ASIF FAISAL CHOWDHURY

AI/ML Engineer | Multimodal AI for Construction & AEC

Email: afschowdhury@gmail.com
Phone: +1(225)506-8796
LinkedIn: linkedin.com/in/afschowdhury
GitHub: github.com/afschowdhury
HuggingFace: huggingface.co/afschowdhury

Summary

AI/ML engineer and PhD researcher applying machine learning to construction and AEC problems. Three years of production NLP and chatbot experience, now conducting NSF-funded research that brings LLMs and vision-language models to construction: predicting worker movement for exoskeletons and benchmarking how well multimodal models read real construction drawings. Concurrently works as an energy analyst on the U.S. DOE ITAC program at LSU, assessing industrial facilities on site and identifying seven-figure savings. Combines a hands-on AI engineering toolkit with direct knowledge of construction documents and field operations.

Education

Doctor of Philosophy (Ph.D.) | ***Construction Management*** 2024 – Present
Louisiana State University Baton Rouge, LA

Bachelor of Science | ***Information and Communication Technology*** 2018 – 2022
Islamic University Kushtia, Bangladesh

Technical Skills

- **Languages:** Python, C/C++, SQL, MATLAB
- **Multimodal & Computer Vision:** vision-language models (VLMs), CLIP, ImageBind, Moondream, document & drawing understanding, image classification, OpenCV, Whisper (speech-to-text)
- **LLM & Agentic AI:** agentic / multi-agent pipelines, tool use & function calling, structured (JSON) outputs, retrieval-augmented generation (RAG), prompt engineering, chain-of-thought, human-in-the-loop, context management, evaluation/benchmarking
- **Models & APIs:** GPT-4o / GPT-4o-mini, GPT-5.4, OpenAI API, Gemini, Gemma, Qwen-VL, Llama
- **ML & Deep Learning:** PyTorch, TensorFlow, Hugging Face Transformers, scikit-learn, fine-tuning, embeddings & semantic retrieval, Rasa
- **On-device & Edge / MLOps:** Ollama, local VLM deployment, latency optimization, Git, Docker, Linux, AWS (EC2, S3), REST APIs, Flask
- **Spoken Languages:** English (fluent), Bengali (native)

Research & University Experience

Graduate Research Assistant (NSF-funded) 2024 – Present
Bert S. Turner Dept. of Construction Management, LSU Baton Rouge, LA

- Apply LLMs, vision-language models, and agentic pipelines to construction tasks on an NSF-funded project (Grant No. 2222881), including a first-authored locomotion-prediction system for exoskeletons and a benchmark for reading construction drawings.
- Take projects from first idea to working prototype and write them up as peer-reviewed papers.
- Design multimodal datasets and machine-learning models that assess heavy-equipment operator skill from machine telemetry, gaze, and physiological signals on construction simulators.

Energy Analyst Sep 2024 – Present
LSU Industrial Training & Assessment Center (U.S. DOE ITAC program) Baton Rouge, LA

- Conduct on-site energy assessments at manufacturing plants and large facilities, serving as lead analyst on several and contributing to more than a dozen overall.

- Identify seven-figure (US$1M+) annual savings across recommendations, with a substantial share already implemented by client facilities.
- Recommend several million kWh per year of electricity reductions, equal to an estimated few thousand metric tons of CO_2 avoided annually.
- Author client technical reports and perform utility-bill analysis, on-site measurement, and energy-balance calculations across electricity, natural gas, and other fuels.

Research Assistant Jan 2021 – Sep 2022

ICT Innovation Lab, Islamic University Kushtia, Bangladesh

- Developed and tested machine-learning models for computer-vision and IoT projects, delivering working prototypes.
- Built and deployed IoT and embedded-systems projects with attention to security and scalability, and contributed to several computer-vision efforts.

Selected Projects

- **Locomotion-prediction agent for construction exoskeletons** (*first author, led*): Built an agent that predicts a worker's next movement so an exoskeleton can assist. A local vision-language model (Moondream via Ollama) describes the scene every two seconds; a GPT-4o-mini step combines that with the worker's voice command and recent movements, and asks the user to confirm when confidence is low. Reached 94.3% accuracy at 0.5 s median latency in a simulated live pilot.
- **AI Coaching Agent for Personalized Excavator Operator Training** (*first author, led*): Closed-loop multi-agent vision-language system that trains novice excavator operators on a simulator. Specialized agents split the work of reading live screen, camera, gaze, and joystick data, diagnosing each operator's skill gaps, and generating targeted corrective feedback, while a retrieval-augmented generation (RAG) knowledge base validates every recommendation so the feedback stays grounded and credible. The coaching is built on a multimodal study of how experts differ from novices (machine telemetry with eye-tracking, heart-rate, and productivity signals), which locates the perceptual and cognitive gaps that personalized training must close.
- **Agentic report-validation system for industrial energy audits** (*sole author*): Built a multi-agent LLM system that validates ITAC energy-audit reports for correctness, to cut manual review time. An orchestrator assigns each check to specialized agents that call whatever tool the task needs, from live evidence retrieval to programmatic computation, and the agents cross-verify so every number and recommendation stays traceable and correct.
- **ConstructionQA, a multimodal-LLM benchmark for drawing sets** (*co-author*): Benchmarked five multimodal LLMs on a real 100% Construction Documents set (150+ sheets, 8 engineering disciplines, 120 question-answer pairs) across single-sheet lookup, sheet-reference retrieval, and cross-sheet compliance reasoning. Found that frontier models lead but every model fails at sheet-reference retrieval, the main blocker for AEC deployment.
- **Zero-shot locomotion prediction with GPT-4o** (*co-author*): Compared a fine-tuned CLIP model, GPT-4o run zero-shot with chain-of-thought prompting, and ImageBind on vision and speech from smart glasses (Whisper transcription), using data augmentation to balance classes. Fine-tuned CLIP reached 90.05% F1; GPT-4o followed closely at 87.87% F1 with no task-specific training.
- **Onusondhan**, **Retrieval-MPNet-BN**, and **GAN image generation** (*open source*): A transformer that answers Bengali questions from a passage (GitHub); an MPNet passage-retrieval model for English/Bengali QA (Hugging Face Hub); and a generative adversarial network that synthesizes anime and human faces (GitHub).

Industry Experience

Prompt Engineer (Project-based) Aug 2024 – Jan 2024

Combokid (Remote) Hong Kong

- Built and integrated a multi-agent system into a child-healthcare chatbot, delivering tailored health advice and image-based features.
- Enhanced retrieval-augmented generation (RAG) for contextually relevant responses and tuned back-end performance to improve user engagement.

AI/NLP Engineer Dec 2022 – Jan 2024

Headless Technologies Limited Dhaka, Bangladesh

- Implemented and deployed AI chatbots for an e-commerce platform (engaze.ai) serving 1,500+ online stores and about 1 million user queries, automating customer support.
- Developed a national heritage chatbot *Mujib Jigyasa* (data curation + NLP fine-tuning) and a machine-learning classifier that routes government-office complaints.