NIHAL KHAN

AI ENGINEER

Ontario, Canada

+1 (437)-556-9983

mohammednihal281001@gmail.com

ABOUT ME

AI & Full Stack Developer at NexApproach, proficient in Python and AWS, engineered high-performance platforms that supported 500+ users. Demonstrated strong problem-solving skills by architecting resilient systems with sub-100ms latency, and successfully integrated AI workflows, enhancing service availability and optimizing operational efficiency.

SKILLS

Programming: Python, C++, C, TypeScript, JavaScript
Frontend: React.js, Next.js, Tailwind CSS, Framer Motion
Multimedia: H.264, HEVC, AV1, FFmpeg, OpenCV
Cloud & DevOps: AWS, GCP, Docker, Kubernetes, API Gateway, CI/CD, Hugging Face
AI & ML: LangChain, OpenAI APIs, RAG Pipelines, Prompt Engineering, Vector Search, SHAP, N8N
Security & Governance: Threat Modeling, API Access Control, Secure AI Workflow Design, AI Governance Principles
Data & Automation: Lightweight ETL, Workflow Automation

LANGUAGES

ENGLISH

HINDI

URDU

PERSONAL DETAILS

Nationality: Indian
Visa status: Valid Canadian Work Permit
Marital status: Single

WORK EXPERIENCE

NEXAPPROACH
Toronto, Ontario
May 2025 - Present

AI & FULL STACK ENGINEER

- Engineered a high-performance Next.js platform with LangChain, supporting 500+ simultaneous users via real-time WebSockets and advanced SSR optimization.
- Architected a resilient AWS environment, Lambda (Node 18.x), API Gateway (custom auth), and DynamoDB, achieving sub-100ms latency and 99.5% service availability.
- Delivered secure, scalable REST APIs leveraging JWT, controlled traffic with 100 req/min rate limits, and integrated GPT-4 with cost-optimized, fault-tolerant workflows.

EDUCATION

UNIVERSITY OF WINDSOR
Windsor, Ontario
Sept 2023 - Dec 2024

M. ENG IN ELECTRICAL & COMPUTER ENGINEERING

SATHYABAMA UNIVERSITY
Chennai, Tamil Nadu
June 2019 - May 2023

B. TECH IN INFORMATION TECHNOLOGY

PROJECTS

JUNE - SEPTEMBER 2025

RETRIEVAL-AUGMENTED GENERATION (RAG) AI ASSISTANT

Engineered a RAG-based LLM assistant using LangChain and OpenAI APIs, showcasing potential for financial knowledge retrieval and low-latency decision support. Designed agentic workflows with custom input/output handling, reducing hallucinations by 20%. Integrated lightweight ETL and deployed on Hugging Face Spaces via Docker for rapid prototyping. *https://github.com/nihalkhan2810/NihalLLM*

MARCH - MAY 2025

REACT PORTFOLIO

Built a fully responsive portfolio using React.js, Tailwind CSS, and Framer Motion to showcase ML projects and publications. Integrated GitHub and HuggingFace APIs to auto-sync repositories and deployed the site on Vercel.
https://github.com/nihalkhan2810/react-portfolio
https://nihalportfolio-rho.vercel.app/

JANUARY - MARCH 2025

SECURE LLM AGENT FRAMEWORK:

Developed LLM-based agent framework with layered API access control and secure prompt routing., Integrated LangChain with OpenAI and external APIs while mitigating prompt injection vulnerabilities. Aligned architecture with emerging AI governance standards including transparency and control policies.