

Ekaagra Gupta

✉ ekaagrag2006@gmail.com — 📞 +91-9773366055 — 📍 Jaipur, India

🌐 Portfolio — 🐙 GitHub — 🔗 LinkedIn

Education

Manipal University Jaipur	2024 – 2028
Bachelor of Technology in Computer Science and Engineering Specialization in Artificial Intelligence and Machine Learning	
Gyan Ashram School	Batch 2024
Senior Secondary (Class XII) Warren Academy School	Batch 2022
Secondary (Class X)	

Technical Skills

Artificial Intelligence & Machine Learning: Machine Learning, Deep Learning, Computer Vision, XAI

Generative AI & Retrieval Systems: LLMs, RAG, LangChain, LlamaIndex, Hugging Face Transformers, vLLM, FAISS, ChromaDB, BM25, ColBERT, HNSW

Geospatial AI & Image Processing: Satellite Image Processing, NDVI/NDWI Spectral Analysis, Geospatial Analysis, Remote Sensing Concepts

Programming Languages: Python, C++, C, SQL, javascript

Full-Stack Development: MongoDB, Express.js, React.js, Node.js (MERN Stack)

Libraries & Frameworks: PyTorch, Scikit-learn, OpenCV, Pandas, NumPy, Matplotlib

Course-work: OS, OOPs, DSA, Computer network, DBMS

Tools & Platforms: Git, GitHub, Linux, Docker, Jupyter Notebook, Google Colab, CUDA

Internship Experience

AI Intern *May 2026 - Present*
Malaviya National Institute of Technology Jaipur (MNIT)

- Led a modular Retrieval-Augmented Generation (RAG) pipeline with hybrid dense and sparse retrieval.
- Implemented document chunking, embedding generation, HNSW indexing, and BM25-based retrieval using Qwen embedding models.
- Integrated multi-stage reranking using Qwen Reranker, ColBERT, and Reciprocal Rank Fusion (RRF) for retrieval optimization.
- Optimized large-scale inference workflows using vLLM, Hugging Face Transformers, and PyTorch on NVIDIA GPUs.

Technologies: Python, PyTorch, Hugging Face Transformers, vLLM, Qwen Embeddings, ColBERT, BM25, HNSW, FAISS, RAG, Information Retrieval, Linux, Git, CUDA

AI/ML Intern *Mar 2026 – May 2026*
CoreGen

- Incorporated a **Super-Resolution Image Reconstruction** project for satellite imagery, building a multi-frame (3+1) deep learning pipeline that combines optical flow with Fourier-based neural warping to reconstruct high-resolution detail from low-resolution inputs ($64 \times 64 \rightarrow 256 \times 256$), with a custom GPU-optimized dataset and training pipeline for Earth observation applications.

- Operationalized a full-stack **MERN dashboard** (React frontend, Node/Express backend) to visualize and serve reconstruction outputs, connecting the model pipeline to a usable interface for reviewing results.
- Explored lightweight object detection (MobileNet SSD) and real-time anomaly detection for traffic analysis, evaluating ROS, TensorFlow Lite, and OpenCV for efficient on-device inference.

Technologies: Python, PyTorch, FAISS, ChromaDB, LangChain, LlamaIndex, HuggingFace Transformers, MobileNet SSD, ROS, TensorFlow Lite, OpenCV, MongoDB, Express.js, React.js, Node.js

Projects

MemoRAG: Memory-Augmented RAG Framework



LLM / RAG System

- Engineered a memory-augmented RAG pipeline that builds a compressed KV-cache representation of long documents before retrieval, enabling globally-aware query rewriting and surrogate-question generation for multi-hop and summarization queries that standard chunk-based RAG handles poorly.
- Integrated a dense retrieval layer using FAISS and BGE-M3 embeddings, paired with a lightweight variant (Qwen2.5-1.5B) for lower-resource deployment.
- Evaluated pipeline performance on long-context QA and summarization tasks, comparing standard RAG against the memory-augmented approach for retrieval accuracy and answer completeness.

Technologies: Python, HuggingFace Transformers, FAISS, LangChain, LLM APIs, Retrieval Systems

SEVAS: Satellite-based Environmental Violation Analysis System



AI/ML System

- Orchestrated a satellite image analysis system combining a custom U-Net segmentation model with **NDVI/NDWI spectral indices** to detect sand mining, land encroachment, and vegetation loss; U-Net achieved 73% validation on a 4-class pixel segmentation task (mining, vegetation, water, background).
- Leveraged Gemini Vision API for natural-language violation descriptions alongside the segmentation pipeline, combining computer vision with vision-language analysis for richer detection output.
- Designed a severity-scoring pipeline to prioritize violation reports for field inspection, reducing manual triage effort through automated ranking.
- Devised a Flask REST API supporting single and batch satellite image analysis (up to 50MB per file) across JPG, PNG, TIFF, and GeoTIFF formats.

Technologies: Python, TensorFlow, Keras, OpenCV, NumPy, scikit-learn, Flask, REST APIs, Geospatial Analysis, Gemini Vision API

GETHER: Generative Emission Temporal Hybrid Explainable Regressor



GitHub

AI/ML System

- Formulated a temporal deep learning pipeline for AQI forecasting using a stacked LSTM architecture (128→64 units) with engineered lag and rolling-statistic features, benchmarked against linear regression and random forest baselines.
- Engineered a causal discovery module using Granger causality and dynamic causal graph construction to surface relationships between pollutants (PM2.5, NO2, SO2) and AQI

variations.

- crafted a SHAP-based explainability layer to rank pollutant feature importance and interpret individual model predictions.
- Configured a counterfactual policy simulator to estimate AQI impact under hypothetical emission-reduction scenarios, deployed via an interactive Streamlit dashboard.

Technologies: Python, TensorFlow, Scikit-learn, SHAP, Streamlit, Pandas, NumPy, Matplotlib

Patent application filed (under examination)

Leadership and Achievements

Google Developer Groups — Technical Co-Lead May 2026 – Present

- Led technical planning and execution of developer-focused workshops on web, cloud, and AI/ML technologies — *300+ attendee / 3 workshops* .

IEEE CIS — Joint Head, Sponsorship and Curation Jan 2026 – Present

- Led sponsorship acquisition for technical events — *150+ sponsors secured* and contributed to the webmaster team.

DevForge — Technical Joint Head Jul 2026 – Present

- Oversaw full-stack development across *24+* projects, focusing on modular architecture and performance optimization.

GSSoC'25 — Open Source Contributor Jul 2025

- Contributed AI/ML and web features across *90+* merged pull requests following structured open-source workflows.

Certifications

- **Prompt Engineering & Programming with OpenAI** — Columbia University
Credential Link
- **Essentials in Generative AI** — Microsoft & LinkedIn
Credential Link
- **Generative AI Studio** — Google Cloud
Credential Link
- **Generative AI: The Evolution of Thoughtful Online Search** — LinkedIn Learning
Credential Link

Publications

- **Mapping the Latency-Explainability-Accuracy Pareto Frontier: A Systematic Benchmark and Framework for Attribution Methods in Real-Time Satellite Anomaly Detection**

Manuscript in preparation, targeting IEEE Transactions on Geoscience and Remote Sensing