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

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

Vishwanath (Vish) Guruvayur
Data Scientist II

EXPERIENCE

Machine Learning Researcher
Tysons, VA (Remote)

LMI Inc. and UVA School of Data Science
Nov 2024 – Apr 2025

Led end-to-end development of a multimodal RAG-based LLM product for educational content retrieval, from conceptualization to deployment, achieving high adoption in academic settings.

- Developed an **end-to-end multimodal RAG pipeline** using Python, SQL, Sentence Transformers, and CLIP embeddings, indexing over **6,000 vectors** from educational text, images, and audio for semantic retrieval via Pinecone DB.
- Evaluated three RAG frameworks (Zero-Shot, Text-only, Image-augmented) to enhance retrieval accuracy and summarization quality, presenting results to academic and technical stakeholders for strategic adoption.
- **Awarded "Most Compelling Data Visualizations"** at the Capstone Showcase for designing intuitive Plotly Graphs with semantic clustering, improving understanding of LLM RAG System.

Lead Data Scientist
Bethlehem, PA (Remote, Mumbai Office)

Piramal Pharma Ltd.
Jan 2022 – Jun 2024

Led a high-performing Data Science and Analytics Team, driving data-driven product development and strategic analytics for financial planning and market expansion, influencing C-suite decisions.

- **Saved $1.2M quarterly** by developing a time-series forecasting product using S-ARIMA and Monte Carlo simulations in Python and SQL, optimizing inventory replenishment cycles and enhancing financial visibility.
- **Collaborated with CFO to launch market expansion initiatives** by designing Power BI and Plotly Dash analytics dashboards, integrating market data, segmentation, and competitive intelligence for strategic financial planning.
- **Received "Piramal Excellence Award"** for building over **30 SQL-driven analytics pipelines** using Python and Azure Databricks, reducing reporting turnaround from monthly to daily and boosting efficiency by 85%.
- **Drove talent development strategies** by developing an NLP-based product using VADER sentiment analysis and topic modeling, enabling CHRO to personalize growth plans and improve retention.

Data Science Intern
Mumbai, India

Piramal Pharma Ltd.
Jul 2021 – Dec 2021

Developed and launched data-driven products to enhance market penetration and employee retention, collaborating with business and technical teams.

- Built the company's first **product recommendation engine** using the FP-Growth algorithm in Python and SQL, driving a **15% quarterly increase in market penetration** across 400,000 retailers.
- Developed an **attrition prediction product** using XGBoost and Logistic Regression, achieving **93% retention of high-performing sales reps** and informing HR strategies.
- Recognized by the CEO for impactful product development, leading to an offer to head a **new analytics team.**

SKILLS

Languages: Python, R, SQL, JavaScript   **Tools:** Bash, Docker, Git, Conda, Jupyter, MLflow, Kubeflow
AI and ML: Machine Learning (Scikit-Learn), Deep Learning (PyTorch, TensorFlow), LLMs (LangChain), NLP (NLTK, SpaCy)
Visualization: Power BI, Plotly Dash, Matplotlib, GGPlot, d3.js, QlikSense
Big Data and Cloud: Spark, AWS (SageMaker, EC2, S3), Azure (DataBricks, Data Factory), Snowflake

EDUCATION

University of Virginia
Charlottesville, VA

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

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

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

CERTIFICATIONS

- AWS Certified Solutions Architect – Associate, *Amazon Web Services* — *May 2025 - May 2028*
- AWS Certified Machine Learning – Specialty, *Amazon Web Services* — *In Progress, Scheduled Jun 2025*