

Shriyans Naik

✦ shriyansnaik@gmail.com ✦ +91 7777077180 ✦ shriyansnaik.vercel.app ✦ linkedin.com/in/shriyansnaik

WORK EXPERIENCE

National Stock Exchange

Feb 2025 – Present

Data Scientist

Mumbai, MH

- Streamlined DRHP risk review process by building a domain-filtered vector database of historical risks and implementing semantic similarity matching, **reducing review time from 1.5 months to 3 days**.
- Implemented silent updates by hosting versions on AWS S3 and enabling automatic software updates.

Jio Platforms Ltd.

Oct 2023 – Jan 2025

Data Scientist

Mumbai, MH

- Achieved similar sentiment analysis accuracy to LLaMA 3.1 405B by **fine-tuning LLaMA 3.1 8B** with QLoRA, **reducing GPU utilization by 98%**.
- Deployed LLaMA 3.3 70B** in production using vLLM's continuous batching to serve 128 concurrent requests, enabling support for 10+ applications.
- Optimized data ingestion pipelines using **Ray** to parallelize API calls and embedding generation, **reducing processing time from 1 hour and 23 minutes to 12 minutes**.
- Built an LLM-based summary pipeline using Llama-3.3 70b model, condensing **80-page** Oil Market Reports into **6-page** insights, **saving 480 hours and ₹1.2 million annually**.
- Automated DevOps ticket validation for Kafka topic creation using Qwen 2.5, **reducing ticket closure time from 2 days to 3 hours** by ensuring required fields and notifying creators via Teams (Power Automate).
- Developed a Retrieval-Augmented Generation (RAG) pipeline, enabling question answering on over **1,000+ pages** of documentation stored on Azure DevOps, improving knowledge retrieval efficiency.

Jio Platforms Ltd.

Jan 2023 – Jun 2023

Data Science Intern

Mumbai, MH

- Replaced Azure OCR with a YOLOv5-based OCR pipeline for digit recognition in seven-segment displays, boosting accuracy by 50%, reducing latency by 90%, and **saving ₹1.7 million annually**.

EDUCATION

K. J. Somaiya Institute of Technology

May 2023

Bachelor of Technology in Computer Science (CGPA: 9.55)

Mumbai, MH

PROJECTS

Up2Date - AI Newsletter

Feb 2025 - Mar 2025

- Curated a daily "Generative AI" digest (~100 links) to **50+ subscribers** by scraping Google Alert emails via Google Apps Script, extracting URLs with regex, and storing them in Firebase every 2 hrs.
- Orchestrated a **GitHub Actions workflow** to fetch links at 10 AM, scrape pages with Python, enrich content via Gemini 2.0 Flash, and send formatted newsletters.

Multimodal RAG

Jan 2025 - Feb 2025

- Built a **multimodal RAG pipeline** that structured PDFs using **Unstructured**, summarized images with Gemini 2.0 Flash, and embedded text and image summaries using Gemini Text Embedding 004.
- Achieved **87% accuracy** on 100 benchmark questions by designing the system to include images in generated answers, enhancing contextual understanding and user experience.

TECHNICAL SKILLS & OTHER PROFILES

- Languages:** Python, JavaScript, SQL
- Technical Skills:** Generative AI, Computer Vision, PyTorch, LangChain, LangGraph, Milvus, FastAPI, Ray framework, Git, Langfuse, Bash, Numpy, Pandas, Redis, MongoDB, Node.js, React.js, React-Native
- Other Profiles:** leetcode.com/shriyansnaik, medium.com/@shriyansnaik, github.com/shriyansnaik