MD. HASAN IMON

ASPIRING AI/ML ENGINEER

• Savar, Dhaka, Bangladesh • emon.mlengineer@gmail.com • 01834363533 • Portfolio • Linkedin • GitHub

SUMMARY

Generative AI & Agentic AI Specialist with expertise in fine-tuning, multi-agent systems, RAG pipelines, and tool-augmented AI applications using LangGraph and modern orchestration frameworks. Skilled in delivering production-ready AI solutions with advanced reasoning, planning, and intelligent tool usage. Also experienced in Machine Learning and Deep Learning, including classification, regression, clustering, MLOps, NLP, transformers, and neural networks for scalable real-world applications.

EXPERIENCE

Junior Backend AI Engineer

NeuroStack AI Solutions - Indiana, USA (2025 - Present)

- Engineered tool-augmented reasoning pipelines, integrating APIs, databases, and web search, while fine-tuning LLMs for high-performance, scalable real-world applications.
- Designed and deployed enterprise-grade multi-agent AI systems with LangGraph orchestration, modular agents, and intelligent fallback workflows for robust automation.

EDUCATION

Bachelor of Computer Science & Engineering

City University, Bangladesh (2022 - Present)

CERTIFICATION

- **Machine Learning Specialization** by Stanford University (Coursera)
- **Programming for Everybody** by University of Michigan (Coursera)

SKILLS

- **Generative AI & Agentic AI:** Multi-Agent, Fine-Tuning(PEFT, LoRA, QLoRA), AgentOps, RAG Pipelines, HITL, Tool-Augmented AI, Memory-Augmented, Prompt Engineering
- **Machine Learning & Deep Learning:** Classification, Regression, Clustering, NLP, LLMs, Neural Networks (CNN, RNN, LSTM, GRU), Attention Mechanisms
- **Frameworks & Libraries:** LangChain, LangGraph, LlamaIndex, CrewAI, PyTorch, TensorFlow, Keras, Scikit-learn, Hugging Face, NLTK, spaCy, Transformers
- **Databases:** MySQL, MongoDB, ChromaDB, Pinecone, Vector Databases
- **Other Tools & Technologies:** DVC, Docker, CI/CD, FastAPI, FlaskPython, Git, GitHub, API Integration, Pandas, NumPy, Matplotlib, Seaborn

PROJECTS

TrueWealth AI: Your AI-powered financial strategist – GitHub

Multi-agent financial advisor with **Query breakdown & multi-stage planning.**

- Developed a LangGraph-orchestrated multi-agent AI pipeline (Planner, Retriever, News Agent, Memory Agent, etc) to analyze and synthesize financial information from multiple sources for structured financial reasoning.
- Leveraged dynamic Tool Router and multi-source reasoning for accurate and adaptive investment guidance.
- Designed a user interface with seamless interaction to assist users in real-time financial decision-making.

InformaTruth: AI-Driven News Authenticity Analyzer – GitHub

Build a real-time fake news detection AI **using Fine-tuned models.**

- Developed a multi-agent system leveraging fine-tuned Roberta (classification) and Flan-T5 (explanation) to detect misinformation for explainable outputs, ensuring both predictive and user interpretability from different input.
- context-aware predictions, ensuring higher accuracy and reliability. Dynamic Tool Router and fallback agents to retrieve verified sources from DuckDuckGo and Wikipedia.
- Designed transparent, user-friendly outputs that public trust, ensuring explainable AI for news evaluation.

AutoDocThinker: Intelligent Agentic Search Engine – GitHub

Build an intelligent document QA system **using Agentic RAG**

- Developed a multi-format ingestion pipeline with database storage, enabling the system to parse, chunk, and semantically retrieve knowledge from heterogeneous documents with high accuracy and scalability.
- Implemented dynamic Tool Routing and fallback strategies that adaptively select the most reliable information source (RAG, Web Search, or Memory) based on context, ensuring trustworthy and resilient information delivery.
- Designed a user interface enabling seamless interaction with documents of any format.

Translatica: English to Spanish Translation – GitHub

Developed a **Fine tuned(with PEFT + LoRA)** for high-fidelity AI literary translator

- preserving nuanced tone, stylistic elements, and contextual meaning. Integrated evaluation metrics (BLEU, METEOR, and human evaluation) to ensure translation quality and consistency across diverse literary genres.
- Optimized model for low-latency inference suitable for CPU deployment and Built a user-friendly Flask interface to provide smooth, real-time literary translations.

MediGenius: Medical AI Assistant – GitHub

Build an AI-driven medical assistant with M**ulti-agent communication control.**

- Multi-agent medical AI assistant orchestrated with LangGraph, enabling dynamic communication between Planner, Retriever, Generator, and Fallback agents. Leveraged RAG to deliver doctor-like, empathetic responses from medical PDFs, research papers, and fallback web sources like various website.
- Integrated short-term memory and state-based reasoning for context-aware, multi-turn consultations.
- Designed an accessible, responsive web interface optimized for low latency, accurate guidance, and adaptive feedback, ensuring a professional-grade for deliver doctor-like, empathetic responses.
