Md Abu Sayed Shaun

Dhaka, Bangladesh • +880 1903 248784 • sayedshaun4@gmail.com • LinkedIn • GitHub • Portfolio

Experience

AI Engineer • Synesis IT PLC – Dhaka, Bangladesh June 2026 – Present

- Applied optimization techniques, including quantization, to open-source SLMs and LLMs such as Gemma and Qwen, improving inference speed and reducing model size for efficient deployment.
- Built and fine-tuned Bangla speech recognition (ASR) and text-to-speech (TTS) models, combining them with LLMs to power a hands-free voice bot supporting fully spoken, real-time interaction.
- Designed and integrated LLM-driven features into Convay, an online meeting platform, delivering intelligent meeting experiences and productivity enhancements.

Machine Learning Engineer • Giga Tech Limited – Dhaka, Bangladesh September 2024 – June 2026

- Pretrained and fine-tuned Bangla language models (BERT, GPT-2, T5, LLaMA) on large-scale in-house corpora, and built downstream NLP models for NER, POS tagging, QA, next-word prediction, and WSD, among others.
- Developed an agentic RAG-based HR chatbot that automated 90% of employee query resolution across SQL databases and vector stores using open-source LLMs.
- Built an AI-powered web crawler extracting structured and unstructured data from heterogeneous websites via adaptive parsing and LLM-based content understanding.

Trainee Software Engineer • Syntax Solution Limited – Dhaka, Bangladesh June 2024 – August 2024

- Worked on a recommendation engine, replacing keyword search with a hybrid architecture combining semantic similarity and keyword-based retrieval.
- Built a Retrieval Augmented Generation powered chatbot pipeline to automate repetitive customer support queries.

Education

Bachelor of Science • **World University of Bangladesh** January 2020 – April 2024

- **Major:** Computer Science and Engineering
- **Coursework:** Data Structures, Algorithms, OOP, Statistics, Data Mining, Neural Networks, System Design.
- **Thesis:** "Advancement in Neuroimaging: Automated Identification of Brain Strokes through Machine Learning."

Publications

- **BanSuite: A Unified Toolkit and Software Platform for Low-Resource NLP in Bangla**. *ACL 2026*. Link
  Md. Abu Sayed, Faisal Ahamed Khan, Jannatul Ferdous Tuli, Nabeel Mohammed, Mohammad Ruhul Amin, and Mohammad Mamun Or Rashid.

Projects

DeepHarness Link

- A framework for orchestrating multiple specialized AI agents that collaborate on complex tasks, coordinating task delegation, context sharing, and parallel execution across agents.

OneCrawler Framework Link • App Link

- An async Python crawling framework for URL discovery, link extraction, and structured scraping, with browser-based extraction, sitemap discovery, and optional GenAI-powered structured outputs.
- Built a hosted web application on top of the framework so users can configure and run crawls in the browser without writing code.

Corpus Link

- Core contributor to this large-scale unified NLP platform, working on corpus development, syntactic annotation, and low-resource language models for Bangla NLP researchers.

LangTrain Link

- A PyTorch-based high-level Python package for rapid language model development, supporting large-scale tokenization, distributed training, and evaluation.

Technologies

AI/ML: PyTorch, Transformers, LangChain, LangGraph, TensorFlow, Scikit-Learn, Pandas, NumPy, Lightning, HuggingFace, vLLM, Ollama, llama.cpp, Onnx, Nvidia Triton, LitServe, VectorDBs, Weight & Bias, Prefect.

Backend/Ops: Python, FastAPI, Docker, PostgreSQL, MongoDB, JWT Authentication, Redis, Celery, ElasticSearch, CI/CD.

Certifications

- Generative AI with Large Language Models • DeepLearning.AI • AWS • Coursera Link
- Machine Learning Specialization • DeepLearning.AI • Stanford University Link