Md. Fahim Arefin

Assistant Professor, Computer Science and Engineering, University of Dhaka
fahim@cse.du.ac.bd | +880 172 760 5430 | farefin.com
Google Scholar | ORCID | Dhaka, Bangladesh

Research Profile

I study how students learn to program when AI can write much of their code, and how to tell what they actually understand. I lead development of CoLAS, a collaborative programming platform that links students' plans, code revisions and AI-tutor exchanges, and I proposed CodeStream (Data in Brief), a dataset of programming submissions and verdict histories. In a cognitive diagnostic study I supervised (CSEDM at EDM), large language models graded strictly and missed skills present in the code, while small models gave credit for skills that were absent. Two papers accepted at LAK 2027 examine how reliable LLM skill diagnosis is and how well multimodal LLMs grade handwritten programming exams. I am now studying how peer testing works in AI-assisted labs. I also spent four years on contract model-evaluation work with frontier AI researchers at Turing and Pareto.

Education

M.Sc., Computer Science and Engineering 2020
University of Dhaka
CGPA: 3.98; ranked 1st of 30; Provost Award for Academic Excellence (2019).

B.Sc., Computer Science and Engineering 2018
University of Dhaka
CGPA: 3.89; ranked 1st of 65; Dean's Award (2017); President's Gold Medal (2018).

IELTS Academic (October 2026): overall 8.5 Listening 9.0; Reading 9.0; Speaking 8.5; Writing 7.5.

Research Contributions

- **LAK 2027.** Supervised two papers: **disagreement between models flags handwritten exam answers for human review**, and **LLM-generated skill profiles change across repeated runs**.
- **ACM TOSEM.** Supervised two undergraduate researchers and defined research questions on whether feedback from multiple models improves model performance on competitive-programming problems (A-ProS), in collaboration with Dr. Tarannum Shaila Zaman (UMBC).
- **Findings of EMNLP.** Supervised two undergraduate researchers; suggested evaluation axes and rubrics to score model explanations of textual and visual metaphors in Bangla memes (BanglaMemeX).
- **CSEDM at EDM.** Supervised *Strict Graders and Hallucinated Mastery* (how LLMs grade skills).
- **Data in Brief.** Proposed CodeStream, a dataset of submissions and verdict histories (Mendeley Data); designed its tasks and data collection.
- **CoLAS (May 2026 – Present).** Led development of CoLAS, a programming platform used in multiple labs and piloted at another university; it links each student's plan, code revisions and AI-tutor exchanges.
- **Data mining (2016 – 2019).** First-author research on sequential correlation and contextual item similarity, supervised by Prof. Chowdhury Farhan Ahmed.
- **Current co-supervision.** Four students, with Dr. Md. Mosaddek Khan.

Academic Appointments

University of Dhaka, Computer Science and Engineering
Assistant Professor — September 2026 to present
Lecturer — March 2023 to August 2026
Supervised five undergraduate thesis groups to their first peer-reviewed publications, at venues including ACM TOSEM, Findings of EMNLP and LAK.

Lecturer, Computer Science and Engineering January 2020 to January 2023
Green University of Bangladesh

Industry Experience in Model Evaluation

Consultant, RL Data Production October 2025 to September 2026

Pareto

- Led **reinforcement-learning (RL) data production** for data science and agentic tasks with distributed experts, engineers and vendors. Built review workflows with automated checks and interactive samples for researcher evaluation.
- When evaluating a model that scored close to Astra 6 on Terminal-Bench, noticed it used 6–7 times as many tokens as Astra 6 and took 3 times as long. **Analyzed its execution traces to identify the model's weaknesses**.

Consultant, Model Evaluation and Research Data July 2022 to September 2025

Turing

- Worked with AI researchers to design evaluation datasets and define success metrics, and used model failures to guide fine-tuning data collection. Scaled the programs to **hundreds of trainers** using qualification checks and AI-assisted coaching.
- Designed training datasets for tool-calling agents in 2022, teaching language models to use the Azure SDK and external APIs.
- Identified **buttons that opened blank pages** in computer-use training environments in 2024. Flagged the risk of teaching agents to avoid valid actions, prompting researchers to hide the affected buttons.

Domain Expert 2025 to 2026

Bangladesh Computer Council

Evaluated speech recognition, speech synthesis, and spelling and grammar tools for the national Bangla language platform (bangla.gov.bd).

Funded Research

Consultant, Text-to-SQL Research April 2026 to present

Research project financed by Asian Development Bank / UGC

Conduct Text-to-SQL research for an enterprise analytics platform using retrieval-augmented generation.

Co-Principal Investigator, UGC Research Grant 2024 to 2025

Contributed to the design and final report for the Utility-based Hypergraph Mining project.

Undergraduate Research Assistant, Rupai 2016 to 2017

Developed a Scratch-inspired robotics kit using Raspberry Pi, Arduino, JavaScript and Python; supported by ICT Division (**BDT 3 million** total project funding).

Entrepreneurship & Social Impact

Co-founder and Advisor, AlterYouth July 2019 to present

CTO: July 2019 to February 2023; subsequently Advisor

- Co-founded a platform connecting global scholarship providers with financially struggling students in Bangladesh's government primary schools, now supporting **3,000+ monthly scholarships** worth approximately **US$400,000 annually**.
- Built the web platform, scholarship dashboard, cross-platform user app, and Android teacher app for validating student progress.
- Automated payouts to mothers' bKash accounts and payment reconciliation; over 99% of payouts are disbursed on time.
- Added payment and KYC checks that cut manual reviews by 80% and let users fund a scholarship in under two minutes.

Co-founder and Advisor, DhakaDoctor July 2020 to present

CTO: July 2020 to February 2023; subsequently Advisor

Helped launch online medical consultations for AlterYouth's teachers during COVID-19, and later opened the service to the public. The platform has handled **5,000+ paid appointments**. Hired two junior developers to maintain AlterYouth and DhakaDoctor before moving into advisory roles.

Founder, Foodtong August 2017 to December 2018

Grameenphone Accelerator, 12-week program

Built apps that let students order home-cooked meals from nearby chefs. During the Grameenphone Accelerator, tested pre-orders, subscriptions, on-demand meals and corporate orders with customers. Foodtong closed in December 2018.

Selected Academic Service

2024 Co-coach, University of Dhaka team, 46th ICPC World Finals.

2022, 2024 Problem Setter and Judge Coordinator, Code Samurai.

Selected Honors & Awards

2021 Best Paper Award, Signal Detection and Processing track, IEEE IEMCON 2021.

2017 First Runner-Up, Health Category, EATL-Prothom Alo Apps Contest 2016.

2016 Champion, Best Student Software, Red Hat Code Healthy with OpenShift.

2016 Best Mobile Application, Microsoft VS Marketplace Hackfest, Java Tools Challenge.

Selected Publications

† Papers from thesis research I supervised.

[1] Bedar Baksh Chowdhury, Khondker Bin Yamin, **Md. Fahim Arefin**†. **Automated Rubric-Based Grading of Handwritten Programming Examinations with Multimodal LLMs**. *International Conference on Learning Analytics and Knowledge (LAK27), full paper, accepted, Mar 2027.*

[2] Rubaiya Tarannum Mrittika, Md. Saif Mahamud, Anika Tabassum, Marzan Binte Abid, **Md. Fahim Arefin**†. **Agreement and Run-to-Run Stability of LLM-Generated Programming Skill Profiles: An Exploratory Study**. *International Conference on Learning Analytics and Knowledge (LAK27), short paper, accepted, Mar 2027.*

[3] Anika Tabassum, Md Sifat Hossain, **Md. Fahim Arefin**†, Tariqul Islam, Tarannum Shaila Zaman. **A-ProS: Towards Reliable Autonomous Programming Through Multi-Model Feedback**. *ACM Transactions on Software Engineering and Methodology, Nov 2026.* Paper

[4] Anika Tabassum, Mushahid Intesum, **Md. Fahim Arefin**†, Tarannum Shaila Zaman. **VibeCheck: Assessing the Quality of LLM-Generated Unit Tests: A Multi-agent Empirical Study across Heterogeneous Repositories**. *1st International Workshop on PromptOps and Vibe Coding (POVC), at ASE, pp. 9-16, Oct 2026.* Paper

[5] Md. Sadman Sakib, Zisan Mahmud, **Md. Fahim Arefin**†, Md Fahim. **BanglaMemeX: Advancing Cultural Metaphoric Image Interpretation in Bangla with a Multimodal Explainable Dataset**. *Findings of the Association for Computational Linguistics: EMNLP 2026, Oct 2026.* Paper

[6] Nazira Jesmin Lina, Syed Mumtahin Mahmud, Mahmudul Hasan, **Md. Fahim Arefin**†, Redwan Ahmed Rizvee, Md Mahmudur Rahman, Md Mosaddek Khan. **CodeStream: A Dataset of Iterative Programming Submissions with Sequential Verdict Traces and Attempt Histories**. *Data in Brief, vol. 66, article 112843, Jun 2026.* Paper Dataset

[7] Sartaj Solaiman, **Md. Fahim Arefin**†. **Strict Graders and Hallucinated Mastery: A Cognitive Diagnostic Study of Language Models on Dynamic Programming**. *Computer Science Education Data Mining (CSEDM) Workshop at EDM, CEUR Workshop Proceedings, vol. 4287, paper 6, Jun 2026.* Paper

[8] Md Sifat Hossain, Anika Tabassum, **Md. Fahim Arefin**†, Tarannum Shaila Zaman. **LLM-ProS: Analyzing Large Language Models' Performance in Competitive Problem Solving**. *International Workshop on Large Language Models for Code (LLM4Code), at ICSE, May 2025.* Paper

[9] Redwan Ahmed Rizvee, Chowdhury Farhan Ahmed, **Md. Fahim Arefin**, Carson K. Leung. **A New Tree-Based Approach to Mine Sequential Patterns**. *Expert Systems with Applications, vol. 242, article 122754, May 2024.* Paper

[10] Sartaj Solaiman, B.M. Obaydur Rahman, **Md. Fahim Arefin**†, Chowdhury Farhan Ahmed, Carson K. Leung, Evan W.R. Madill. **Unveiling Market Sentiments: A Comprehensive Analysis of Stock Market Responses to Diverse News Events Using Data Mining Techniques**. *18th International Conference on Ubiquitous Information Management and Communication (IMCOM), pp. 1-8, Jan 2024.* Paper

[11] Redwan Ahmed Rizvee, **Md. Fahim Arefin**, Marzan Binte Abid. **A Robust Objective Focused Algorithm to Detect Source Code Plagiarism**. *IEEE Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), Oct 2022.* Paper

[12] Redwan Ahmed Rizvee, **Md. Fahim Arefin**, MD. Rayhun Khan, MD. Nezarul Islam, Kh. Fazle Rabbi. **An Automated System to Calculate Marks from Answer Scripts**. *IEEE Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 48-54, Oct 2022.* Paper

[13] Sanjida Jabin, Mahbuba Sumia Suhi, **Md. Fahim Arefin**†, Khan Md. Hasib. **Comparison of Different Sentiment Analysis Techniques for Bangla Reviews**. *IEEE Region 10 Humanitarian Technology Conference (R10-HTC), pp. 288-293, Sep 2022.* Paper

[14] **Md. Fahim Arefin**, Chowdhury Farhan Ahmed, Redwan Ahmed Rizvee, Carson K. Leung, Longbing Cao. **Mining Contextual Item Similarity without Concept Hierarchy**. *16th International Conference on Ubiquitous Information Management and Communication (IMCOM), pp. 1-8, Jan 2022.* Paper

[15] Tanzirul Islam, Mofazzal Hossain, **Md. Fahim Arefin**†. **Comparative Analysis of Different Text Summarization Techniques Using Enhanced Tokenization**. *International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1-6, Dec 2021.* Paper

[16] Shadman Sakib, Khan Md. Hasib, Ihtyaz Kader Tasawar, Abyaz Kader Tanzeem, **Md. Fahim Arefin**, Saharul Islam, Mohammad Shafiul Alam. **A Data-Driven Hybrid Optimization Based Deep Network Model for Short-Term Residential Load Forecasting**. *IEEE Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 187-193, Oct 2021.* Paper **Best Paper Award, Signal Detection and Processing track.**

[17] Redwan Ahmed Rizvee, **Md. Fahim Arefin**, Chowdhury Farhan Ahmed. **Tree-Miner: Mining Sequential Patterns from SP-Tree**. *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), May 2020.* Paper

[18] **Md Fahim Arefin**, Maliha Tashfia Islam, Chowdhury Farhan Ahmed. **Mining Sequential Correlation with a New Measure**. *Industrial Conference on Data Mining (ICDM), pp. 29-43, Jul 2018.* Paper