Md Shariar Hossain

Senior AI & Machine Learning Engineer

Dhaka, Bangladesh • +8801853297010 • shmozumder2@gmail.com
shariar-mozumder.github.io • LinkedIn • GitHub • Google Scholar

Summary

Senior AI & Machine Learning Engineer with 5+ years of experience designing and shipping production-grade AI systems, including multi-agent orchestration platforms, Retrieval-Augmented Generation (RAG) pipelines, fine-tuned Large Language Models, and real-time computer vision services. Proven expertise in LangGraph, LangChain, and Model Context Protocol (MCP) based agentic architectures, with end-to-end delivery from model fine-tuning to scalable FastAPI deployment on AWS and Google Cloud Platform. Experience in building models from scratch using core ML and Deep Learning algorithms, Neural network,RNN, CNN, LSTM and Transformers architecture and fine-tuning LLMs with custom dataset. Author of three peer-reviewed publications in machine learning and security, and creator of the open-source **Retune** framework for AI agent observability, evaluation, and optimization.

Skills

Agentic AI & LLMs: Multi-Agent Systems, LangGraph, LangChain, Phidata, ReAct Agents, Reflect Agents, Supervisor Agents, MCP Servers & Clients, OpenAI API, Gemini API, Ollama, Hugging Face Transformers, Llama 3, Prompt Engineering, Tool Calling

Machine Learning & Deep Learning: PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas, Transfer Learning, LLM Fine-Tuning, YOLOv8, YOLOv11-seg, Stable Diffusion, BERT

Retrieval-Augmented Generation: FAISS, Vector Databases, Sentence-Transformers, Prompt Templates, Semantic Search, Embedding Pipelines

Voice AI: Whisper (Speech-to-Text), Google TTS, Coqui TTS

Backend & APIs: Python, FastAPI, Django, Flask, Celery, .NET Core, Microservices, REST, GraphQL, gRPC

Cloud, DevOps & Deployment: Docker, Kubernetes, GitHub Actions, Jenkins, AWS (Lambda, AppSync, SQS, SNS), Google Cloud Platform, CI/CD

Data & Messaging: PostgreSQL, Microsoft SQL Server, MongoDB, Redis, Vector Databases, RabbitMQ, gRPC, MQTT

Web Automation & Scraping: Selenium, Playwright, BeautifulSoup, Google Search API

Programming Languages: Python, C#, C++, Java

Tools & Collaboration: Git, GitHub, GitLab, Bitbucket, Jira, Postman, Label Studio, VS Code, PyCharm, IntelliJ IDEA

Work Experience

FlowGenX AI **Remote**

AI & Machine Learning Engineer *Jan 2025 – Present*

- Engineer an enterprise **LangGraph-based Multi-Agent Workflow Orchestration Platform** (positioned alongside n8n, MuleSoft, and Boomi) that enables users to design, deploy, and monitor AI-driven automations end to end.
- Architected the runtime around a node- and edge-based graph with **ReAct Agents, Supervisor Agents, Tools, Router Nodes, Loop Nodes, and Trigger Nodes**, supporting dynamic tool calling and deterministic orchestration.
- Delivered structured input/output handling, an **Expression Resolver**, dynamic prompting, and multi-provider LLM routing, alongside persistent short- and long-term memory backed by **Redis and PostgreSQL**.
- Integrated **MCP Servers and Clients** as first-class internal and external tools, enabling composable, standards-based tooling across agent workflows.
- Built authenticated **App Connectors** with OAuth and API-key flows to integrate enterprise applications into automated workflows.

ESAP AI (Empowering Energy) **Remote**

Machine Learning Engineer *Jul 2024 – Dec 2024*

- Designed and delivered a **fully automated multi-agent ERP solution** with HR and smart price-quotation modules, driven by voice-input distributed AI agents with dynamic tool calling.
- Architected a **Collaborative Multi-Agent System (MAS)** of autonomous specialist agents coordinated by a Coordinator Agent, with mandatory data validation across dynamic dashboards.
- Built agentic pipelines using **LangGraph, LangChain, and Phidata**, and integrated **Selenium, Playwright, Google Search API, and BeautifulSoup** for live vendor-price scraping.
- Developed an **Interior Recommendation System (YOLOv11-seg + AR)**: fine-tuned YOLOv11-seg on a custom-annotated dataset (Label Studio) to segment walls, floors, ceilings, and furniture, driving AR-based color recommendations.

NybSys Inc. **Bangladesh**

Software Engineer – Machine Learning & Backend *May 2022 – Jun 2024*

- Built a **custom agentic RAG chatbot** for e-commerce customer support by fine-tuning **Llama 3.2-3B** on proprietary company-policy data, with FAISS and Sentence-Transformers retrieval and a FastAPI inference pipeline.
- Delivered an **AI-based surveillance system** for public safety using a fine-tuned **YOLOv8** human-detection model, OpenCV for frame counting/saving/optimization, and FastAPI for real-time inference.
- Engineered microservices for **Nexus NMS (Network Management System)** using Python, FastAPI, and Celery, enabling proactive performance optimization, automated issue resolution, and routine-task scheduling.
- Built the backend for **Nybsys Sentra PTT App and Dispatch**, an IoT-based private-wireless solution for zonal and emergency communication, using C# and .NET Core with RabbitMQ and gRPC messaging.

Pro Data Ltd. **Bangladesh**

Software Engineer – Android *Mar 2021 – Apr 2022*

- Developed Android applications for financial and banking clients including **National Life Insurance** and **Mercantile Bank Ltd.**
- Integrated REST and GraphQL APIs using HTTP Client, Retrofit, and Volley.

Projects

- **Retune – AI Agent Observability, Evaluation & Optimization Framework** [Open Source] [PyPI | GitHub]
  - Designed, authored, and published an open-source framework for agent observability following an **Observe → Evaluate → Optimize** loop; distributed via `pip install retune` and open for community contributions.
  - Ships a Python SDK and a management dashboard for action, tool-call, and I/O inspection; internally uses a Deep Agent architecture for evaluation and self-optimization.
- **Cholche – Live-Bidding Ride-Sharing Platform (Founder)** [cholche.com]
  - Founded and launched a scalable inter-city ride-sharing product for Bangladesh with a live-bidding system for rider-driver fare negotiation; web platform live, native mobile app in development.
- **RAG Q&A on the Holy Quran Translation and Tafseer** [GitHub]
  - Built a retrieval-augmented Q&A system using **Llama 3.2-3B** for generation, LangChain for orchestration, and FAISS with Sentence-Transformers for verse-level retrieval over Quran translation and Tafseer.
- **Smart Product & Vendor Quotation Engine** [GitHub]
  - Built an AI-driven scraping pipeline for product and vendor pricing using Google Search API, DuckDuckGo, Selenium, Playwright, and BeautifulSoup, with embedding-based filtering via Ollama open-source LLMs.
- **Stable Diffusion Fine-Tuning for Image-to-Image Generation** [GitHub]
  - Fine-tuned `runwayml/stable-diffusion-v1-5` on a custom fashion dataset for image-to-image generation.
- **Medical Q&A Model on Hugging Face Spaces** [HF Spaces]
  - Fine-tuned a BERT-based Text-to-Text generation model on a medical symptoms-to-disease dataset and deployed an interactive Q&A interface via Gradio on Hugging Face Spaces.

Publications

- Mahmud, A., **Hossain, M. S.**, Meya, M. R., Didar, D. I., Islam, H., Dhar, P., & Akter, P. (2025). *Prompt-Based Fine-Tuning of Bangla Language Models for Sequence Classification.* Lecture Notes in Networks and Systems, pp. 409–424. https://doi.org/10.1007/978-981-96-2721-9_27
- **Hossain, M. S.**, & Haque, A. B. (2022). *Blockchain-Based Services in Education: A Bibliometric Analysis.* International Conference on Machine Intelligence and Emerging Technologies, Springer Nature, pp. 348–362. [Link]
- **Hossain, M. S.**, & Riaz, M. H. (2022). *Android Malware Detection System: A Machine Learning and Deep Learning Based Multilayered Approach.* Intelligent Computing & Optimization (ICO 2021), Springer International Publishing, pp. 277–287. [Link]

Education

Noakhali Science and Technology University, Bangladesh *Feb 2017 – Feb 2021*
BSc in Computer Science and Telecommunication Engineering • CGPA 3.30 / 4.00
Thesis: Android Malware Detection System – A Machine Learning and Deep Learning Based Multilayered Approach
Undergraduate Research Assistant (Mar 2020 – Jan 2021), Advisor: Md Hasnat Riaz, Assistant Professor, Dept. of CSTE.

Feni Govt. College, Bangladesh *2015*
Higher Secondary Certificate (Science) • GPA 4.83 / 5.00

Central High School, Bangladesh *2013*
Secondary School Certificate (Science) • GPA 5.00 / 5.00