Ritwik Kashyap

Github | LinkedIn | Portfolio | ritwik21485@iiitd.ac.in | +91 8447164748 | New Delhi, India

Summary

Data Scientist skilled in end-to-end development and deployment of machine learning and deep learning models, including data processing, model training, and pipeline creation. Experienced in owning projects from conception to deployment, building production-ready solutions, and delivering measurable business impact.

Education

2021 - 2025 **B.Tech (Honors), Computer Science & Biosciences, IIIT Delhi** (CGPA: **8.67**/10.0)
Relevant Coursework: Machine Learning, Deep Learning, Linear Algebra, Probability and Statistics, Natural Language Processing, Computer Vision, Database Management, Object Oriented Programming, Explainable AI.

Work Experience

Data Scientist, ImagingIQ (Gurugram) June 2025 – Present

- Collaborating in an **8-member team** to build an end-to-end medical imaging pipeline orchestrated using LangGraph with **LLaMA 3** for agentic decision-making, automated preprocessing, and image analysis workflows.
- Implemented **diffusion-based inpainting** using AnoDDPM, improving downstream segmentation accuracy by nearly **20%**.
- Developed a model-serving infrastructure with **Triton Inference Server**, enabling the seamless integration of ML modules as nodes within **LangGraph** workflow and **reducing CPU inference time from 13 minutes to 5 minutes.**
- Implemented **real-time monitoring** and performance analysis of deployed services using **Grafana and Prometheus** dashboards, ensuring high availability and reliability.

Research Intern, TCS Research (Bhubaneswar) May 2024 - July 2024

- Collaborated in a **4-member research team** to explore antibody sequence generation using LLMs trained on protein sequences.
- **Fine-tuned** the **ProstT5 model** on an antigen-antibody dataset, achieving **30%** sequence similarity to the reference set and validating the effectiveness of LLM-based approaches for protein design.

Skills

Programming Languages	Python, C++, SQL
Data & Model Development	pandas, NumPy, NLTK, OpenCV, PySpark, PyTorch/Tensorflow, Transformers, scikit-learn
Deployment & Monitoring	FastAPI, AWS, Docker (basic), Git/GitHub, Prometheus, Grafana, Triton Inference Server
Generative AI	LangChain/LangGraph, RAG (basic), Agentic AI, LLM fine-tuning

Personal Projects

Automated Fetal Age Estimation from Ultrasound Videos Project Link

- Trained a hybrid **U-Net + LSTM model** to segment key anatomical structures from relevant ultrasound slices.
- Used OpenCV techniques **(ellipse fitting, bounding boxes)** to measure structures and calculate gestational age.
- **Reduced** sonographer time by **70%**, enabling faster, more consistent, and accurate fetal age estimation workflows.

University Database Migration and Optimization Project Link

- Designed a relational **PostgreSQL** schema for a university with 50,000 enrollments across multiple departments.
- Modeled a document-based MongoDB schema, reducing six SQL tables to four documents and improving query efficiency.
- Formulated an **ETL pipeline** for data migration and optimized **MongoDB** queries using indexing and **Adaptive Query Execution (AQE)** in Spark.

Awards & Publications

- **Dean's List for Academic Excellence (2022-23)**: Secured 9+ SGPA in an academic year at IIIT Delhi.
- **Graduated with Honors**, awarded for academic excellence, extra credits, and thesis work **(top 5% of batch)**.
- CommuteQA: Visual Question Answering for Bus Transport - R.R. Shah, **R. Kashyap**, et al., BDA, IIIT Bangalore, 2025.