

Rajarshi Dutta

☎ +91 9735952230 , ✉ rajarshidutta2015@gmail.com , 🌐 [Rajarshi1001](#) , in [linkedin](#)

PROFESSIONAL EXPERIENCE

American Express

February 2025 to Present

Data Scientist, Commercial Marketing Decision Sciences, Credit and Fraud Risk

- Architected multi-offer ranking framework using response-to-gross scores with **96%** match rate thereby replacing existing 3 **XGboost** models with **2** segmentations, delivering **\$60 M** revenue impact for the **Email** channel across commercial and consumer card offers
- Developed an overall **Contact History** variable by integrating multiple channel information via multiple tables, enabling daily refresh and production onboarding with comprehensive User Acceptance Testing (UAT), resulting in a **3%** improvement in Gini
- Built Spend-Centric propensity model using **XGBoost** with **90+** features (spend, risk, contact history, real-time triggers), consolidating **8+** channel-specific models deployed for daily scoring across the entire US AIF base, driving **\$30M** in incremental GCM
- Owned end-to-end **deployment** workflow for the model including base creation, risk suppressions, variable attachment to scoring, tracking and analytical runs while ensuring adherence to model governance and validation procedures prior to production release
- Implemented an **LSTM**-based response model, utilizing scores to optimize segment allocation and integrated attention-weight visualization and sequence-level **SHAP** contributions for interpretability with estimated **15%** increase in conversion rate
- Spearheaded migration of **3** marketing models from **PySpark**-based **ETL** to **Google BigQuery**, refactoring feature engineering pipelines, implementing automated model tracking and scoring script and achieving **30%** runtime reduction

Decision Point Analytics

July 2024 to December 2024

Data Scientist, Pricing and Promotion modelling

- Developed a **Hybrid RAG** with Reciprocal Rank fusion based Sales insights generation tool using **LangChain**, **Pinecone VectorDB**, incorporating **CoT reasoning** over metadata-filtered sales report corpora reducing manual review time by 5%
- Streamlined SKU-level promotion modeling across **500+** SKUs in **12** markets using **discount adjustments**, **VARMAX**, **OLS**
- Built an **MLflow** based automated forecasting pipeline for monthly model scoring on SKU-level with periodic models refresh

INTERNSHIP

MITACS Research Internship (Link)

May 2023 to August 2023

Prof. Oleksandr Voznyy, Clean Energy Group, UofT Scarborough, Toronto

- Implemented **graph message passing networks** using **relational ConvNet** incorporating bidirectional updates and residual connections, **CGCNN** for band gap and orbital energy predictions of bilinear **zincblende** and **rocksalt** crystal structures
- Created a custom **maggie**, **orbital-aware and structural** embedding scheme thereby achieving an MSE score of **0.03** and **0.23**
- Devised an **XGBoost** band-gap prediction model from **2048** sized organic **Morgan fingerprints**, achieving an MSE of **0.30**

KEY PROJECTS

Safe Exploration in Continuous Action Spaces (Link)

March 2024 to April 2024

- Developed RL models like **DDPG**, **TD3** for safe exploration in custom **OpenAI gym** environments like Box-2D, spaceship etc
- Designed a **3-layer feedforward safety network** for action correction after every deterministic policy update in **Pytorch**
- Attained **zero safety violations** across **10,000+** episodes, recovering **85-95%** of optimal unsafe policy performance

E-Commerce Review Score Forecasting using MLflow (Link)

April 2024

- Implemented an **MLOps** pipeline using **ZenML** across 4 stages (ingestion, cleaning, training, evaluation), achieving F-score of **0.88**
- Automated model deployment using **MLflow** model registry and a custom **ZenML** stack, reducing manual deployment steps by **60%**
- Deployed a live regression model server using **MLflow**'s model deployer, enabling real-time inference requests

Neural Machine Translation Challenge (Link)

November 2023

- Explored SOTA models for neural machine translation from English to seven major Indian languages and vice-versa
- Implemented **LSTM/GRU** with attention and **Transformer** with periodic positional embeddings in **Pytorch**
- Used tokenization strategies like **BoW**, **TF-IDF** and embedding methods like **FastText**, **GloVe** for effective encoding
- Achieved **BLEU** score of **0.101** with Transformer architecture on curated validation set for multilingual Indian language translation

TECHNICAL SKILLS

Programming	Python, C/C++, R, MATLAB, Bash, Javascript
Databases	MongoDB, SQL, SQLite, ChromaDB, PineconeDB, PostgreSQL
AI/ML tools	PyTorch, Scikit-learn, Keras, MLflow, OpenCV, OpenAI, Gymnasium, LangChain/LangGraph, LLM, MLflow
Other Tools	Docker, Git, Advanced Excel, Power BI, Tableau, BigQuery, PySpark, Map-Reduce, Hive, GCP, Claude Code

EDUCATION

Indian Institute of Technology (IIT) Kanpur

2020 to 2024

B.Tech, Materials Science and Engineering

8.2/10.0

ACHIEVEMENTS

- Secured an AIR **5610** in JEE Advanced 2020 among 1.5 lakh shortlisted candidates
- **Analyst of the Quarter Award - Q3'25**, American Express for Email Response model usage optimization analysis
- Earned **Bronze Medal** in Inter IIT Tech Meet 11.0 - **ISRO's The Chandrayaan Moon Mapping Challenge**
- Secured **2nd Position** in Inter IIT Tech Meet 10.0 - **ISRO's Web Based X-Ray Burst Identification**