Harsh Amin

587-664-9195 | aminharsh317@gmail.com | harshamin.me | hamin06 | hamin2006

Education

University of British Columbia April 2027
BSc Co-op Computer Science · Math *Vancouver, BC*
GPA: 4.00

Awards: Trek Academic Excellence Scholarship - awarded to top 5% of faculty
Extracurricular: UBC Launch Pad Design Team, Intramural Hockey, Weight Lifting Club, Entrepreneurship Club

Technical Skills

Languages: Java, Python, C/C++, Racket, R, HTML/CSS, JavaScript, TypeScript, SQL, GraphQL
Frameworks: LangChain, LlamaIndex, Agents, Ollama, React, Flask, FastAPI, Node.js/Deno, Express.js, SQLAlchemy
Developer Tools: AWS, AWS CDK, REST, OpenAPI, PostgreSQL, MongoDB, Docker, Git, Linux/Unix, Jenkins

Experience

Software Developer Intern May 2025 – Dec 2025
Amazon Web Services - Cloud Innovation Centre *Vancouver, BC*

- Built **open source Generative AI** applications **end-to-end** for **3+** sponsor teams, leveraging **AWS Bedrock Guardrails, IAM, WAF, Shield, Cognito, Secrets Manager** to deliver secure and scalable **AI/ML** solutions
- Developed an **AI Agent** using an **MCP Server** that maintained context across **10+** multi-turn interactions, improving dataset query accuracy by **25%+** for researchers, storing conversational history in **DynamoDB**
- Migrated from unreliable local Docker workflows to a standardized **AWS CI/CD** pipeline (**CodeBuild, CodePipeline, ECR**), cutting cross-environment inconsistencies by **33%** and reducing deployment time by **40%**
- Collaborated with a cross-functional team of **5+** project developers, managers, and stakeholders, presenting technical solutions and documenting each phase for seamless deployment and knowledge transfer

Technical Software Team Lead Sep 2024 – Present
UBC Launchpad - Project: Forum AI (LLM based student question board for UBC classes) *Vancouver, BC*

- Designed the **Retrieval Augmented Generation** model and updated it to be **20%** quicker
- Co-lead a **15+** member team of developers and designers, arranging weekly syncs and tickets over 8 months
- Migrated to **FastAPI** endpoints and implemented trigger-based embedding retrieval from a **PGVector** database, resulting in a **27%** reduction in API response time
- Led the migration to the **Deno** runtime, optimizing containerized function performance and achieving a **20%+** improvement in execution speed within **Docker** environments for resource heavy functions (ex. RAG)

Technical Projects

OpenEd - AI Study Companion | *AWS CDK,WAF,RDS, Bedrock, LangChain, Docker, RAG, OpenAPI* Dec 2025

- Developed a cloud-native application sponsored by BCcampus that will enable **100,000+** students/teachers to streamline studying, teaching, and learning material formation by $\sim$ **30%**
- Built a textbook-specific chat assistant leveraging **AWS Bedrock** and a **LangChain RAG** pipeline, scaled to process **750+** textbooks with **15+** hours of audio/video transcripts, storing embeddings in **RDS PostgreSQL**
- Deployed a production-ready **API Gateway** stack from an **OpenAPI** spec, integrating **10+ Lambda** functions (via **Docker** images and **Lambda Layers**) with automated **CI/CD** integration scaling to **1K+ requests/day**

Jarviz | *LangGraph, Asyncio, OpenCV, OpenAI Whisper, ElevenLabs, WebSockets, PyQt5, OCR, Gemini* Jan 2026

- Engineered a multimodal **AI agent** using **LangChain** and **LangGraph**, orchestrating **5+** complex tools (real-time vision, OCR translation, geolocation) to enable context-aware reasoning and multi-turn conversations
- Built a highly concurrent Python asyncio backend integrating local **Whisper STT** and **ElevenLabs TTS**, minimizing latency via WebSocket streaming to facilitate seamless voice-to-voice interaction
- Developed an immersive, transparent HUD with **PyQt5** and **OpenCV**, rendering dynamic widgets and webcam feeds at **60 FPS** while synchronizing agent state in real-time

Auto Pilot | *Jetson Nano, YOLOv8, TensorRT, OpenCV, PyTorch, Jetson.GPIO* Apr 2025

- Developed an edge-based self-driving car software stack on a **Jetson Nano** to perform real-time object detection, lane following, and steering control using **deep learning** and **computer vision**
- Trained and optimized a **YOLOv8** model using **PyTorch** for road sign and obstacle recognition; exported to **ONNX** and accelerated inference with **TensorRT** to achieve real-time performance on embedded hardware
- Utilized **OpenCV** for **perspective transformations** to assist in path planning and steering control logic