

# Vish Guruvayur

(571) 602-2034 | vishguru2704@gmail.com | github.com/vishugp | linkedin.com/in/vishgp | vishguru.com

## SUMMARY

Data Scientist with 3+ years of experience delivering **30+ ML and NLP products** across healthcare, retail, and e-commerce. Skilled in Python, SQL, PyTorch, leveraging HuggingFace, LangChain, and AWS (Certified). MS in Data Science from University of Virginia (4.0 GPA) with **peer-reviewed research** on multimodal retrieval. Passionate about applying **advanced Gen-AI and applied ML expertise** to turn complex data into interpretable solutions that drive meaningful impact.

## EXPERIENCE

### AI Researcher

University of Virginia, School of Data Science

Charlottesville, VA (Remote)

Nov 2024 - Present

- Led applied research on application of **multimodal and hybrid RAG** methods on instructional content (text, images, and videos from classroom materials) to improve content retrieval, comprehension, and summarization.
- Built a scalable RAG pipeline with **10,000+ multimodal embeddings** using LangChain and Pinecone DB, conducting retrieval experiments to optimize image vs. text vectors, improving answer relevance and **reducing LLM compute cost by 20%**.
- Co-authored a **peer-reviewed paper** ([link](#)) presented at AI for Education 2025, introducing novel techniques for educational content retrieval and evaluation; awarded "Most Compelling Data Visualizations" at UVA's Capstone Showcase.
- Currently conducting R&D on a **GraphRAG-based LLM assistant**, extending prior research to leverage structured retrieval over temporally-linked educational content, enhancing student support and informing instructional content design.

### Senior Data Scientist

Piramal Pharma Ltd.

Bethlehem, PA (Remote)

Jan 2022 - Jun 2024

- Established and **led the Data Science team** for the Critical Care division, embedding high-impact analytics into key business functions while mentoring junior scientists to deliver scalable end-to-end solutions with best practices.
- Prevented \$1.2M in quarterly losses** by providing analytical visibility into inventory risk conditions using causal inference and SARIMA-based time-series forecasting of inventory, write-offs, and sales under volatile demand.
- Partnered with CFO and senior leadership to develop market segmentation models and competitive intelligence dashboards (AWS, Power BI), guiding capital allocation, price elasticity, and portfolio **growth decisions across 5 global regions**.
- Reduced Advertising Cost of Sales **from 45% to 8% in 6 months** for high-priority e-commerce products using alternative data (web-scraping and NLP) to drive market segmentation and A/B-tested ad optimizations.
- Received the "Piramal Excellence Award" for building **30+ ML and analytics pipelines** using Databricks and PySpark to automate model retraining triggers and KPI reporting dashboards, improving overall operational efficiency by 85%.

### Data Scientist

Piramal Pharma Ltd.

Mumbai, India

Jul 2021 - Dec 2021

- Drove a **15% quarterly increase** in market penetration across 400,000 retailers by building the company's first product recommendation engine, coding a tree-based custom FP-Growth algorithm from scratch.
- Boosted high-performing sales reps' **retention to 93%** by developing an attrition prediction model using an ensemble of XGBoost and Logistic Regression on historical sales and employee data.

## SKILLS

*Languages:* Python, R, SQL, Scala, JavaScript, Bash

*LLMs & GenAI:* Transformers, LangChain, HuggingFace, RAG, Prompt Engineering, PEFT (LoRA, IA3), Fine-tuning

*Machine Learning:* Scikit-learn, XGBoost, NLP (SpaCy, VADER), Statistical Modeling, A/B Testing, Recommendation Engines

*Visualization:* Power BI, Plotly Dash, Seaborn, Matplotlib, Excel, VBA

*Cloud and Tools:* AWS (Certified), Azure, Databricks, Snowflake, Git, Conda, Flask

*Data Engineering:* Apache Spark, Snowflake, Hadoop, MLflow, Airflow, dbt, Docker, APIs, Web Scraping

## CERTIFICATIONS

- AWS Certified Solutions Architect – Associate, Amazon Web Services May 2025 - May 2028
- AWS Certified Machine Learning – Specialty, Amazon Web Services Jun 2025 - Jun 2028

## EDUCATION

### University of Virginia (UVA)

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

Data Science Ambassador • Teaching Assistant - Statistical Learning • Range Resident

Charlottesville, VA

Jun 2024 - May 2025

### Birla Institute of Technology and Science (BITS Pilani)

B.E. Mechanical Engineering, Minor in Data Science, GPA: 3.8/4.0

Hyderabad, India

Aug 2018 - Dec 2021