

Ajinkya Mawal

ajinkyamawal.iitb@gmail.com | +91 86981 02164 | linkedin.com/in/ajinkya-mawal | github.com/a-j-i-n-k-y-a | Mumbai, India

FDE shipping production agentic systems inside enterprise BFSI environments. Experience across multi-agent workflows, RAG, evaluation and observability, and owning delivery end to end, from scoping with client stakeholders through on-prem rollout and iteration. Worked across on-prem cloud (AWS) and baremetal deployments

SKILLS

Harness	Multi-Agent Orchestration, Tool and Function Calling, Context Engineering, Deterministic Guardrails and Fallback Routing, Evaluation Harnesses, Regression Gating, Tracing and Observability
AI / ML	Agentic AI, RAG, LLM Evaluation, Fine-Tuning, PyTorch, Hugging Face Transformers, LangChain, LangSmith, Vector Databases
Engineering	Python, SQL, FastAPI, PostgreSQL, Redis, Docker, Kubernetes, AWS, Linux, Git
Delivery	On-Prem and Hybrid Deployment, Legacy System Integration, Client Scoping, Stakeholder Alignment

EXPERIENCE

Forward Deployed Engineer & AI Engineer, OnFinance AI – Mumbai, India May 2026 – Present

- Own enterprise agentic AI deployments end to end for BFSI institutions, from scoping with client stakeholders through production rollout inside their environments
- Build and maintain the agent harness behind those deployments: orchestration loops, tool interfaces, context assembly, and deterministic fallback paths that keep outputs auditable
- Design multi-agent workflows automating regulatory compliance audits, financial risk modeling, and accounting operations
- Integrate on-prem and hybrid AI pipelines into legacy core banking systems alongside client engineering, security, and risk teams
- Run continuous evaluation and tracing against live deployments, gating releases on regression checks and holding latency and reliability budgets in production

AI Product Intern, Karta – Bengaluru, India Jun 2025 – Aug 2025

- Developed an AI observability platform using LangSmith traces and agent run-evaluations to monitor token latency, flag regressions, and cut customer escalations
- Built a prompt-based automation engine and standardized SOPs, reducing customer onboarding time by 60% and improving CX accuracy by 10%
- Optimized multi-step agent workflows and fallback routing, scaling support resolution throughput

Data Science Intern, Swiggy – Bengaluru, India Apr 2025 – Jun 2025

- Developed an enterprise RAG chatbot over internal operational knowledge bases for self-serve retrieval across teams
- Architected an automated candidate-screening pipeline for HR, modernizing resume ingestion and recruitment

PROJECTS

Protein Language Models for Antimicrobial Peptide Design *ESM-2, PyTorch, Hugging Face*
Bachelor's Thesis, IIT Bombay Jan 2025 – Dec 2025

- Fine-tuned a 35M-parameter ESM-2 protein language model via masked language modeling to design novel beta-hairpin antimicrobial peptides, improving training loss from 4.93 to 1.05
- Engineered multi-stage biophysical screening filters on net charge, amphipathicity, and structural propensity, achieving a 91.4% validation pass rate across 180+ candidates
- Built data curation pipelines to clean contaminated structural data and generate synthetic training sequences

IoT Computer Vision Pipeline for Automated Waste Segregation *CNNs, Vision Transformers, IoT*
Machine Learning Engineer, Team Zero Waste, IIT Bombay Nov 2024 – Apr 2025

- Architected an IoT-integrated computer vision pipeline using CNNs for real-time waste classification and automated physical segregation
- Built low-latency ingestion pipelines across sensors and microcontrollers for fault-tolerant edge-to-cloud telemetry, collaborating across mechanical, firmware, and software teams

EDUCATION

Indian Institute of Technology Bombay – B.Tech, Metallurgical Engg. & Materials Science Oct 2022 – Apr 2026

- Granted L2 funding by IDEAS, IIT Bombay; Convener, SilverScreen (Film and Media Club); Team Lead, ShARE IITB