

Abdul Rehman Khan

Toronto, Canada, +1 561-614-0057, a13khan@torontomu.ca, www.codedbykhan.com

Artificial Intelligence Engineer | LLM Systems & Generative AI Expert

Professional Summary

Strategic AI engineer with 5+ years of experience building cloud-native ML systems and production-ready intelligent solutions. Skilled in deploying LLM-powered agents, conversational AI, and real-time inference pipelines using AWS Bedrock, LangChain, and OpenAI APIs.

Technical Skills

Foundation Models & LLMs: OpenAI GPT, Claude, LLaMA, AWS Bedrock, LangChain, LangGraph
Model Engineering: Fine-tuning, prompt engineering, embedding-based retrieval (RAG), agent orchestration
NLP Tasks: Document understanding, summarization, classification, data extraction, chatbots
Programming Languages: Python, JavaScript, HTML, CSS, C++, C#, Java, Golang, PHP, Kotlin
Frontend & Backend : React.js, Angular, Vue.js, Tailwind CSS, jQuery, Flask, Node.js, Express.js, Spring Boot, Django,
Databases: MySQL, PostgreSQL, MongoDB, Oracle
Cloud, DevOps & Tools: Git, GitHub, Docker, Jenkins, Jira, AWS Lambda, Heroku, Azure, OpenShift
CI/CD: Continuous Integration (CircleCI, Jenkins), Unit tests, Test-driven development, CMake
Other: Next.js, Shopify, Klaviyo, Jest, Mocha, Selenium, Springboot, Kubernetes, Archer

Work History

AI Engineer, Toronto Metropolitan University, Toronto

10/2024 – Present

- Designed and deployed scalable AI-driven backend services to support research applications and internal tools using Python, Flask, and AWS Lambda.
- Built and maintained NLP pipelines for structured data extraction, classification, and summarization from academic documents.
- Integrated Lang Chain-based orchestration and prompt-engineered OpenAI models into student service chatbots and content agents.
- Led sprint planning sessions, enhancing internal buy-in for new solutions through persuasive communication.

Software Developer, Shaw Communications Inc, Toronto

09/2021 – 07/2023

- Developed cloud-based AI applications in telecom operations, including real-time alert systems and recommendation engines for customer support using AWS Sagemaker and Lambda.
- Utilized my proficiency in programming languages such as Java, Python, C#, C++, and C to develop robust and scalable AI software applications.
- Took ownership and ensured proactive closure for problem reports that fall within my area of responsibilities, worked closely with QA and technical support to quickly and successfully closeout reported problems.

Software Developer, Intern, International Business Machines, Markham

05/2019 – 09/2020

- Conducted data analysis, designed surveys, and utilized critical thinking to understand user behavior, identify patterns, and address anomalies.
- Proposed innovative solutions to enhance the IBM Technology Zone's user experience and supported tasks as directed by the team lead.

Front End Engineer, Intern, Amazon Development Centre, Toronto,

05/2017 – 09/2018

- Delivered software components and improvements to software features with my team's software.
- Participated in code reviews and daily stand-ups.

Education

Designing & Building AI Solutions and Services. Massachusetts Institute of Technology, USA 09/2023 – 02/2024

B. Eng. Computer Engineering, Toronto Metropolitan University, Canada

09/2016 – 07/2021

Projects

Developed an AI-powered LMS using React, Node.js, PostgreSQL, and AWS with GPT-based tutoring, auto-grading via NLP, personalized learning recommendations, dropout risk prediction using ML, and full CI/CD deployment with Docker and Kubernetes. www.bytesinstituteoftechnology.com