MD. ASIF IQBAL FAHIM

PhD Researcher, Agentic AI & LLM Code Generation (UCD)•Lead AI Engineer•Kaggle Competition Expert(x2)•Reviewer (ACM TOSEM, Elsevier)

@ md.fahim@ucdconnect.ie +353892360266 linkedin.com/in/aifahim github.com/AIFahim kaggle.com/AIFahim
researchgate.net/profile/Md-Asif-Fahim datascience.stackexchange.com/AIFahim

EDUCATION

Ph.D. in Computer Science (NFQ Level 10, in progress)

University College Dublin, Ireland Sep 2025 – Present

- **Research area:** Building individualised software: self-evolving pipelines that autonomously generate, test, and integrate new functionality at runtime in response to natural-language requests, without a developer in the loop.
- Two first-author papers accepted during the first year of study: **ASE 2026** (CORE Rank A*) and **SEAMS 2026** (CORE Rank A, co-located with ICSE).
- **Outstanding Poster Award**, UCD School of Computer Science Graduate Research Poster Event, January 2026.
- **ACM SIGSOFT CAPS travel grant**, 2026.
- **Lab Demonstrator** (Jan 2026 – Present): Recommender Systems, Advanced Machine Learning, Data Science in Python, and Software Engineering.

B.Sc. in Computer Science and Engineering

Green University of Bangladesh Sep 2017 – Dec 2020

- **CGPA: 3.57 out of 4.00 —** Ranked 2nd in Batch — Received (2 times) **Dean Awards** for excellent Academic Performance.

WORK EXPERIENCE

Lead AI Engineer

InfinitiBit, Germany(Remote) | Dec 2024 – Aug 2025

- Led the development of an innovative Multimodal Agentic RAG system, implementing advanced PDF processing capabilities for extracting and analyzing text, tables, images, and diagrams using unstructured.io and ChromaDB
- Collaborated with **Siemens** on developing intelligent energy solutions, creating AI Agent Workflows for grid optimization and predictive maintenance applications
- Orchestrated AWS infrastructure using Terraform for automated EC2 instance scaling and management, optimizing resource allocation for AI agent workflows and ensuring high availability
- Developed and documented a metadata-aware RAG system with embedding caching strategies using Redis and semantic caching to optimize OpenAI API usage and reduce costs
- Implemented causal reasoning models and inference frameworks for time-series analysis, enabling more robust prediction capabilities in renewable energy forecasting applications
- Designed and deployed AI agent workflows for automated data processing pipelines, incorporating customized reasoning engines and decision-making frameworks
- Spearheaded optimization efforts for PyPSA (Python for Power System Analysis) library, including code profiling, debugging, and performance enhancement
- Implemented proof-of-concept for LLaMA 7B (PEFT) embedding extraction, enhancing the system's natural language processing capabilities

Software Engineer – Machine Learning/Recommendation Systems

IdeaScale Bangladesh Limited, Dhaka | December 2023 – November 2024

- Developed end-to-end recommendation systems using **ALS-based matrix factorization** and implemented **Context-Aware Recommendation Engine** with Elasticsearch for personalized user experiences
- Optimized LLM deployments using PEFT, LoRA, and **ONNX**-accelerated **Sentence-BERT**, while implementing OpenSearch **Vector Database** with **HNSW** for efficient semantic search
- Implemented adaptive **BERTopic** modeling with **incremental training** using IncrementalPCA and MiniBatchKMeans, deployed via **Celery**-powered parallel processing pipelines
- Architected and deployed ML solutions as microservices using Docker, Kubernetes, and Flask RESTful APIs, while mentoring junior team members

Machine Learning Engineer-Full time

REVE System Ltd., Dhaka | December 2022 – December 2023

- Developed and optimized **Speech-to-Text (STT)** solutions using OpenAI Whisper & wav2vec2, creating phoneme and grapheme models with dataset-specific vocabulary enhancement
- Implemented various **Text-to-Speech (TTS)** architectures (Align TTS, GlowTTS, FastSpeech2, vits, ProDiff) with multiple vocoders (HiFiGAN, MultiBandMelGAN, WaveRNN, UnivNet)
- Enhanced model performance through **Speech Augmentation** techniques and **dynamic training**, leveraging **JAX** framework, Peft, and LoRA for large-scale optimization
- Developed WhisperX **streaming service** with voice activity detection and forced phoneme alignment, deploying containerized solutions via FastAPI and Flask

Research & Teaching Assistant - Full time

RIoT Center Research LAB & Independent University of Bangladesh, Dhaka | September 2022 – 14th December 2022

- Worked on Research project: Response of Bevacizumab effectiveness in Ovarian Cancer. from WSI Dataset w/ fine-tuned State-of-the-art ViT model & achieved 96% Test Accuracy & 99% ROC AUC Score. **(Accepted in ICACIS 2024 and publication to the Springer Nature Singapore (Scopus indexed CCIS Series))**
- Taught the course under the **CSE department** of the Independent University of Bangladesh, - CIS 101: Fundamentals of Computer Systems LAB for three sections total students 40/Class.

Research Assistant - Full time

Institute of Advanced Research (IAR) LAB, United International University, Dhaka | February 2022 – August 2022

- Worked in Research project, "Semantic Biomedical Image Segmentation Using Attention Mechanism Driven Modified U-Net Based on Deep CNN" (funded by the Institute of Advanced Research, UIU, Grant Reference: UIU/IAR/02/2021/SE/25.) **(Published in High Impact Factor Journal(6.0) Neurocomputing, Elsevier)**

Junior Research Assistant - Part time

Consortium Research Center(Previously ResearchHub), Dhaka | Sep 2021 – Mar 2022

- Did Research on Efficiently Convert Time Series Data to Image & Deeplearing Modeling for Credit Card Fraud Detection.

NLP Specialization Mentor & Alpha Tester

Deeplearning.AI, USA | February '21 – May '21

- Tested curriculum before published to the Coursera & Provided feedbacks to course staffs. Guide learners if didn't understand the course contents.

Data Scientist - Independent Contractor

Dentalfynd(Previously ilekun Health), USA | May 2021 – July 2021

- Aggregate NYC borough's dentist data by web scraping (**Selenium & Scrapy**), Data cleaning, Predictive modeling.

Data Science - Trainee

CRIDA, Canada Based Organization, Dhaka | March 2021 – June 2021

- **Build Attention Based LSTM Networks**, Used TPU Strategy - Distributed Training for a large amount of Dataset, Used Frameworks Like PyTorch, TensorFlow.

Android Application Developer - Intern

Leotech Ltd.(Govt. Project), Dhaka | November 2019 – May 2020

- Worked with **MVVM Pattern, Rest API, Firebase, MLKit, Google Sceneform** in **fast-paced environment**, wrote maintainable and **reusable Java and Kotlin Code**.

RESEARCH & PEER REVIEW

- **Md. Asif Iqbal Fahim**, Alessio Ferrari, "ReqEvolve: User-Oriented Software Self-Evolution through Automatic Requirement Interpretation," **ASE 2026** Research Track (41st IEEE/ACM Intl. Conf. on Automated Software Engineering, **CORE Rank A***), Munich, Germany, October 2026. DOI: 10.1145/3832783.3834416.
- **Md. Asif Iqbal Fahim**, Alessio Ferrari, "Software Self-Extension with SelfEvolve: an Agentic Architecture for Runtime Code Generation," **SEAMS 2026** (21st Intl. Conf. on Software Engineering for Adaptive and Self-Managing Systems, **CORE Rank A**, co-located with ICSE), Rio de Janeiro, Brazil, 2026.
- Md. Rayhan Ahmed, **Md. Asif Iqbal Fahim**, et al., "DOLG-Next: Deep Orthogonal Fusion of Local and Global Features with ConvNext for Biomedical Image Segmentation," **Neurocomputing** (Elsevier), IF 6.0.
- **Md. Asif Iqbal Fahim**, et al., "Improved Transfer Learning Architecture to Classify Covid-19 Affected Chest X-Rays using Noisy Student Pre-training," ICIEV & ICIVPR, Japan.
- **Active reviewer for Elsevier journals** (Oct 2024 – May 2025): 11 completed reviews across *Neurocomputing* (5), *Natural Language Processing Journal* (3), and *Engineering Applications of Artificial Intelligence* (3). Reviewer for IVPR and ICIEV international conferences.
- **Reviewer** for *ACM TOSEM* (3 reviews, 2026), *Empirical Software Engineering* (2026), and the NeurIPS 2026 MusIML Workshop.

ML ACHIEVEMENTS & ACTIVITIES

- Selected as **Datascience Advisor** for **KaggleX BIPOC Program, 2024, Goal:** ***Help fellows to fine-tune Google LLM Gemma Model for downstream task***.
- Selected as **Datascience Mentor** for **KaggleX BIPOC Program, 2023, Goal:** ***Help mentee to complete end-to-end data science projects***.
- **"Expert" in Kaggle ML/DL Competitions** - Rank **Top 0.8%** among 182K+ & **"Expert" in Kaggle Discussions Forum** - Rank **Top 0.3%** among 269K+ active Kaggle users worldwide.
- **Silver Medal** (190th of 4,186 teams) in the OpenAI-hosted Kaggle competition "AI Agent Security – Multi-Step Tool Attacks", Sept 2026.
- **Bengali.AI Speech Recognition Comp., Kaggle** - **Silver Medal** Rank in Top 6%
- **PetFinder.my - Pawpularity Comp., Kaggle** - **Silver Medal** Rank in Top 2%
- **G-Research Crypto Forecasting Comp., Kaggle** - **Silver Medal** Rank in Top 5%
- **G2Net Gravitational Wave Detection Comp., Kaggle** - **Bronze Medal** Rank in Top 10%

- **Bristol-Myers Squibb – Molecular Translation Comp., Kaggle** - Rank in Top 13%
- **Cassava Leaf Disease Classification Comp., Kaggle** - Top 16%.
- **Dhaka AI Traffic Challenge Comp., GUB**- Finished at 45th.
- **Top 6%** Contributor to **Data Science Stack Exchange** in 2021. 13th Q. Answered, 1 Gold, 3 Silver, 14 Bronze Medals achived.
- AWS Machine Learning Scholarship Program, Udacity - **Scholar**

KEY PROJECTS & PRODUCTION SYSTEMS

Multimodal RAG System with Intelligent Caching - InfinitiBit

- Architected production RAG system processing PDFs with text, tables, images using ChromaDB vector storage
- Implemented Redis-based semantic caching reducing OpenAI API costs by 60%, deployed with Terraform on AWS

Real-time Speech Recognition Streaming Service - REVE Systems

- Developed WhisperX streaming service with WebSocket support, voice activity detection, and phoneme alignment
- Built containerized deployment with FastAPI, achieving sub-200ms latency for real-time transcription

Scalable Recommendation Engine with Vector Search - IdeaScale

- Built context-aware recommendation system using OpenSearch vector database with HNSW indexing
- Deployed ONNX-optimized Sentence-BERT models achieving 5x latency reduction in production

Distributed ML Training Infrastructure - REVE Systems

- Designed dynamic training system for 900-hour speech dataset with automated GPU scheduling
- Implemented JAX-based optimization pipeline reducing training time by 40% using model pruning

Incremental Learning Pipeline with BERTopic - IdeaScale

- Built production pipeline for topic modeling with incremental training using IncrementalPCA and MiniBatchKMeans
- Implemented Celery-based distributed processing handling 10K+ requests/minute with automated scaling

AI Energy Grid Optimization - InfinitiBit/Siemens Collaboration

- Developed AI agent workflows for predictive maintenance and grid optimization
- Implemented time-series analysis with causal reasoning models for renewable energy forecasting

CI/CD Pipeline for ML Model Deployment

- Built end-to-end CI/CD pipeline using GitHub Actions for automated model testing, validation, and versioning
- Integrated AWS CodePipeline for automated deployments with rollback capabilities and A/B testing support

AI Face Recognition Android App - Edge Deployment

- Deployed MobileNet SSD on mobile devices using TFLite with real-time OpenCV inference

Bengali.AI Speech Recognition - Kaggle Competition

- **Silver Medal (Top 6%)**: Developed optimized speech recognition models with production-ready data pipelines

Additional projects: https://aifahim.github.io/#project

CERTIFICATIONS

- Deep Learning Specialization, DEEPLEARNING.AI, Coursera.
- AWS MACHINE LEARNING NANODEGREE PROGRAM, AWS, UDACITY.
- Tensorflow Developer Specialization, DEEPLEARNING.AI, Coursera.
- AI for Medicine Specialization, DEEPLEARNING.AI, Coursera.
- Generative Adversarial Networks (GANs) Specialization, DEEPLEARNING.AI, Coursera.
- Natural Language Processing Specialization, DEEPLEARNING.AI, Coursera.
- Intro to ML: Image Processing, Qwiklabs.
- Intro to ML: Language Processing, Qwiklabs.
- Android App Development Specialization, Vanderbilt University, Coursera.

ADMINISTRATIVE EXPERIENCE

- **Advisor**, IEEE GUB Student Branch
- **Programming Secretary**, Green University Programming Club
- **Joint General Secretary**, Green University Reading Society
- **Core Team-member & Contributor** in International AI-based Dhaka Traffic Detection Challenge, Green University of Bangladesh, Dhaka

TECHNICAL SKILLS

- **Languages:** Python, R, Java, Kotlin, C/C++, JavaScript
- **Lib/Frameworks:** PyTorch, TensorFlow, JAX, LangChain, FastAPI, Flask, ONNX, Docker, Kubernetes, Terraform, React
- **Tools:** Google Colab Environment, Kaggle Environment, Android Studio IDE, Pycharm, VS Code
- **Platforms:** Google Cloud Platform, Heroku, AWS

• **Databases:** SQL, Cloud Databases, Vector Databases

HACKATHONS/PROJECT COMPETITIONS

Nasa Space Apps Challenge ’20 Champion — Global Nominee (LINK)
Nasa Space Apps COVID-19 Challenge ’20 Honorable — Global Problem Solver. (LINK)
Facebook Messaging Hackathon ’20 Honourable — Swag. (LINK)
Nasa Space Apps Challenge ’19 Top 40 in Regional — Honorable. (LINK)
COVID-19 Global Hackathon 2.0 Honorable (LINK)

PROBLEM SOLVING ACTIVITIES

- CSE CARNIVAL, GUB 18 & 19 - **Champion**
- ACM International Collegiate Programming Contest ’18 & ’19 - **Honorable**
- National Collegiate Programming Contest ’20, MIST - **Honorable**
- IUPC SUST ’19 - Ranked 27th
- Inter University Programming Contests: BUET ’18, GUB ’18, and RUET ’19 - **Honorable**

VOLUNTARY EXPERIENCE

- **Volunteered Programming Instructor**, Green University Programming Club - Taught classes for 500+ students.
- **Academic Student Mentor**, Green University of Bangladesh - Helped and supported students in academic courses for 200+ students.
- Participated in **Inter University Parliamentary Debate Competition**.

SOFT SKILLS

Leadership, Ability to Work Under Pressure, Decision Making, Time Management, Adaptability, Teamwork, Creativity

LANGUAGES

English, Bengali ●●●●●