# Sujeet Gund

+91-7620420850 | sujeetgund@gmail.com | sujeetgund.in | github.com/sujeetgund | linkedin.com/in/sujeetgund

## SUMMARY

AI Engineer with 3+ years building production-grade Generative AI and ML systems, specializing in RAG pipelines, LLM orchestration, and agentic AI architectures. Delivered real-world systems: a Human-in-the-Loop email automation agent (LangGraph), an explainable legal AI copilot with structured LLM output contracts, and an anomaly detection platform achieving 94.16% accuracy on telecom data. Proficient across the full AI stack including prompt engineering, vector retrieval (FAISS, pgvector, Pinecone), and cloud deployment on AWS and GCP, with a strong focus on explainability, reliability, and production readiness.

## EDUCATION

**VIT Bhopal University**                                                                                              Bhopal, India
*Integrated M.Tech in Artificial Intelligence, CGPA: 9.24/10.0*                              *Expected Graduation: 2028*

## TECHNICAL SKILLS

**Programming**: Python, SQL, TypeScript
**LLMs & Agentic AI**: LangChain, LangGraph, LangSmith, OpenAI API, Gemini API, Prompt Engineering, RAG, HITL Agents, FAISS, Pinecone, pgvector, Hybrid Search
**ML/AI Frameworks**: HuggingFace Transformers, scikit-learn, TensorFlow, PyTorch, Pydantic v2
**Backend & Data**: FastAPI, PostgreSQL, MongoDB, Redis, Docker
**Cloud & DevOps**: AWS (EC2, S3, Lambda), GCP (Cloud Run, Vertex AI), Git

## PROJECTS

**MailMind: Agentic Email Processing System** | *LangGraph, FastAPI, Next.js, PostgreSQL, FAISS*          GitHub

- Built a production-grade, event-driven email automation agent using LangGraph StateGraph to classify intent and autonomously execute actions (reply, label, forward, escalate) via Resend webhooks in real time
- Engineered HITL checkpoints with interrupt-based pause/resume and persistent state, ensuring human approval for low-confidence or high-stakes decisions before execution
- Integrated FAISS-powered RAG for context-grounded auto-replies with a FastAPI + SSE backend and Next.js dashboard for live monitoring and full pipeline auditability

**Legal Sahayak: Explainable RAG Legal Copilot for India** | *LangChain, FastAPI, Next.js, FAISS*          GitHub

- Built a FAISS-powered hybrid retrieval pipeline spanning 6+ Indian legal domains using dense HuggingFace embeddings with semantic + keyword search, improving context relevance over naive RAG baselines
- Enforced structured LLM output contracts (answer, confidence, legal_references, action_plan) via Pydantic v2 schema validation with demographic-aware prompt orchestration for personalized legal guidance
- Delivered a Dockerized, production-ready system with traceable source references and confidence scoring, prioritizing AI transparency and trust through a secure bearer-token API and Next.js 15 frontend

**IPDR Graph Engine** | *Python, FastAPI, Machine Learning, MongoDB, GCP, Next.js*          GitHub

- Architected an ML-powered telecom investigation platform ingesting large-scale IPDR logs (CSV/JSON/Excel), automatically constructing communication graphs with intelligent A→B party extraction and timeline reconstruction
- Trained custom anomaly detection models achieving **94.16% accuracy** in identifying suspicious communication sessions, enabling rapid threat detection across thousands of network records per run
- Deployed end-to-end on GCP Cloud Run (FastAPI) and Vercel (Next.js) with MongoDB for scalable persistence, automated audit report generation, and interactive graph visualizations for security analysts

## CERTIFICATIONS

**Complete Data Science, Machine Learning, DL, NLP Bootcamp** - Udemy (Certificate, December 2025)
**Applied Machine Learning in Python** - University of Michigan (Certificate, December 2024)
**Supervised Machine Learning: Regression and Classification** - DeepLearning.AI (Certificate, July 2023)
**AWS Cloud Essentials** - Amazon Web Services (Badge, June 2023)