ARPIT SHROTRIYA

✉ as4893@njit.edu ▪ +1 9734184125 in LinkedIn ○ Github </> Leetcode 👤 Portfolio

Education

New Jersey Institute Of Technology   **Sept 2024 - May 2026**
*MS in Computer Science; **GPA: 3.72***   *Newark, NJ, USA*

- **Coursework:** Deep Learning With Computer Vision, DSA, Reinforcement Learning, Machine Learning, Capstone Project, Data Management System Design, Data Mining, Intro to BigData

Experience

Veeyo Tech   **Jul 2023 - Aug 2024**
Associate Data Scientist   *New Delhi, India*

- **Designed and deployed** LLM-powered conversational AI systems using Gemini APIs and fine-tuned BERT models, enabling autonomous task handling and resolving 65% of customer support queries without human intervention
- **Developed and optimized** Retrieval-Augmented Generation (RAG) pipelines integrating vector-based retrieval for grounding LLM responses on structured and unstructured documents, **reducing manual review time by 45%.**
- **Built enterprise-grade** MLOps pipelines using AWS Lambda, S3, and SageMaker for automated ingestion, preprocessing, deployment, monitoring, and **scaling of ML models over 1M+ daily records.**
- Collaborated with product teams to translate business workflows into **AI automation features, cutting payment reconciliation times by 35%.**
- **Engineered scalable data pipelines** using Python, SQL, and AWS Lambda to ingest and process high-volume time-series and transactional datasets, **improving system efficiency by 30%.**

Sail Analytics   **Jun 2022 - Jun 2023**
Data Engineer   *New Delhi, India*

- Automated report generation pipelines using Python and HTML to email Excel files, **reducing manual work by 30%**
- Applied OCR via Python to extract key data from over **1,000 receipts with 95% accuracy, significantly speeding up client insights**
- **Developed ETL workflows** to consolidate large-scale structured datasets into analytical dashboards, **supporting downstream ML modeling and business insights.**
- Enabled secure SAP HANA to BigQuery data transfers, **reducing data movement time by 20%**

Roboiotics Services   **Feb 2022 - May 2022**
Data Analyst   *New Delhi, India*

- Designed and implemented 8+ custom PyQT5-based GUIs, **reducing average data-entry time by 30% and cutting training ramp-up for new users from 5 days to 3 days.**
- **Architected secure authentication schemas** and RESTful APIs supporting 5,000+ daily sessions with 99.8% uptime, emphasizing reliability and fault tolerance.
- Collaborated with UX, backend, and QA teams in 12+ cross-functional sprints to **deliver features 20% faster than projected timelines.**

Technical Skills

- Core: Python, SQL, PyTorch, TensorFlow, Scikit-learn, Transformers, HuggingFace, Pandas, NumPy
- AI/ML: Large Language Models (LLMs), Agentic AI Systems, Retrieval-Augmented Generation (RAG), Prompt Engineering, NLP, OCR, Model Optimization Fine-Tuning
- Tools: AWS (Lambda, S3, SageMaker), GCP (BigQuery), MLOps, Model Monitoring Observability, CI/CD for AI Pipelines
- Cloud/Data: AWS, GCP, ETL Pipelines, SAP HANA, Data Engineering
- Tools & Platforms: Windsurf, Git, Postman, MySQL, MongoDB, Datadog.

Project

• **Search Engine with LLM** : • Architected an agentic LLM-based research assistant using LangChain and Streamlit, implementing RAG workflows for grounded multi-source knowledge retrieval.

• **MathSolveX** : Constructed and deployed an interactive churn prediction tool using TensorFlow and Streamlit, empowering business users to forecast churn risk with 95% accuracy based on user input.

• **Rusty-Playground** : A step-by-step adventure through Rust programming with practical mini-projects. Follow along as I build small applications from simple calculators to more complex systems, documenting my learning process and code improvements.