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 **predictive** models and **statistical** techniques to identify **healthcare overpayment patterns**, reducing fraud-related costs and increasing model efficiency with a **36% improvement** in **True Positive Rate** (TPR).
- Extracted **sequential patterns** from **ICD-10 diagnosis** and **CPT (procedure)** codes using **BioBERT** embeddings on **AzureML**, engineering risk features to enhance model accuracy by **15%**.
- Implemented robust **model and data drift monitoring** using the **Population Stability Index (PSI)** to ensure model performance and reliability.
- Leveraged **SHAP** values for transparent problem-solving, improving **communication** and **collaboration** across **cross-functional teams** and leadership, and supporting strategic decision-making.

Amazon Web Services   **Nov 2022 - Jul 2023**
Machine Learning Engineer - Internship
- Designed and deployed **recommendation systems** using **GRU4Rec** models with **PyTorch**, achieving a notable **NDCG@10** score of **0.42**.
- Optimized **data preprocessing pipelines**, reducing manual intervention and improving query execution efficiency by **50%**.
- Enhanced **model explainability** using **SHAP** and **TimeSHAP**, providing actionable insights for **product strategy**.

Fractal Analytics   **Nov 2021 - Jun 2022**
Data Scientist / Engineer
- Conducted **product analytics** using **Python** and **SQL** to identify customer segments, boosting **conversion rates** by **18%** through **RFM (Recency Frequency Monetary)** segmentation.
- Collaborated with cross-functional teams to develop **client data warehouses** using **GCP** (Google Cloud Platform) products (**Airflow, Dataflow, BigQuery, Cloud Storage**), reducing data processing time by **30%**.
- Built **sentiment analysis models** using **NLP** techniques to analyze customer reviews, leading to a **20% reduction** in **complaints**.

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%.
- Designed **recommendation engines** with **Alternating Least Squares (ALS)**, resulting in a **14%** rise in **RPTI** (revenue-per-thousand impressions) i.e. improving application engagement and growth metrics.
- Designed and executed **A/B testing frameworks** to evaluate **marketing strategies**, increasing campaign **ROI** by **12%** and aligning with product development roadmaps.
- Conducted **sales forecasting** to optimize **inventory management**, reducing **stockouts by 20%**.

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 with two-hour data lag.
- Built comprehensive **dashboards** for non-technical stakeholders using **Tableau** and **Kibana**, enabling Product Managers and Leaders to monitor key metrics and **business KPIs** easily and improve problem resolution.

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