Rishon Jain

Noida, UP

+91 7042497878 | rishon.jain@hotmail.com | rishonjain | rishonjain | rishonjain

ABOUT ME

CS student specializing in AI/ML, passionate about building and shipping end-to-end ML pipelines from data to deployment. Currently a Data Science Intern at a GovTech AI startup.

EDUCATION

Bennett University **2023 – 2027**

B.Tech Computer Science and Engineering (AI/ML) *Greater Noida*

TECHNICAL SKILLS

Programming Languages: Python (Advanced), SQL (Advanced), Unix/Shell Scripting, C++ (Intermediate)
ML & Analytics: PyTorch, TensorFlow, Hugging Face Transformers, Scikit-learn, LangChain, LangGraph, FAISS, Ollama, OpenCV, YOLO, Pandas, NumPy
Big Data & Cloud: PySpark, Kafka, PostgreSQL, AWS, Azure, Docker, Kubernetes, Git, CI/CD
Tools & Practices: Power BI, Tableau, Streamlit, REST APIs, Debugging, Unit Testing, Code Reviews

WORK EXPERIENCE

Blitz AI **July 2026 – Present**

Data Science Intern *Remote*

- Training and fine-tuning YOLOv26 object detection models for multi-class document layout analysis (text, tables, images, legends) on architectural/construction plan sets, as part of an AI-driven building-permit compliance and sufficiency review pipeline
- Designing a cost-optimized, tiered inference architecture — routing text-based compliance checks through a self-hosted open-weight LLM and reserving frontier vision-language models for image-dependent and low-confidence cases — to reduce reliance on paid API calls at scale

Guvi HCL **July 2025 – September 2025**

AI Content Research Intern *Remote*

- Spearheaded research into **30+** AI/ML topics, translating intricate technical concepts into accessible content for broader audiences. Analyzed engagement metrics using Google Trends and YouTube Analytics to identify high-impact content opportunities.
- Collaborated with engineers and content teams across multi-functional groups to translate technical research into educational resources. Developed strong technical communication and coordination skills.

PROJECTS

Hallucination-Aware RAG System | Python, LangChain, LangGraph, FAISS **March - 2026**

- Built a fully local, self-correcting RAG chatbot over **88K** Wikipedia articles with a DeBERTa-based NLI hallucination checker that verifies every answer against retrieved context, triggering up to **3**× LangGraph retry loops before responding.
- Engineered the full pipeline—FAISS indexing across **150K+** chunks, **Llama 3.2:3B** inference via Ollama on **4GB VRAM**, and a Streamlit UI exposing grounding scores, citations, and pipeline traces—running entirely offline.

PolyDetect | (Paper) | Python, PyTorch, TensorFlow, Hugging Face **October - 2025**

- Built a hybrid multilingual machine-generated-text detector combining frozen transformer embeddings **(XLM-R, BERT, MiniLM)** with perplexity and lexical-diversity features; best variant **(MiniLM+XGBoost)** achieved **0.9907** AUC-ROC on MULTITuDE, beating fine-tuned baselines by **11+** Macro-F1 points.
- Fixed a per-sample perplexity computation bug and added per-encoder feature scaling, enabling **20** model variants **(4 encoders × 5 classifiers)** to train in under **24 hours** on a single **4GB VRAM** workstation, no cloud required.

RoadSense | Python, PyTorch, TensorFlow, Streamlit, OpenCV, Ultralytics **August - 2025**

- Built an end-to-end road damage detection system using **YOLOv8n**, achieving **78% mAP@0.5** across 4 damage classes on RDD2022, with real-time inference at **30+ FPS** on GPU.
- Developed a Streamlit dashboard supporting video upload, asynchronous YOLO inference, and automated PDF report generation with frame-level annotations and severity classification.

CERTIFICATIONS

- Data Analysis using Python
- Statistics for Data Science with Python
- Object-Oriented Data Structures in C++
- Improving Deep Neural Networks
- Machine Learning: Regression
- Databases and SQL for Data Science
- Fundamentals of Network Communication
- Generative Adversarial Networks (GANs)