SAUMYA LADDHA

📞 8989705185 ✉ saumyaladdha23@gmail.com in LinkedIn ⌂ GitHub ⊕ LeetCode

WORK EXPERIENCE

AI/ML Engineer
Profolo, Hyderabad, India   *April 2025 – December 2025*

- Achieved **40% higher** skill-match precision by architecting a **30K+ skill taxonomy** across domains using embeddings and LLM automation, deployed on AWS with Docker containers for scalable performance.
- Built an AI interview system using **LLMs and LangGraph**, generating **100% adaptive**, persona-driven interview.
- Improved evaluation **reliability by 20%** through designing an AI-powered proctoring and assessment platform leveraging computer vision for monitoring, NLP-based response analysis, and real-time scoring.
- Implemented a JD–Resume Matching System with **LlamaIndex and vector databases** for semantic ranking.

Software Engineer Intern
Katalyse AI, Remote   *January 2025 - March 2025*

- **Improved product discovery by 30%** by engineering multi-agent recommendation systems for Similar and Frequently Bought Together products using Weaviate, Ollama, and MongoDB, integrated with the Shopify API.
- Implemented an image-based search pipeline leveraging **vision-language models and semantic embeddings**.

AI Engineer Intern
DataToBiz, Remote   *April 2024 - September 2024*

- Enhanced testing efficiency by 30% by developing intelligent assessment modules featuring **RAG-based** question generation, real-time evaluation, and integration of text-to-image and image-to-text models.
- Developed and deployed full-stack services using **React.js and Flask API via PuTTY** for production.

EDUCATION

Rajiv Gandhi Institute of Petroleum Technology   CGPA: 8.4
Bachelor of Technology in Computer Science and Engineering   *July 2021 - May 2025*

PROJECTS

Multi Agent Travel RAG | *Python, Langchain, Qdrant, SQLite, Docker*   ⌂

- Engineered a multi-agent **LLM-based RAG** system using LangGraph for contextual, end-to-end travel assistance.
- Architected stateful routing and memory checkpointing in LangGraph, boosting response **accuracy by 30%**.
- Integrated **LangSmith observability** for tracing of agent reasoning, tool usage, and performance monitoring.

Neural Music Generation System | *Python, TensorFlow, NumPy & Pandas*   ⌂

- Built deep learning pipelines with LSTM, GRU, and GANs for symbolic music generation using **MIDI datasets**.
- Preprocessed and converted MIDI files using Music21 and Magenta for sequence learning and model training.
- Optimized **GAN-based** generation achieving 35% faster convergence and enhanced musical diversity over RNN.

TECHNICAL SKILLS

- **Core CS Concepts:** Data Structures & Algorithms, OOPs, DBMS, Computer Networks
- **AI & Data Science Concepts:** Machine Learning , AI, NLP, LLMs, Data Mining
- **AI/ML Frameworks:** LangChain, LangGraph, TensorFlow, PyTorch, Scikit-learn
- **Web & Backend Frameworks:** React.js, Node.js, Flask, FastAPI
- **Databases & Caching:** MongoDB, MySQL, Vector Databases
- **Languages:** C/C++, Python, JavaScript, SQL
- **DevOps & Cloud:** Docker, Git, AWS
- **Tools & Utilities:** Pandas, NumPy, OpenCV

ACHIEVEMENTS

- Awarded Best Intern at DataToBiz for outstanding performance and contributions during the internship.
- Certifications: Machine Learning by Andrew Ng (Coursera), Generative AI for Everyone (Coursera).