Ratul Dash

ratuldash.dev | ratuldash12@gmail.com | linkedin.com/in/ratul-dash | github.com/ratul-d

Education

Vellore Institute of Technology — Amaravati, Andhra Pradesh
B.Tech in Computer Science and Engineering (Spc. Data Analytics) | CGPA: 8.62/10 — *Sep. 2023 – June 2027*

Experience

AI Engineer Intern — Mar. 2026 – Present
TailorTalk.ai

- Improved AI sales-agent spreadsheet task accuracy from 60% to 87% by building evaluation pipelines and redesigning agent workflows using custom task-level metrics.
- Developed an asynchronous image-retrieval pipeline using Gemini Embedding 2 and Milvus for image-vector similarity search; benchmarked 10+ retrieval strategies and achieved 98% exact-match accuracy.
- Built an asynchronous LLM evaluation pipeline to benchmark intent classification, lead status, escalation, blocking, and attribute extraction across 7,000+ customer interactions, enabling automated model evaluation and regression testing.
- Built a multi-agent reporting workflow that analyzes 3,000+ lead conversations and CRM records per client agent to automatically generate weekly reports containing customer insights, lead-quality metrics, conversion trends, and lead-source attribution.
- Implemented an agent-facing escalation tool that routes complex customer interactions to human operators via WhatsApp, with conversation summaries and lead details.

Achievements

2nd Place Winner | Udhgam 24Hr Hackathon, Woxsen University — Feb. 2025

- Built **NexTrail**, an ML-driven career guidance platform using Flask-based backend serving inference APIs for a **custom neural network model** to map user skill embeddings to relevant job roles, and generating personalized learning roadmaps via LLM-driven flow visualization; secured **2nd place among 130+ teams**.

Projects

AptivHire – Recruiting Multi-Agent System — Link

- Architected a multi-agent recruiting system using LLM agents for CV parsing, job-description parsing, candidate matching, and interview email generation with structured Pydantic outputs.
- Engineered LLM-based CV and job-description parsing pipelines that transform unstructured documents and text into structured candidate and job profiles for downstream matching.
- Built the FastAPI backend with PostgreSQL and SQLAlchemy to orchestrate agent workflows, candidate-job matching, interview scheduling, and authenticated recruiter APIs.

Semantic Graph RAG — Link

- Implemented a semantic graph RAG pipeline using semantic chunking, knowledge graph construction, and hierarchical community summaries for document retrieval.
- Built local and global graph-based retrieval with embedding-based ranking and local LLM answer generation using Ollama.

Minimal Neural Network Framework — Link

- Implemented a scalar reverse-mode autodiff engine with computation-graph traversal and gradient-based backpropagation from scratch.
- Built an MLP framework with configurable activations, He/Xavier initialization, loss functions, and SGD optimization.

Technical Skills

Languages: Python, SQL
Backend: FastAPI, Flask, REST APIs, Pydantic
AI/ML: PyTorch, TensorFlow, Sentence Transformers, spaCy
LLM/GenAI: LangGraph, LangChain, PydanticAI
Cloud/Platforms: AWS, Amazon Bedrock, GCP Vertex AI, Google AI Studio, IBM watsonx
Databases: PostgreSQL, MongoDB, Redis, Milvus
Tools: Git, Docker