 203 2174 | San Francisco, CA, USA | [www.aditshrimal.com](http://www.aditshrimal.com) | [LinkedIn](http://www.linkedin.com/in/aditshrimal) | [Github](https://github.com/aditpshrimal)

**PROFESSIONAL EXPERIENCE**

**Optum (UnitedHealth Group) Aug 2023 - Present**

*Senior Data Scientist*

* Developed a financial abuse detection system for high-risk healthcare claims using XGBoost (PySpark, SQL), improving TPR from 52% to 71% while maintaining low false positives.
* Engineered temporal and behavioral fraud signals from sequential diagnosis/procedure codes using BioBERT embeddings in AzureML.
* Designed model and data drift monitoring using Population Stability Index (PSI) to ensure model reliability.
* Applied SHAP values to produce localized, review-ready explanations for flagged claims, reducing investigation turnaround time.
* Created dashboards and presented risk model insights to leadership, enabling data-backed fraud mitigation strategies.
  + Enabled estimated savings of $12M+ annually by reducing undetected overpayments and enhancing manual review efficiency.

**Amazon Web Services Nov 2022 - Jul 2023**

*Machine Learning Engineer - Internship*

* Designed and deployed a GRU4Rec-based recommendation system using PyTorch, achieving NDCG@10 of 0.42 on sequential benchmark data as part of an internal proof-of-concept for explainable recommendations.
* Led the initiative on explainability using SHAP, KernelSHAP, and TimeSHAP, and co-authored an internal research paper with the AWS team (under review).

**Fractal Analytics Nov 2021 - Jun 2022**

*Data Scientist / Engineer*

* Architected data-driven solutions for client data warehouses using GCP (Airflow, Dataflow, BigQuery, Cloud Storage), reducing data processing time by 30%.
* Boosted conversion rate from 20% to 23.6%, by implementing RFM segmentation in marketing strategies.
* Built a customer propensity model using logistic regression to identify behavioral patterns, enabling 15% lift in cross-selling and fraud flagging.

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

*Data Scientist*

* Devised a product ranking auto-sort algorithm, boosting CTR by 26% and A2C by 12%.
* Utilized ALS in the implementation of a personalized recommendation engine, resulting in a 14% rise in RPTI.
* Conducted A/B testing on homepage layout variants to compare static vs. dynamic product widgets, leading to a 6% increase in conversion rate for new users.
* Developed a sentiment analysis model for customer reviews using NLP, leading to a decrease in customer complaints.

**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.
* Built dashboards using Tableau and Kibana, enabling non-technical stakeholders to monitor key 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, Oracle, 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