

## EDUCATION

University of Pennsylvania (School of Engineering and Applied Science)

Master of Science in Engineering (MSE) in Data Science; GPA 3.9/4.0

**Coursework:** Machine Learning, Databases, Computer Systems, Deep Learning, Generative Modeling, NLP, Machine Perception

Philadelphia, PA

Aug 2023 – May 2025

Kirori Mal College (KMC – University of Delhi)

Bachelor of Science (BS) in Statistics; GPA: 8.11/10

Delhi, India

July 2017 – Aug 2020

## EXPERIENCE

AI Researcher, [Penn Medicine](#) - [Computational Social Listening Lab](#)

Dec 2024 – Present

Project 1: (IH Risk Model)

- Reduced **model tuning time by 10x** by developing an **AutoML** pipeline using sklearn for hernia prediction (**recall: 0.9**)
- Deployed GPT-4 few-shot pipeline on 10k+ redacted surgical notes, **reducing 30% noise vs. BERT** for feature labeling.
- Integrated **vLLM** & **LangChain** for scalable local LLM inference deployed on A100 GPU and **modularized the backend** (OpenAI, Mistral) using **parallelized batch inference** and **memory-aware chunking** across variable length notes.
- Spun up a FastAPI server + TypeScript frontend for RAG workflow, **slashing** surgical data analysis time from **hours to seconds** through ChromaDB vector embeddings and GPT-4 integration.

Project 2: (Misinformation)

- Built an NLP pipeline (RoBERTa) to flag health misinformation in ~20k posts, **boosting precision 20%** via entailment.
- Extracted linguistic features from posts using [DLATK](#) and applied LDA for topic modeling and performed correlation analysis.
- Consolidated **10K+ survey responses** from multiple platforms into a secure MySQL server and engineered features in **PySpark**.

Data Science Intern (Full-time), [Universal Media \(PA, USA\)](#)

May 2024 – Aug 2024

- Led the development of **3+ data pipelines** using Azure Data Factory, facilitating seamless ingesting into Azure
- Developed python scripts for data transformation, stored them in Blob storage and executed them via batch activity in ADF.
- Drove product insights by building **Mixed Media time series models in Azure Synapse**, analyzing marketing channel impacts on media diversity metrics. Built **Power BI** dashboards to deliver actionable insights for optimizing client strategies.
- Improved query efficiency by 30%** via stored procedures, parameterization, and indexing of high-frequency access paths.

Assistant Manager (Full-time), [IIFL Finance Ltd](#)

Apr 2022 – July 2023

- Analyzed ETL process failures and created **10+** paginated reports in SSIS to help the management track 1000+ branches.
- Optimized & migrated complex SQL queries from an obsolete database server that improved the **reporting services by ~40%**.
- Digital Adoption:** Led a product-focused initiative to identify digitally savvy customers by engineering features and building ADF pipelines to track campaign behavior. Trained and deployed a Random Forest model (**with a 90% accuracy**) in Azure ML Studio; **exposed it as a REST endpoint consumed by marketing campaigns**, driving digital **disbursal adoption by 50%**.

## SELECTED PROJECTS

- Particle (2024):**
  - Built a context-aware chatbot for product and order support using FastAPI, LangChain, and Pinecone, implementing **RAG for product queries and tool-based logic with LLM fallback** for transactions.
  - Engineered multi-turn memory and entity tracking, **enabling >95% accurate follow-up handling**. Automated ingestion of synthetic data, & delivered **React + Vite frontend, improving support efficiency by 40%** and user satisfaction. [\[Link\]](#)
- Ride Duration Prediction (2025):** Developed a production-ready ML pipeline to predict NYC taxi ride durations, using **Airflow** for orchestration and **MLflow** for experiment tracking. Achieved **~30% RMSE reduction** via automated hyperparameter tuning & designed modular, reproducible workflows to simulate real-world deployment as a web service via **Flask**. [\[Link\]](#)
- BSE quarterlies (2025):** Achieved **500x faster responses (~100ms)** in fetching quarterly closing prices from Bombay Stock Exchange for a client by replacing **Playwright browser automation** with **Requests + BeautifulSoup (with ASP.NET WebForms token)** posts, cutting **cold-start and memory**. Dockerized & deployed an application on Google Cloud Run. [\[Link\]](#)
- Multithreaded Image Processing (2025):** Engineered a parallelized box blur algorithm in C++ using POSIX threads, achieving a 2.8x speedup (3251 ms → 1165 ms) on 4 cores by optimizing memory access, leveraging shared-memory synchronization, and **partitioning workloads across non-overlapping thread-local regions**. [\[Link\]](#)

## TECHNICAL SKILLS

**Programming Languages:** Python, C/C++, SQL, R programming, JavaScript

**ML Libraries/Frameworks:** PyTorch, scikit-learn, XGBoost, LightGBM, HuggingFace, spaCy, OpenCV, MLflow, A/B Testing

**Databases/Web Frameworks:** MySQL, PostgreSQL, SSMS, MongoDB, Neo4j, React, NodeJS, Django, Flask, FastAPI, Typescript

**Cloud/Big Data Orchestration:** AWS (S3, Glue), Azure (Data Factory, Synapse), GCP (BigQuery), Kafka, Airflow, DBT, PySpark

**Tools/DevOps:** DataBricks, Apache Spark, Docker, Jenkins, Git, Kubernetes, pytest, CI/CD