

SUMMARY

AI engineer who takes models from prototype to production in steel plants and naval deployments: edge computer vision on NVIDIA Jetson, GraphRAG platforms, on-prem GPU clusters on DGX Spark, and the backends, data pipelines and IoT systems around them. Cut a production genetic algorithm's runtime from 5 hours to under 2 minutes.

EXPERIENCE

QED Analytics

Senior AI Engineer

Delhi, India · Hybrid

Jul 2026 – Present

- Clustered two NVIDIA DGX Spark nodes for on-prem model hosting, low-latency inference and air-gapped workloads.
- Built a GenAI platform (FastAPI, React 19, Neo4j, ChromaDB) with hybrid GraphRAG, SSE streaming and multimodal OCR.
- Revived a critical offline platform on RHEL and Docker Swarm: upgraded 15+ microservices and led a PostgreSQL disaster recovery, salvaging data from a replica, rebuilding the corrupted primary and completing an end-to-end migration.

AI Engineer

Nov 2024 – Jul 2026

- Vectorized a coal-blend genetic algorithm's population scoring, cutting runtime from 5 hours to under 2 minutes (150x).
- Built a thermal-video digital twin that measures hot sinter height so operators can tune fuel rates in real time.
- Deployed multi-regression models predicting Internal and Blast Furnace Return Fines from live PI Server process data.
- Deployed naval ship identification on Jetson AGX Orin (YOLO + hull-number OCR) at 92% accuracy in operational conditions.

Krishworks Technology & Research Labs

System Software Engineer

Bengaluru, India

Mar 2024 – Dec 2024

- Architected an AWS IoT solution (IoT Core, Lambda, RDS, S3) for device onboarding and over-the-air firmware updates.
- Engineered Raspberry Pi telematics units combining 4G, GPS and OBD-II for real-time vehicle analytics.
- Built hotspot/coldspot detection with a U-Net segmentation model and a React + Tailwind CSS front end.
- Implemented a Samsung SmartThings Edge Driver with custom webhook integration.

County of Simcoe

Project Lead

Ontario, Canada

May 2023 – Aug 2023

- Led a team of 3 building an emergency-dispatch simulator with LSTM call-volume forecasting and Voronoi coverage zones.

Krishworks Technology & Research Labs

System Software Engineer – Trainee

Bengaluru, India

Jun 2022 – Dec 2022

- Cut LMS app load times from 6 seconds to milliseconds by parallelizing data fetching.
- Led the AWS Amplify integration for the Dorabhai learning app, which reached 5,000+ Play Store downloads.
- Wrote a custom iOS BLE library that doubled command throughput in the MYRO PaintBot React Native robot controller.
- Built a full-stack property-management CRM with React, Flutter, Firebase, Django and Twilio.

PROJECTS

**Compose Keyboard** (Kotlin, PyTorch, 2026): Android keyboard with a neural swipe decoder, a 634K-parameter TCN trained on ~900K real swipes with CTC and trie-constrained beam search; 89% top-1 with the shipped dictionary (92% with every test word in it), hand-written Kotlin inference, weights on Hugging Face.

**WorkFromPhone** (Flutter, FastAPI, 2025–26): mobile development environment for remote Linux machines, with persistent PTY terminals, Git and file management, server monitoring, dev-server previews and autonomous AI coding tasks.

**P2P Network Protocol** (Python, 2024): peer-to-peer protocol for decentralized peer connection, domain name resolution and resource discovery across intranet and internet, with Ed25519-signed records.

SKILLS

|                       |                                                                                               |
|-----------------------|-----------------------------------------------------------------------------------------------|
| AI / ML               | PyTorch, TensorFlow, YOLO, OpenCV, OCR, U-Net, TCN/CTC, LSTM, regression, genetic algorithms  |
| GenAI & LLMs          | GraphRAG, Neo4j, ChromaDB, Claude API, Model Context Protocol, vLLM, Whisper                  |
| Infrastructure & edge | NVIDIA DGX Spark, Jetson AGX Orin, Docker, Docker Swarm, RHEL, air-gapped deployment          |
| Backend & data        | Python, FastAPI, PostgreSQL, Django, SSE, PI Server, AWS (IoT Core, Lambda, RDS, S3, Amplify) |
| Frontend & mobile     | React, TypeScript, Kotlin, Jetpack Compose, Flutter, Dart, React Native, Tailwind CSS         |
| IoT                   | Raspberry Pi, OBD-II, GPS/4G, Bluetooth Low Energy, Samsung SmartThings                       |

EDUCATION

Postgraduate Diploma in AI: Architecture, Design & Implementation · Georgian College, Canada

B.Sc. Information Technology · University of Mumbai

Jan 2023 – Aug 2023

Jul 2018 – May 2022

CERTIFICATIONS

|                                                  |                                                           |
|--------------------------------------------------|-----------------------------------------------------------|
| Building with the Claude API · Anthropic, 2026   | Model Context Protocol: Advanced Topics · Anthropic, 2026 |
| AWS Academy Machine Learning for NLP · AWS, 2023 | AWS Cloud Technical Essentials · AWS, 2023                |