Ratnesh Dagli

dagliratnesh05@gmail.com | +91 70211 33954 | linkedin.com/in/ratneshdagli | github.com/ratneshdagli

Education

Dwarkadas J. Sanghvi College of Engineering Aug 2024 – Aug 2028

B.Tech in Computer Engineering

Experience

DHL Express India | *Project Trainee (IT)* Jun 2026 – Aug 2026

- Conducted an in-depth technical analysis of enterprise-scale local applications including EMM, BCA, SCM, DIET, and KYC to identify critical integration points for artificial intelligence and automation workflows.
- Formulated detailed, DPDP-compliant architectural proposals to bridge the gap between legacy systems and modern machine learning capabilities, establishing clear pathways to reduce manual processing time.

Zeex AI | *Artificial Intelligence Intern* Apr 2026 – May 2026

- Implemented a highly scalable fraud analytics engine for trilingual audio calls on AWS, leveraging FastAPI and Pyannote diarization to reliably flag mimicry, forced answers, and anomalous speech patterns.
- Fine-tuned the Indic Conformer 600M acoustic model, dropping Word Error Rate from 0.80 to 0.36, and deployed fault-tolerant inference pipelines integrating EC2, S3, SageMaker, ASGs, SQS, and DLQs.

Projects

Genie - Personalized AI Assistant | *Flutter, Qwen3-4B, llama.cpp, LoRA*

- Implemented an offline-first, GPU-accelerated Android agent powered by a LoRA fine-tuned Qwen3-4B via llama.cpp, integrating an Adaptive Compute Allocator and orchestration, cutting latency to <0.7s.
- Developed a grammar-constrained ReAct loop with deterministic JSON decoding and an 18-tool capability layer, alongside a 3-layer semantic memory system with CLIP search, shrinking prompt overhead by >70%.

GenAI Research Suite (RAG + Agent) | *LangChain, Streamlit, FAISS* View Project

- Built a high-performance conversational RAG system enabling context-aware Q&A over extensive private document repositories, using FAISS databases for scalable semantic search and sub-second retrieval.
- Integrated an autonomous research agent orchestrating multi-step queries across ArXiv, Wikipedia, and DuckDuckGo APIs to synthesize technical data and generate comprehensive multi-domain reports.

Blogs & Articles

How to improve your Research Paper RAG Chatbot | Medium Oct 2024

- Authored a deep-dive on RAG pipelines, covering hybrid search and re-ranking optimization, achieving 150+ views. **[Read Article]**

Extra-Curricular Activities

Research Mentee, DJS ACM Student Chapter Oct 2024 – Present

- Conducting preliminary research for a technical paper on Generative AI and Healthcare architectures.

AI Mentee, DJS Code AI Aug 2024 – Present

- Studying ML model architectures and methodologies while organizing student community hackathons.

Skills

Languages & Core: Python, Java, Dart, SQL, C, C++, Git, Docker
AI, ML & Audio: TensorFlow, Scikit-learn, LangChain, FAISS, RAG, LoRA, HuggingFace, Pyannote, Speaker Diarization
Cloud & AWS: Amazon Web Services (AWS), EC2, S3, SQS, Dead Letter Queues (DLQ), Auto Scaling Groups (ASG), SageMaker
Backend & DBs: FastAPI, Flask, SQLAlchemy, MySQL, MongoDB, Firebase, Vector Databases, Async Processing
Frontend & Mobile: Flutter, Streamlit