

Atharva Deshmukh

New York, NY | 585-910-6910 | atharvad061199@gmail.com | [in/atharva-s-deshmukh](https://www.linkedin.com/in/atharva-s-deshmukh) | [GitHub](https://github.com)

SUMMARY

AI/ML Engineer with 3+ years of experience building production RAG, agentic AI, and clinical decision-support systems. At CVS Health, led multimodal RAG and LLM evaluation frameworks powering 10K+ daily predictions at sub-500ms latency. Delivered a full-stack AI lead-generation platform in 30 days. Skilled in LangChain, PyTorch, LLM fine-tuning, FastAPI, and MLOps across GCP, and Azure.

EXPERIENCE

ContextQA

Feb 2026 - Mar 2026

Forward Deployed AI Engineer (Contract)

Remote

GenAI Product Engineering · Client-Facing AI Delivery · Agentic Workflows · Full-Stack

- Built a React + TypeScript CRM from scratch to capture, qualify, and manage 115+ AI-sourced leads, reducing sales rep lead processing time by 60% through a unified interface for contacts, status tracking, and follow-up workflows
- Generated 115 leads and 18 qualified prospects (16% qualification rate) within 2 weeks of deployment, validating the agentic pipeline's production readiness and demonstrating rapid POC-to-production delivery
- Collaborated with sales and GTM teams to demonstrate technical value to prospects, owning the full lifecycle from problem scoping through production deployment

CVS Health

Jan 2025 - Feb 2026

GenAI Engineer

New York City, NY

Clinical AI · Healthcare · Pharmacy Workflows · HIPAA-Compliant Systems

- Improved medication non-adherence recall from 71% to 84% by training gradient-boosted risk prediction models on EHR and pharmacy claims features, directly improving care for 50K patients
- Architected retrieval-augmented generation clinical decision support using pgvector and semantic search with reranking, achieving 95% guideline retrieval accuracy for pharmacists
- Built multimodal RAG pipelines combining patient charts, medication history, and lab results to auto-generate contextual clinical summaries, cutting documentation time by 45%
- Implemented LLM evaluation pipelines using RAGAS and human-in-the-loop assessment, measuring faithfulness, relevance, and hallucination rates across clinical summaries to ensure responsible AI outputs
- Deployed FastAPI inference services with guardrails on Azure AI Foundry, achieving <500ms latency and observability monitoring across 10K daily pharmacy workflow predictions
- Established MLflow MLOps pipelines with A/B testing and experiment tracking, mentoring junior engineers and reducing model iteration cycles by 30%

VR Digital Solution

Mar 2021 - Jul 2023

Machine Learning Engineer

Mumbai, India

Marketing Analytics · Computer Vision · eCommerce

- Built churn prediction and lead scoring models with XGBoost across 2M+ customer records, boosting marketing ROI by 17%
- Designed a computer vision pipeline (PyTorch + YOLO) for automated defect detection – reducing manual QC effort by 40% and saving \$200K annually
- Implemented MLflow MLOps standards improving reproducibility and reducing deployment errors by 25%

TECHNICAL SKILLS

- **GenAI & LLMs:** LangChain, LangGraph, RAG, OpenAI API, Transformers, Multimodal Agents, LLM Evaluation (RAGAS, BLEU, ROUGE, Human Eval), Prompt Engineering, Fine-tuning (QLoRA/LoRA), Guardrails, Responsible AI, Hugging Face, n8n, CrewAI
- **ML & DL:** PyTorch, TensorFlow, Scikit-learn, XGBoost, LightGBM, NLP (NER, Sentiment Analysis, Summarization), CNNs, RNNs, SHAP, A/B Testing, Feature Engineering
- **MLOps & Cloud:** AWS SageMaker, GCP Vertex AI, Azure ML & AI Foundry, MLflow, Docker, Kubernetes, Docker Compose, Airflow, Terraform, Observability/Monitoring, CI/CD, Git
- **Full-Stack & Infra:** FastAPI, Bun, Node.js, React, Streamlit, Gradio, REST APIs, Docker Orchestration, Nginx
- **Databases & Vector Stores:** PostgreSQL, pgvector, MongoDB, Redis, FAISS, Pinecone, ChromaDB, MySQL
- **Languages:** Python, TypeScript, SQL, R, Bash, JavaScript, Java

EDUCATION

Rochester Institute of Technology

Aug 2023 - May 2025

M.S, Information Technology & Analytics

Rochester, New York

University of Mumbai

Jul 2017 - May 2021

B.E, Computer Engineering

Mumbai, India

CERTIFICATIONS

- **Microsoft Certified: Azure Data Scientist Associate**