ANIKET SHARMA

+91 8482069077 ⋄ Bengaluru, Karnataka

aniketsharma1378@gmail.com ⋄ linkedin.com/in/aniketsharma20/

OBJECTIVE

Data scientist with 2 years of experience in banking industry, building and analysing machine learning models with a focus on model explainability.

EDUCATION

Master of Science in Quantitative Economics, Indian Statistical Institute, Kolkata   2021 - 2023
Relevant Coursework: Econometrics, Statistics and Time Series Modelling.

Bachelor of Economics, University of Calcutta   2018 - 2021

EXPERIENCE

Wells Fargo   June 2023 - Present
Quantitative Analytics Specialist [Fraud Modelling]   *Bengaluru, Karnataka*

- Developed first-party fraud prediction system for credit card application, using XGBoost along with Bayesian Optimization for $2B retail services portfolio, achieving an estimated on $1.2 M savings in fraud loss prevention.
- Conducted Generative AI-based research on first-party fraud detection using Variational Auto-Encoder (VAE), generating synthetic data to model fraudulent behaviors enhancing early fraud detection and improving recall by 18% for high-risk applicants at onboarding.
- Developed a challenger model for the existing credit scorecard model using lightGBM for personal loans, improving model AUC by 20% and KS by 24%.
- Built a challenger model to predict check return using logistic regression for credit card portfolio to compare with industry leading vendor models.
- Automated analysis and reporting of model performance for vendor models, reducing computation time by 40% through Spark optimizations. Created a Tableau dashboard for reporting to provide enhanced business insights.

Expedia Group   June 2022 - August 2022
Machine Learning Science Intern [Intents Modelling]   *Gurugram, Haryana*

- Developed a Python script for automated performance analysis of intent prediction models.
- Conducted segmentation evaluations and identified target leakage in the model pipeline.
- Extended evaluation frameworks to granular, event-level performance analysis.

PROJECTS

RAG-LLM Developed a Retrieval-Augmented Generation (RAG) pipeline integrating LLMs with vector databases for efficient context-aware question answering. Implemented document chunking, semantic search, and evaluation of multiple embedding models, enabling scalable and explainable AI-driven retrieval systems.

Credit Card Transaction Fraud Detection Developed an XGBoost-based third-party fraud detection model achieving AUC-ROC of 0.91 and recall of 0.53 on highly imbalanced data. Created end-to-end pipeline with feature engineering (velocity, days since features), hyperparameter tuning using optuna, and model interpretability (SHAP).

SKILLS

Programming & Tools	Python, PySpark, SQL, SAS, Tableau, Databricks, LangChain, ChromaDB
Machine Learning	Logistic Regression, XGBoost, lightGBM, SHAP, Bayesian HPT, RAG
Deep Learning	ANN, GANs, VAE, Pytorch
Data Analysis and Visualization	Pandas, NumPy, Matplotlib, Seaborn
Big Data Technologies	Hadoop, Spark