Gainesville, VA
+1 (571) 602-2034
vishguru2704@gmail.com

Vishwanath (Vish) Guruvayur
Data Scientist

vishguru.com
github.com/vishugp
linkedin.com/in/vishgp

SUMMARY

Full-Stack Data Scientist with over 3 years of experience in evaluating generative AI models, building recommendation systems, and analyzing user and market behavior to drive personalized experiences. Delivered 30+ cross-functional impactful projects and created music tools like the Spotify Artist Map and Carnatic Raga Viz, fueled by a passion for music discovery.

EXPERIENCE

AI Researcher
Charlottesville, VA (Remote)

LMI Inc. and UVA School of Data Science
Nov 2024 – Present

- Applied research on multimodal and hybrid Retrieval-Augmented Generation (RAG) systems, exploring their impact on educational content retrieval, summarization quality, and learner comprehension.
- Engineered a modular multimodal RAG evaluation pipeline using Sentence Transformers, CLIP, and Pinecone, enabling content retrieval and rapid experimentation across diverse embedding backends for research.
- Co-authored a peer-reviewed paper, *"Using Digital Textbook and Classroom Data to Explore Multimodal (Audio, Visual, & Textual) LLM Retrieval Techniques" [PDF]*, accepted at AIED 2025 (iTextbooks Workshop).
- Recipient of the "Most Compelling Data Visualizations" award at the Capstone Showcase for innovative and interpretable visual representations that advanced insight into multimodal RAG models.

Lead Data Scientist
Bethlehem, PA (Remote)

Piramal Pharma Ltd.
Jan 2022 – Jun 2024

- Established and led the Data Science Team for the Critical Care Division, building a high-performing team of data scientists in integrating advanced analytics and AI products with best practices into core business functions.
- **Prevented $1.2M in quarterly losses** by implementing causal inference via counterfactual analysis of inventory schedules, write-offs and sales forecasting using Monte Carlo simulations and S-ARIMA, optimizing depleting stock visibility.
- **Reduced ACOS from 45% to 8%** within 6 months for high-priority e-commerce products via ad campaign optimizations by leveraging alternative data using NLP (VADER, SpaCy), market segmentation, and A/B testing.
- **Partnered directly with the CFO** to design interactive market story dashboards combining segmentation analysis and competitive intelligence, shaping global portfolio expansion strategies across five regions.
- Designed unified monitoring workflows in Python to track both business KPIs and ML model health (data drift, prediction quality), enabling early retraining and aligning model performance with business goals.
- Received the "Piramal Excellence Award" for engineering **over 30 end-to-end analytics pipelines** using Python and SQL on Azure, optimizing reporting frequency by weeks and increasing operational efficiency by 85%.

Data Science Intern
Mumbai, India

Piramal Pharma Ltd.
Jul 2021 – Dec 2021

- Achieved **15% quarterly increase** in market penetration **across 400,000 retailers** by building the company's first product recommendation engine using the Tree-based 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.

EDUCATION

University of Virginia
Charlottesville, VA

Master of Science in Data Science, GPA: 4.0/4.0
Jun 2024 – May 2025

Data Science Ambassador • Teaching Assistant - Statistical Learning • Academic Village Resident (The Range)

AWS Certifications: *Solutions Architect – Associate*, *Machine Learning – Specialty*

Birla Institute of Technology and Science (BITS Pilani)
Hyderabad, India

Bachelor of Engineering in Mechanical Engineering, Minor in Data Science, GPA: 3.8/4.0
Aug 2018 – Dec 2021

SKILLS

Languages: Python, R, SQL, Scala, JavaScript, Bash   **Tools:** Jupyter, Flask, Docker, Git, Conda, Linux
Visualization: Power BI, Plotly Dash, Matplotlib, Seaborn, Excel
AI and ML: Stat Modeling (Scikit-Learn, XGBoost), Deep Learning (PyTorch, TF), LLMs (LangChain), NLP (NLTK, SpaCy)
Cloud and Big Data: AWS Certified, Azure Databricks, Spark, Snowflake, Hadoop, MLFlow
Hobbies: Singing and Music Analysis (*Spotify Artist Map*, *Classical Raga Viz*), *Card Games*, Chess, Badminton