Shashank Rai

Kopaganj Mau, U.P

+91 7007871856 shashankjirai4@gmail.com GitHub Repository
linkedin Kaggle Repository Portfolio

Education

Indian Institute of Information Technology, Lucknow 2023 – 2025
Master of Technology(M.Tech) in Computer Science ***GPA: 8.25***

Krishna Engineering College, Ghaziabad 2018 – 2022
Bachelor of Technology(B.Tech) in Computer Science ***GPA: 7.56***

Projects

AI Chatbot Agent | LangChain, LangGraph
- Built a multi-tool LLM agent that uses Tavily (search), YouTube API, and weather data to answer real-time queries. Used LangGraph for flow control and LangChain tools to manage API interactions and agent memory.

Context-Aware Question Answering | LangChain, LangGraph, Multi-Stage RAG
- Built a multi-stage RAG system using LangChain, Chroma, and Google GenAI embeddings to retrieve context-aware answers. Incorporated chat history and custom prompts for multi-turn reasoning and refined question answering.

Fine-Tuned DeepSeek-8B for Medical QA | Python, PyTorch, Hugging Face, LoRA, Unsloth
- Fine-tuned 4-bit **DeepSeek-8B (R1 distilled)** using **LoRA** and **Unsloth** on a verified medical QA dataset, applying chain-of-thought prompting to improve reasoning accuracy; deployed with quantization on T4 GPUs for efficient inference.

Vanilla GAN for Image Generation | Python, PyTorch
- Implemented a Generative Adversarial Network (GAN) from scratch to generate synthetic images. Trained on image datasets using a standard generator–discriminator architecture and optimized with adversarial loss for realistic output generation.

Self-Drive Car | Computer Vision, YOLO
- Focused on road detection, lane segmentation, and steering angle prediction. Merged multiple datasets, trained three models, and built a real- time driving simulation UI using OpenCV, strengthening my deep learning and computer vision skills.

Work Experience

IIIT Lucknow | *Teaching Assistant* August 2023 – Present
- Working as a Teaching Assistant under university professors; responsible for conducting lab sessions, clarifying student doubts, and evaluating assignments and answer sheets for B.Tech and M.Sc. courses.

Skills

- **Machine Learning:** Supervised & Unsupervised Learning, Boosting.
- **Deep Learning:** ANN, RNN, LSTM, GRU, Transformers, CNN architecture(ResNet-50, InceptionNet, Imagenet, VGG-16, RCNN, Fast-RCNN, UNet, YOLO).
- **LLMs & GenAI:** LangChain, LangGraph, RAG, Multi-Agent Systems, Fine-Tuning (LoRA, QLoRA)
- **Programming & CS Fundamentals:** Python, C, OOP, DSA, Operating Systems, DBMS, SQL.
- **Libraries & Frameworks:** PyTorch, TensorFlow, Keras, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, FastAPI, Streamlit.
- **MLOps:** Modular Coding, Git, Logging, Exception Handling, Virtual Environments, Cookiecutter, DVC (local & AWS S3), MLflow (local, DagsHub, AWS), Experiment Tracking.

Publications (In Progress)

A Two-Phase Data-Driven Framework for Detecting Cyberattacks on Smart Charging Infrastructure for Electric Vehicles-ML-based anomaly detection and classification research.