ANUSHKA JAIN

San Francisco Bay Area | jain.anushka.tech@gmail.com | 240-355-0262 | Portfolio | LinkedIn | GitHub

Education

University of Maryland, MS in Business Analytics — Aug 2024 - Dec 2025
– Coursework: Data Mining, Predictive Analytics, Cloud Computing, Database Management Systems, Price Optimization

Manipal Institute of Technology, Bachelor of Technology, Biotechnology — Jul 2018 - Jun 2022

Skills

– **Machine Learning and Statistics:** Predictive Modeling, Classification, Regression, Ensemble Methods (XGBoost, LightGBM, Random Forest), Time Series Forecasting (SARIMA), Clustering, Feature Engineering, A/B Testing
– **Programming:** SQL, Python (pandas, numpy, XGBoost, LightGBM, Scikit-learn), R, PySpark
– **Cloud and Tools:** Azure Databricks, Azure Data Factory, AWS, Git, Advanced Excel, MS Office, JIRA
– **Visualization Tools:** Tableau, PowerBI, Matplotlib, Seaborn

Experience

Walmart, Data Scientist – Capstone Project — Oct 2025 - Dec 2025
– Developed a department-level inbound forecasting pipeline in Python for **100+** stores, achieving **6.6% MAPE (less than 1% bias**) using ensemble time-series models (LightGBM, XGBoost, SARIMA) to support labor planning and optimization
– Engineered 30+ time-series features including lag variables, rolling statistics, and holiday window features capturing Q4 seasonal surge and weekly patterns across departments
– Built an LLM-powered Agentic AI interface to operationalize forecasting models, enabling stakeholders to generate predictions via natural language prompts (store number, department, forecast horizon)

University of Florida, Research Intern – Product | *Gainesville, Florida* — Jun 2025 - Aug 2025
– Developed Python implementation of T-GERGM algorithms for temporal social network modeling, performing exploratory analysis to identify structural patterns in time-dependent network data for downstream modeling
– Collaborated with 3+ cross-functional teams to synthesize modeling approaches, translate results into actionable insights and facilitate R-to-Python integration

Tredence Inc. | *Bengaluru, India*
Consultant – Customer Analytics — Jan 2024 - Jul 2024
– **Reduced query latency by 20%** and accelerated feature extraction for downstream models by implementing a Medallion-architecture Customer Data Mart in Azure Databricks
– **Improved data accuracy by 30%** and reduced manual reporting by 20 hours/week by deploying an automated data quality and reporting framework in PySpark with CI/CD pipelines
– Optimized and maintained SQL-based ETL pipelines for high-volume datasets, ensuring reliable data transformation and consistent downstream analytics

Data Analyst – Retail & Marketing Analytics — Jul 2022 - Dec 2023
– Reduced cancellations and stockouts by **15%** and improved customer satisfaction by **leading a 5-member analytics** team to design demand forecasting and inventory optimization models
– Identified key drivers of customer churn and behavioral patterns by using logistic regression and K-means clustering models, translating results into features that informed pricing experiments and customer segmentation
– Automated marketing KPI tracking and A/B testing workflows using hypothesis testing in Python and SQL, reducing reporting latency by **15%**

Projects

Airbnb Perfect Rating Prediction Model *GitHub*
– Engineered 45+ features from listing attributes and text data (TF-IDF, sentiment analysis) to build classification models (LightGBM, XGBoost) achieving 76% accuracy, 80% AUC, and <10% false positive rate
– Conducted model validation and performance optimization across multiple algorithms, tracking metrics to balance precision-recall tradeoff for reliable host recommendations

Qatar Tourism Tableau Dashboard *Tableau Link*
– Built an interactive Tableau dashboard analyzing 10 years of tourism data, uncovering insights on visitor spending patterns and regional travel trends to inform strategic planning