

EDUCATION

University of Virginia

Master of Science in Data Science, GPA: 4.0/4.0

Charlottesville, VA

Jun 2024 – May 2025

Teaching Assistant - Statistical Machine Learning • Data Science Ambassador • Academic Village Resident (The Range)
Coursework: Large Language Models (LLMs) • Statistical and Bayesian Machine Learning (ML) • Deep Learning • Big Data Systems • Natural Language Processing (NLP) • Data Structures and Algorithms • Data Engineering • Data Ethics

Birla Institute of Technology and Science (BITS Pilani)

Bachelor of Engineering in Mechanical Engineering, Minor in Data Science, GPA: 3.8/4.0

Hyderabad, India

Aug 2018 – Dec 2021

EXPERIENCE

Machine Learning Researcher

LMI Inc. (Capstone Project)

Tysons, VA (Remote)

Nov 2024 – May 2025

Developed an **Agentic AI-driven Multimodal RAG-based LLM system** to research the impact of multimodal capabilities on improving educational content retrieval, summarization and comprehension.

- Built a multimodal embedding pipeline with SentenceTransformers and CLIP, indexing **over 5,000 data chunks from diverse sources** (text, images, PDFs, lecture slides, and audio transcripts) for efficient semantic search via Pinecone.
- Benchmarked 3 Major RAG Frameworks - (**Zero-Shot, Text, and Image-aided RAG**) identifying improvements in retrieval accuracy and summarization quality across LLaMA and GPT-4 models.
- Implemented the RAGAS framework and key NLP metrics to assess system efficacy with **100+ pre-classified question sets** with varied levels of specificity, with attention visualizations to validate multimodal understanding.

Lead Data Scientist

Piramal Pharma Ltd.

Bethlehem, PA (Remote, Mumbai Office)

Jan 2022 – Jun 2024

Established and led the Data Science and Analytics Team for the Critical Care Division, mentoring a high-performing team of 3 data scientists and integrating advanced analytics with best practices into core business functions.

- Prevented \$1.2M in quarterly inventory losses** by designing a time-series forecasting system using S-ARIMA and Monte Carlo simulations to improve depleting stock visibility and optimize replenishment cycles.
- Engineered **over 30 data analytics pipelines** using Python and Azure Databricks, reducing reporting turnaround across business functions from monthly to daily and **improving manual reporting efficiency by 85%**.
- Developed end-to-end executive dashboards integrating market data, competitive benchmarking, and segmentation analysis **adopted by the CFO to steer data-driven strategic decisions and growth initiatives**.
- Enabled personalized employee growth strategies by analyzing performance reviews using Text Sentiment Analysis model built with VADER and TextBlob, **driving CHRO's talent development strategies**.

Data Science Intern

Piramal Pharma Ltd.

Mumbai, India

Jul 2021 – Dec 2021

Interned in the Consumer Products Division, contributing to high-impact business-facing data science initiatives.

- Achieved **15% quarterly increase in market penetration across 400,000 retailers** by building the company's first product recommendation engine implementing the FP-Growth algorithm from scratch.
- Boosted high-performing sales reps retention rates to 93%** by developing an attrition prediction model using XGBoost and Logistic Regression on historic sales performance and attrition data.
- Recognized for impact by the CEO and **offered a leadership role to build and head a new analytics team**.

SKILLS

Programming Languages: Python, R, SQL, JavaScript

AI and ML: Traditional ML (Scikit-Learn), Deep Learning (PyTorch, TensorFlow), LLMs (LangChain, HuggingFace)

Cloud Data Tools: Microsoft Azure (DataBricks, Data Factory, ML Studio), AWS (SageMaker, Dask, EC2), Git, Docker

Data Visualization: Power BI, Plotly Dash, d3.js, QlikSense

Interests and Hobbies: Singing and enjoying music - Carnatic Ragas, playing card games and badminton, robotics