# SAI ANEESH GANTI

gantisaianeesh@gmail.com | Portfolio
GitHub: Saianeesh2003 | LinkedIn: sai-aneesh-933a95230

## OBJECTIVE

AI Engineer specializing in agentic systems and RAG pipelines. Experienced building production AI applications from scratch and deploying them to real users. Passionate about building reliable, execution-focused AI systems that solve real business problems. Seeking to contribute to early-stage startup building autonomous AI agents.

## TECHNICAL SKILLS

**Programming:** Python , SQL
**AI & GenAI:** Agentic Workflows, RAG Pipelines, Tool Calling, LLM Integration (OpenAI/Claude), Prompt Orchestration, Vector Databases (Qdrant)
**Backend Development:** Currently learning FastAPI
**Python Libraries:** LangGraph, Pandas, NumPy, Scikit-learn, Streamlit
**ML Fundamentals:** Classification, Regression, Model Evaluation, Data Preprocessing
**Tools & Infrastructure:** Git,Jupyter Notebook, Cloud deployment (Streamlit Cloud)

## PROJECTS

**AI Agent with Autonomous Tool Calling** — GitHub — Live Demo

- Built production-ready agentic AI system with autonomous tool selection and multi-step execution workflows
- Implemented tool-calling framework enabling agent to reason about tool selection, parameter extraction, and task orchestration across 5+ integrated tools (weather API, file I/O, code generation)
- Designed prompt orchestration and fallback strategies to handle edge cases and tool failures reliably
- Shipped from idea to production in under 2 weeks, currently handling real user sessions

*Technologies: Python, Claude API, LangChain, Streamlit — Users: 50+ active testers*

**RAG System with Vector Database** — GitHub — Live Demo

- Built end-to-end RAG pipeline for semantic document search and context-aware question answering
- Implemented full pipeline: PDF processing → chunking → embedding generation → vector storage → retrieval
- Used Qdrant vector database for efficient semantic search and context management
- Designed retrieval strategies and prompt engineering for accurate, contextual LLM responses
- Shipped working MVP in 12 days, deployed for real-world document querying

*Technologies: Python, LangChain, Qdrant, Claude API, Google Embeddings, Streamlit*

**Customer Churn Prediction System** — GitHub

- Built ML classification pipeline comparing 3 algorithms, achieving 85%+ accuracy through systematic evaluation
- Implemented full data preprocessing pipeline with feature engineering and proper train/test methodology
- Used multiple evaluation metrics (Accuracy, Precision, Recall, F1, AUC-ROC) for robust model assessment

*Technologies: Python, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn*

## EDUCATION

**Bachelor of Engineering - Electronics and Communication Engineering** July 2025
Osmania University, Hyderabad GPA: 8.09/10

## ACHIEVEMENTS & ADDITIONAL INFORMATION

**Key Strengths:**

- **Ship Fast:** Built and deployed 2 production AI applications in ¡2 weeks each
- **Autonomous Learning:** Self-taught LangChain, vector databases, and agentic AI concepts from scratch
- **Execution-Focused:** Comfortable shipping imperfect MVPs and iterating based on user feedback
- **Ownership Mindset:** All projects built independently from concept to deployment

**Notable Achievement:**

- Selected for national MeitY Chips-to-Startup (C2S) Program from 10,000+ applicants (Top 0.1%)

**Current Focus:** Learning FastAPI for production AI backend systems, exploring multi-agent architectures