Syed Rifat Raiyan

Phone: +8801721886738 **Website:** starscream-11813.github.io **Email:** rifatraiyan@iut-dhaka.edu
Google Scholar: scholar.google.com/citations?user=4L_7vaoAAAAJ
GitHub: github.com/Starscream-11813 **LinkedIn:** syed-rifat-raiyan **Address:** Uttara, Dhaka, Bangladesh

Research Experience

Multi-Agent Mathematical Reasoning | tinyurl.com/mathmage-proposal — Paper *work-in-progress*
MathMage: ***A Multi-Agent Framework for Enhancing the Mathematical Reasoning Aptitude of Large Language Models*** [M.Sc. Thesis] · Talk: — *August 2025 – present*

- Solver agent that iteratively proposes solutions and refines them via feedback-driven reasoning cycles.
- Multi-agent system to enhance a central solver agent through metacognitive in-context examples, featuring a *Council of Verifiers* for evaluation and a meta-reviewer agent that synthesizes feedback for policy optimization.
- Assessment of the proposed framework on standard math benchmarks and pertinent ablation studies.

Prompt Frugalization | arXiv arxiv.org/abs/2510.16439 — Under Review at *LREC*
FrugalPrompt: ***Reducing Contextual Overhead in Large Language Models via Token Attribution*** — *2024–2025*

- Novel, training-free prompt compression strategy for LLMs that controllably filters low-importance tokens based on saliency scores from pre-trained encoders.
- Experiments on four staple NLP tasks (CLS, SUM, QA, RSN) revealed that a 20% prompt reduction retains the performance across most models and tasks with minimal parameter overhead.
- Key behavioral insights into the performance scaling with the inference cost and possible task contamination.

Physics Reasoning | arXiv arxiv.org/abs/2508.00079 — In *Findings of IJCNLP-AACL*
PhysicsEval: ***Inference-Time Techniques to Improve the Reasoning Proficiency of Large Language Models on Physics Problems*** — *2025*

- Evaluation benchmark comprising 19,609 *physics problems* across 19 *categories* and their elaborated solutions, curated by scraping problems and initial solutions from online educational forums.
- New metric, *Physics Proficiency Score (PPS)*, that quantifies a model's physics reasoning ability based on rubrics used in the *Minnesota Assessment of Problem Solving (MAPS)* framework.
- Employed 4 inference-time techniques and agentic frameworks, including the verification of proposed solutions in a cumulatively by other, smaller LLM agents. 7.5 PPS gain for Phi-4-reasoning on hard problems.

LSB Image Steganography | Springer tinyurl.com/screedsolo-proof-acpr25 — Accepted at *ACPR*
SCReedSolo: ***A Secure and Robust LSB Image Steganography Framework with Randomized Symmetric Encryption and Reed–Solomon Coding*** — *2024–2025*

- Image steganography framework that uniquely combines Random Shuffling for obscurity, Fernet Symmetric Encryption for confidentiality, and Reed–Solomon ECC for data integrity, with an LSB embedding scheme.
- Mathematical analysis of error-correction bounds, multi-metric gains, and immunity to passive steganalysis.

Hand Shadow Puppet Classification | CVF tinyurl.com/hasper-iccv25 — Accepted at *ICCV WCCA*
HaSPeR: ***An Image Repository for Hand Shadow Puppet Recognition*** · Talk: — *2023–2025*

- 15,000 *images* of hand shadow puppets across 15 *classes* sourced from 68 *professional* and 90 *amateur clips*.
- Diversity of poses, orientations, background lighting, and silhouette motion via *optical flow estimation*.
- 31 feature extractor models were employed to establish baselines. We found that *skip-connected convolutional models* supersede *attention-based transformers* in silhouette classification, possibly because skip-connections help preserve low-level edge and contour information through identity mappings.
- Analyses of ResNet34's feature fusions, representations, interpretability, explainability, and CLS errors.
- Lightweight (29 MB size) Android app using *Flutter* for real-time classification (~880 μs latency) of hand shadow puppets from camera feeds, showcasing potential for digitized ombromanie teaching/learning tools.

Math Word Problem Solving | aclanthology.org/2023.acl-srw.49 — Accepted at *ACL-SRW 2023*

Variational Mathematical Reasoning: Enhancing Math Word Problem Solvers with Linguistic Variants and Disentangled Attention [B.Sc. Thesis] · Talk:

- Novel framework for Math Word Problem (MWP) solvers based on the generation of linguistic variants of the problem text and electing the predicted expression with the majority of the votes (+5.4% avg. delta, 5 folds).
- Introduced a challenging dataset, ParaMAWPS, consisting of 16,278 paraphrased, adversarial, and inverse variants of 2,373 seed MWPs from MAWPS.

Bangla Sentiment Analysis | aclanthology.org/2023.findings-acl.80 — In *Findings of ACL 2023*

BanglaBook: ***A Large-scale Bangla Dataset for Sentiment Analysis from Book Reviews*** · Talk:

- 158,065 samples classified into 3 broad categories: *Positive*, *Negative*, and *Neutral*.
- We statistically analyzed the dataset and employed multiple machine learning models to establish baselines.
- We found that pre-trained models substantially trumps models that rely on manually crafted features.

Survey on Math Word Problem and Geometry Problem Solving — Paper *work-in-progress 2022–present*

Solving Math Word Problems and Geometry Problems using Natural Language Processing and Multi-modal Reasoning: A Review of the Recent Approaches

- An analytical, critical, chronological, and comprehensive review of the literature (100+ papers) in the domain of MWP and Geometry Problem Solving, and an outline of our future expectations about this research frontier.

Other Academic Projects

Distributed Optimization for FL | Python/PySpark | tinyurl.com/fairopt — CSE 6273: Cloud Comp.

FairOpt: ***Ensuring Client-agnostic Fairness in Distributed Optimization for Federated Learning*** *2024*

- Implemented a client-agnostic fairness framework for federated learning, namely FairOpt, incorporating both data diversity (variance and interquartile range) and size into the client weighting process.
- Found that fairness-aware distributed linear regression models sometimes yield better best-fit lines than the standard FedAvg and non-distributed versions, promoting equitable model performance across clients.

Food Delivery Route Planning | Python/PDDL/Turtle | tinyurl.com/food-yammy — CSE 6213: AAI

FoodYammy: ***Modeling and Visualizing a Food Delivery Platform using PDDL with Turtle Animation*** *2024*

- Formalized a food delivery platform as a PDDL planning problem, defining states, actions, and goals to support optimization of the Single-Vehicle Pickup and Delivery Problem with time and capacity constraints.
- Created an animated simulation of the delivery process using the Turtle graphics library in Python.

Class Scheduling CSP | Python/MiniZinc | tinyurl.com/routine-minizinc — CSE 6213: AAI

Generating Class Routines with MiniZinc as a Constraint Satisfaction Problem (CSP) *2024*

- Considering a 5-day week, with n time-slots per day, r available rooms, s sections, the students of which attend c courses of h credit hours each, created an $O(n^4)$ solution with a $5 \times s \times c \times h \times n \times r$ search space.

Motion-based Gaming | Python/MediaPipe | tinyurl.com/FitQuestExergames — CSE 4849/6275: HCI/AHCI

FitQuest Exergames: ***Motion-based Gameplay for Simple Sedentary Games*** *2023–2025*

- Re-imagined the gaming experience of 2 simple sedentary games, *Chrome Dino* and *Pinball*, into motion-based forms by sensing the user's motion via an external camera.

MWP Solver | Python/TensorFlow/HuggingFace | tinyurl.com/MathBotModel — CSE 4622: ML Lab

MathBot: ***A Transformer-based Math Word Problem (MWP) Solver*** *2022*

- Implemented a Transformer model that translates an MWP statement to a valid math expression, which when evaluated, yields the *solution* to the problem.

Competitive Programming IDE | Python/Flask/Bootstrap | tinyurl.com/CpZenIDE — CSE 4510: SD Lab

CpZen: ***An Online Integrated Development Environment (IDE) for Competitive Programmers*** *2021*

- Created a Codemirror text editor area with Syntax Highlighting, Auto-Indentation, Auto-Brackets Matching, Auto-Brackets Highlighting, and Line Highlighting. Supports a total of 20 *programming languages*.
- Users can *Compile/Run* codes, *Save* their codes/templates, view a list of *Upcoming Contests* on 12 online judges, view *Profile Statistics*, and keep track of the *Algorithms* they learn throughout their CP journey.

Islamic Productivity App | Javascript/PERN Stack | tinyurl.com/PMuslimApp CSE 4508: RDBMS Lab
PRODUCTIVE MUSLIM: ***A Productivity App for the Adherents of the Islamic Faith*** *2021*

– Users can view a wide variety of *Duas* (supplications) categorized based on emotions, view a list of *Salah Waqts*, use a *Fasting Calendar*, maintain a *To-do List*, choose to participate in a *"30-days, 30-deeds"* challenge, and converse with others in a *Discussion Forum*. **Stack:** PostgreSQL, Express, React, Node.js

3D Modeling of Architectural Landmark | Blender | tinyurl.com/ToriiGate CSE 4552: CGMS Lab
TORII GATE: ***A 3D Blender Model of the Itsukushima Shrine Torii Gate*** *2021*

– Created a day-night timelapse scene featuring the Itsukushima Shrine Torii Gate, a Japanese Shinto shrine.

Offline Programming Judge | Java/Swing/Socket/SQLite | tinyurl.com/IUTForces CSE 4402, CSE 4408
IUTFORCES: ***An Offline Judge Application to Automate the Lab Task Evaluation Process*** *2020*

– Lab Instructors can *create problemsets* as programming lab tasks, automatically *assess* the students' solutions, view *Rank-lists*, view a *Status Table* of the submissions, and view the students' *Submission History*. Students can *submit* their code and view the *verdicts* of their submissions.

Sketchbook Application | C++/Qt | tinyurl.com/InQAppQt CSE 4302: OOP Lab
INQ: ***A Digital Canvas for Painting with a Virtual Palette of Colors and Tools*** *2020*

– Users can adjust *Brush Thickness*, select colors of different shades/hues/saturations from a *Color Palette*, draw *Geometric Shapes*, use a *Floodfill* tool, *Zoom* in/out, *Open/Save* image files, and change their *Resolution*.

ASCII Art-based Console Game | C/SDL | tinyurl.com/IntoThePokeverse CSE 4202: SP-II Lab
POKÉMON - INTO THE POKÉVERSE: ***A Simple Pokémon Game with ASCII Art Characters*** *2019*

– Users play as Pokémon trainers, engage in *Pokémon battles* against opponents of varying *difficulty levels*, purchase *items*, *heal* Pokémons, and *explore* a 2D map of the game world.

PUBLICATIONS

† denotes equal contribution and ‡ denotes supervisory co-authorship.

Peer-Reviewed Works

[1] **S. R. Raiyan**, M. N. Faiyaz, M. J. Kabir, M. Kabir, H. Mahmud, and M. K. Hasan, "Math Word Problem Solving by Generating Linguistic Variants of Problem Statements", in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop)*, Association for Computational Linguistics, Jun. 2023, pp. 362–378. DOI: 10.18653/v1/2023.acl-srw.49.

[2] M. Kabir†, O. B. Mahfuz†, **S. R. Raiyan**†, H. Mahmud, and M. K. Hasan, "BanglaBook: A Large-scale Bangla Dataset for Sentiment Analysis from Book Reviews", in *Findings of the Association for Computational Linguistics: ACL 2023*, Association for Computational Linguistics, Jun. 2023, pp. 1237–1247. DOI: 10.18653/v1/2023.findings-acl.80.

[3] **S. R. Raiyan**, Z. Z. Amio, and S. Ahmed, "HaSPeR: An Image Repository for Hand Shadow Puppet Recognition", in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, Oct. 2025, pp. 4446–4456. DOI: 10.48550/arXiv.2408.10360.

[4] **S. R. Raiyan** and M. H. Kabir, "SCReedSolo: A Secure and Robust LSB Image Steganography Framework with Randomized Symmetric Encryption and Reed–Solomon Coding", in *Proceedings of the 8th Asian Conference on Pattern Recognition*, Lecture Notes in Computer Science, Springer, Nov. 2025. DOI: 10.48550/arXiv.2503.12368.

[5] O. Siddique†, J. Alam†, M. J. R. Rafy†, **S. R. Raiyan**†, H. Mahmud, and M. K. Hasan, "PhysicsEval: Inference-Time Techniques to Improve the Reasoning Proficiency of Large Language Models on Physics Problems", in *Findings of the Association for Computational Linguistics: IJCNLP-AACL 2025*, Association for Computational Linguistics, Dec. 2025. DOI: 10.48550/arXiv.2508.00079.

[6] K. M. T. M. Faruk, M. R. Talha, H. M. K. Ahamad, M. G. Shams, N. M. Hossain, **S. R. Raiyan**‡, M. K. Hasan, H. Mahmud, and R. Islam, "ADAB: A Culturally-Aligned Automated Response Generation Framework for Islamic App Reviews by Integrating ABSA and Hybrid RAG", in *5th Muslims in ML (MusIML) Workshop co-located with NeurIPS 2025*, Accepted; to appear, Dec. 2025.

Preprints & Under Review

[7] **S. R. Raiyan**, M. F. Ishmam, A. A. Imran, and M. A. Moni, "FrugalPrompt: Reducing Contextual Overhead in Large Language Models via Token Attribution", *arXiv:2510.16439*, Oct. 2025, Under review at the *15th Language Resources and Evaluation Conference (LREC 2026)*. DOI: 10.48550/arXiv.2510.16439.

[8] A. A. Mohsin, M. Ahsan, N. Maliyat, S. Maria, **S. R. Raiyan**‡, H. Mahmud, and M. K. Hasan, "BanglaNirTox: A Large-scale Parallel Corpus for Explainable AI in Bengali Text Detoxification", *arXiv:2511.01512*, Oct. 2025, Under review at the *15th Language Resources and Evaluation Conference (LREC 2026)*, ACL ARR October 2025 submission. DOI: 10.48550/arXiv.2511.01512.