

MD RAKIBUL HASAN NAYM

Applied AI Engineer — Production Computer Vision · Generative AI (RAG & AI Agents) · GPU-Optimized MLOps

Dhaka, Bangladesh | +880 163 883 0165 | [E-mail](#) | [LinkedIn](#) | [GitHub](#) | [Hugging Face](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

Applied AI Engineer with 3+ years designing, building, and deploying production AI at national scale. Owns applied-AI delivery at HawkEyes Digital Monitoring across retail-execution, manufacturing, and surveillance product lines — architecting and shipping computer-vision systems that serve 2M+ retail outlets at ~99% accuracy for enterprise FMCG and mobile-financial-services clients (BAT, Unilever, Nagad). Pairs deep computer vision (YOLO detection/segmentation, OpenCV) with current generative-AI delivery — production RAG (vector search, LangGraph), self-correcting multi-agent systems, and real-time voice agents — backed by GPU-optimized async inference, containerized MLOps, and self-hosted Linux GPU infrastructure built and run in-house. Currently developing a custom Bangladeshi-accent Bangla text-to-speech model (VITS + grapheme frontend) to power production AI voice agents for inbound/outbound call automation. Maintains 50+ Hugging Face models, with equal focus on model accuracy and clean, observable production engineering.

CAREER HIGHLIGHTS

- 2M+ retail outlets served by a production computer-vision platform running at ~99% detection accuracy.
- Enterprise clients: BAT, Unilever, and Nagad — FMCG and mobile-financial-services deployments.
- 55+ engineering repositories spanning 8 AI domains — computer vision, GenAI/RAG, agents, speech, OCR, MLOps.
- 50+ Hugging Face model repositories maintained across personal and organizational accounts.
- Designed, built, and administer 2 on-premises Ubuntu GPU servers — an in-house alternative to rented cloud GPUs.
- 5.0/5.0 client rating across delivered Upwork machine-learning and computer-vision contracts.
- Owns AI roadmap and cross-functional delivery; mentors a growing AI engineering team.

CORE TECHNICAL SKILLS

Programming: Python, C++, SQL

Machine Learning: PyTorch, TensorFlow, scikit-learn, NumPy, pandas, CNNs, model fine-tuning & evaluation

Computer Vision: Ultralytics YOLO (detection & segmentation), OpenCV, object detection & counting, POSM/planogram/share-of-shelf analytics, face recognition (dlib, face_recognition), MediaPipe

Generative AI / LLMs: LangChain, LangGraph, Retrieval-Augmented Generation (RAG), vector search (MongoDB Atlas Vector Search, FAISS), embeddings (sentence-transformers, E5), Hugging Face Transformers, AI agents, tool-calling (MCP), prompt engineering

Speech & OCR: STT/TTS (faster-whisper, edge-tts, Coqui VITS), Bangla TTS (VITS + grapheme frontend), Bangla OCR/ICR, Google Cloud Vision & Speech, Tesseract, AssemblyAI

Backend & Serving: FastAPI, asyncio, Hypercorn/Uvicorn, WebRTC/WebSockets, REST APIs, httpx/aiohttp

Data & Storage: MongoDB, PostgreSQL, Redis, Label Studio, Roboflow

Cloud & MLOps: Google Cloud Platform (GCP), Vast.ai GPU orchestration, Docker & Docker Compose, Git/Git LFS, GPU optimization, model deployment

Systems & Infrastructure: Linux (Arch daily driver, 3+ years; Ubuntu, Fedora), on-prem GPU server build & administration, self-hosted training/inference, Bash scripting

Leadership: AI roadmap ownership, team mentoring, cross-functional stakeholder management

PROFESSIONAL EXPERIENCE

Applied AI Engineer — HawkEyes Digital Monitoring Ltd., Dhaka, Bangladesh **Mar 2023 – Present**

Owns applied-AI delivery end to end — research through deployment and monitoring — across retail-execution, manufacturing, and surveillance product lines; mentors a growing AI team.

- Built and scaled the flagship retail-execution computer-vision platform serving 2M+ outlets at ~99% accuracy, routing each outlet image to campaign-specific YOLO detection/segmentation models behind scalable async FastAPI services (BAT program: 397 commits over 18 months).
- Delivered multi-client production detection systems for BAT, Unilever (10+ versioned models with cross-model reconciliation), Malaysia, Square, and Nagad — covering POSM detection, planogram compliance, share-of-shelf, and competitor detection.
- Architected production RAG and multi-agent systems (MongoDB Atlas Vector Search, LangGraph, self-correcting Planner→Executor→Critic loops with groundedness checks and SSE streaming) for enterprise data Q&A, including strict-grounding RAG over Bangla PDFs (FAISS).
- Engineered a real-time conversational voice agent (WebRTC/WebSocket streaming, VAD→STT→LangGraph→TTS, per-session memory, MCP tool-calling), reducing perceived latency through parallel TTS and GPU-batched STT.

- Developing a custom Bangladeshi-accent Bangla text-to-speech model (VITS + grapheme technique) for production AI agent-driven inbound/outbound call systems — hospital appointment automation and sales/support agents.
- Built and administer two on-premises Ubuntu GPU servers end to end (provisioning, setup, deployment, maintenance) that host most of the team's AI services and training workloads, cutting reliance on rented cloud GPUs.
- Delivered low-resource Bangla OCR/ICR (custom PyTorch CNN + CTC) and biometric identity/NID face-verification and safety/surveillance systems (fire, accident, methane-leak, face matching); maintains 50+ personal and 11+ organizational Hugging Face model repositories.

Freelance ML / Computer Vision Engineer — Upwork, Remote2022

- Delivered YOLOv5-based computer-vision solutions and project supervision for international clients; maintained a 5.0/5.0 client rating across completed contracts (image processing, deep learning, Python, C++).

Data Analyst — Quantanite, Dhaka, BangladeshJan 2023 – Mar 2023

- Executed large-scale data analysis, quality assurance, and operational reporting to support business-intelligence initiatives across client accounts.

KEY PROJECTS

- **Custom Bangladeshi-Accent Bangla TTS for AI Call Agents (current)** — Building a Bangla text-to-speech model with a grapheme-based frontend to produce a natural Bangladeshi accent for production AI call automation (hospital scheduling, sales/support agents). *Stack: Python, Coqui TTS (VITS), PyTorch/torchaudio, librosa*
- **Real-Time Conversational Voice Agent** — Browser-based low-latency voice assistant — WebRTC audio → VAD → STT → LangGraph agent → TTS, with per-session memory and MCP tool-calling; foundation for telephony call automation. *Stack: Python, FastAPI, WebRTC, LangGraph, faster-whisper, edge-tts*
- **Retail-Execution Intelligence Platform (BAT · Unilever · Nagad)** — Central inference service routing outlet images to campaign-specific YOLO models for POSM, planogram, share-of-shelf, and competitor detection — ~99% accuracy across 2M+ outlets. *Stack: Python, FastAPI, PyTorch, Ultralytics YOLO, async httpx, Hypercorn*
- **Agentic RAG over MongoDB** — Planner→Executor→Critic→Memory agent that generates queries/vector searches, verifies groundedness, and retries within bounds; SSE streaming with full observability. Companion Bangla-PDF RAG with strict grounding. *Stack: Python, FastAPI, MongoDB Atlas Vector Search, sentence-transformers, LangGraph, FAISS*
- **On-Prem AI Server Infrastructure** — Designed, built, and maintain two self-hosted Ubuntu GPU servers that host production AI services and model-training jobs — an in-house alternative to cloud GPU rental. *Stack: Ubuntu Server, Linux (Arch), NVIDIA GPU, Docker, Bash*
- **Bangla OCR / ICR Engine** — End-to-end custom intelligent-character-recognition pipeline for low-resource Bangla script (CNN feature extractor + CTC decoding, synthetic-data generation), deployed as a REST API. *Stack: Python, PyTorch, OpenCV, scikit-learn*
- **Biometric Identity & NID Verification** — REST service matching a selfie against a National ID photo via face embeddings, returning a verification decision — foundation for KYC-style identity checks. *Stack: Python, FastAPI, face_recognition / dlib*

EDUCATION

B.Sc. in Electrical & Electronic Engineering — Bangladesh University of Business and Technology (BUBT) · 2021 · CGPA 3.75 / 4.00

OPEN SOURCE & RECOGNITION

- Maintainer of 50+ personal and 11+ organizational Hugging Face model repositories spanning computer vision, OCR, LLM fine-tuning, and deployment workflows.
- Author of an open-source OpenCV computer-vision tutorial series and production-ready model cards / inference scripts adopted by the wider ML community.
- 5.0/5.0 client rating on delivered Upwork machine-learning and computer-vision contracts.

LANGUAGES

Bangla (native) · English (professional working proficiency)