Md Tanvirul Alam

Rochester, NY, USA ma8235@rit.edu xashru.github.io Google Scholar GitHub LinkedIn

Research Interests

Reasoning in Large Language Models; Reinforcement Learning with Verifiable Rewards (RLVR); Multimodal (Vision–Language) Learning and Evaluation; Cyber Threat Intelligence and Security; Explainable and Trustworthy AI.

Education

2021–2026 (exp.) **Ph.D. in Computing and Information Science**, *Rochester Institute of Technology (RIT)*, Rochester, NY, USA.

- Advisor: Dr. Nidhi Rastogi
- Research focus: LLM reasoning, RLVR, benchmarking, and cybersecurity applications (CTI).

2011–2016 **B.Sc. in Electrical and Electronic Engineering**, *Bangladesh University of Engineering & Technology (BUET)*, Dhaka, Bangladesh.

- Concentration: Electrical and Electronic Engineering

Publications

Peer-reviewed publications

[P1] **Md Tanvirul Alam**, Dipkamal Bhusal, Le Nguyen, and Nidhi Rastogi. *CTIBench: A Benchmark for Evaluating LLMs in Cyber Threat Intelligence.* In *Advances in Neural Information Processing Systems* (NeurIPS 2024), vol. 37. **Spotlight presentation**.

[P2] **Md Tanvirul Alam**, Justin Yang Chae, and Nidhi Rastogi. *Sphinx: Visual Perception and Reasoning Gym.* In *Multimodal Algorithmic Reasoning (MAR) Workshop at NeurIPS 2025*.

[P3] **Md Tanvirul Alam** and Nidhi Rastogi. *Limits of Generalization in RLVR: Two Case Studies in Mathematical Reasoning.* In *MATH-AI Workshop at NeurIPS 2025*.

[P4] **Md Tanvirul Alam**, Aritran Piplai, and Nidhi Rastogi. *ADAPT: A Pseudo-labeling Approach to Combat Concept Drift in Malware Detection.* In *Proceedings of the 28th International Symposium on Research in Attacks, Intrusions and Defenses* (RAID 2025).

[P5] **Md Tanvirul Alam**, Dipkamal Bhusal, and Nidhi Rastogi. *R+R: Revisiting Static Feature-Based Android Malware Detection using Machine Learning.* In *Annual Computer Security Applications Conference* (ACSAC 2025).

[P6] **Md Tanvirul Alam**, Dipkamal Bhusal, et al. *AthenaBench: A Dynamic Benchmark for Evaluating LLMs in Cyber Threat Intelligence.* In *WAITI 2025*.

[P7] **Md Tanvirul Alam**, Dipkamal Bhusal, et al. *SECURE: Benchmarking Generative Large Language Models for Cybersecurity Advisory.* In *Annual Computer Security Applications Conference* (ACSAC 2024).

[P8] **Md Tanvirul Alam**, Dipkamal Bhusal, Youngja Park, and Nidhi Rastogi. *Looking beyond IoCs: Automatically Extracting Attack Patterns from External CTI.* In *Proceedings of the 26th International Symposium on Research in Attacks, Intrusions and Defenses* (RAID 2023), pp. 92–108.

[P9] Le Nguyen, Preet Jain, Krutik Panchal, **Md Tanvirul Alam**, and Nidhi Rastogi. *Assessing Effective Token Length of Multimodal Models for Text-to-Image Retrieval.* In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR 2025).

[P10] Romy Fieblinger, **Md Tanvirul Alam**, and Nidhi Rastogi. *Actionable Cyber Threat Intelligence Using Knowledge Graphs and Large Language Models.* In *2024 IEEE European Symposium on Security and Privacy Workshops* (EuroS&PW).

[P11] Dipkamal Bhusal, **Md Tanvirul Alam**, Monish Kumar Manikya Veerabhadran, Michael Clifford, Sara Rampazzi, and Nidhi Rastogi. *PASA: Attack Agnostic Unsupervised Adversarial Detection using Prediction & Attribution Sensitivity Analysis.* In *9th IEEE European Symposium on Security and Privacy* (EuroS&P 2024).

[P12] Justin Yang Chae, **Md Tanvirul Alam**, and Nidhi Rastogi. *Towards Understanding Self-play for LLM Reasoning.* In *MATH-AI Workshop at NeurIPS 2025*.

[P13] **Tanvirul Alam**, Akib Khan, and Firoj Alam. *Punctuation Restoration using Transformer Models for High- and Low-Resource Languages.* In *Proceedings of the 6th Workshop on Noisy User-generated Text (W-NUT 2020) @ EMNLP*.

[P14] Firoj Alam, Ferda Ofli, Muhammad Imran, **Tanvirul Alam**, and Umair Qazi. *Deep Learning Benchmarks and Datasets for Social Media Image Classification for Disaster Response.* In *International Conference on Advances in Social Networks Analysis and Mining* (ASONAM 2020).

[P15] Firoj Alam, **Tanvirul Alam**, Md. Arid Hasan, Abul Hasnat, Muhammad Imran, and Ferda Ofli. *MEDIC: A Multi-task Learning Dataset for Disaster Image Classification. Neural Computing and Applications*.

[P16] **Tanvirul Alam** and Akib Khan. *Lightweight CNN for Robust Voice Activity Detection.* In *International Conference on Speech and Computer* (SPECOM 2020), pp. 1–12.

Preprints and under-review manuscripts

[U1] **Md Tanvirul Alam**, Saksham Aggarwal, Justin Yang Chae, and Nidhi Rastogi. *SPHINX: A Synthetic Environment for Visual Perception and Reasoning.* arXiv preprint, 2025. *Under review. Extended version of the NeurIPS 2025 MAR workshop paper.*

[U2] **Md Tanvirul Alam**, Dipkamal Bhusal, Youngja Park, and Nidhi Rastogi. *Cyner: A Python Library for Cybersecurity Named Entity Recognition.* arXiv preprint.

Research & Industry Experience

Aug. 2021–Present **Graduate Research Assistant**, *Rochester Institute of Technology*.
- Conduct research at the intersection of machine learning and cybersecurity, focusing on LLM-based cyber threat intelligence and reasoning.
- Benchmark large language models for CTI using dynamic, realistic evaluation suites and investigate their limitations and failure modes.
- Develop reinforcement learning with verifiable rewards (RLVR) methods to enhance LLM reasoning and decision-making in CTI and related domains.
- Design synthetic visual perception and reasoning benchmarks to evaluate and improve multimodal (vision–language) reasoning.
- Design and use CTI-oriented knowledge graphs to support actionable threat intelligence with LLMs.
- Previously worked on distribution-shift robustness for security applications, including malware and intrusion detection.

Jul. 2025–Jan. 2026 **Research Intern**, *Athena Security Group*.
- Designed and implemented a dynamic large language model (LLM) benchmark for cyber threat intelligence (CTI) using authoritative cybersecurity data sources.
- Developed reinforcement learning and supervised fine-tuning pipelines to improve LLM reasoning and analytic performance on complex CTI tasks.

May 2018–Jun. 2021 **Senior Software Engineer**, *BJIT Limited*.
- Led the development of multiple machine learning projects from conception to deployment across diverse application domains.
- Built a facial attribute recognition system from video streams to drive real-time 3D avatar rendering.
- Developed ML-based solutions for pedestrian safety near level crossings, voice activity detection, environmental sound classification, webpage text classification, and pedestrian attribute detection.

Dec. 2016–Apr. 2018 **Software Engineer**, *Semion Limited*.
- Researched and deployed computer vision models for medical image analysis, including abnormality detection from X-ray images on edge devices.
- Worked on natural language processing for sentiment analysis, question answering, and interpretable text classification with neural networks.

Honors & Awards

2024 Spotlight presentation at NeurIPS 2024 for *CTIBench: A Benchmark for Evaluating LLMs in Cyber Threat Intelligence*.

2024 Best Poster Award, NDSS poster session for *MORPH: Towards Automated Concept Drift Adaptation for Malware Detection*.

2023 IEEE Symposium on Security and Privacy (S&P) Conference Travel Grant.

2020 Employee of the Year, Fintech & AI Department, BJIT Limited.

2016 Advanced to Round 2 (top ∼7% globally) in Google Code Jam 2016.

2011 Admission Test Excellence Scholarship, BUET.

2008, 2010 Education Board Scholarships for academic excellence, Government of Bangladesh.

Teaching & Mentoring

2025 **Co-instructor, Explainable AI (graduate course)**, *Rochester Institute of Technology*.
Designed and delivered lectures, assignments, and projects on interpretable machine learning and explainable AI methods.

2025 **Research Mentor, NSF REU Program**, *Rochester Institute of Technology*.
Mentored undergraduate students on research projects in machine learning and cybersecurity, from problem formulation to experimental evaluation.

2024 **Research Mentor, Master's Capstone Projects**, *Rochester Institute of Technology*.
Supervised master's students on capstone projects in cybersecurity and applied machine learning, including project design, implementation, and reporting.

Service

Reviewer for *NeurIPS*, *ICLR*, and multiple workshops in mathematical reasoning, multimodal algorithmic reasoning, cyber threat intelligence, and NLP.

Skills

Programming Python (primary), C/C++, Java, C#.

ML / DL PyTorch, TensorFlow, Keras, scikit-learn, Hugging Face.

LLMs & RL LLMs and LVLMs, Reinforcement Learning with Verifiable Rewards (RLVR), RLHF/RLAIF (PPO/GRPO-style methods), off-policy and curriculum RL, self-play, synthetic dataset and benchmark design, evaluation and benchmarking of LLMs.

CTI Cyber threat intelligence (CTI) frameworks (CVE, CWE, ATT&CK, CAPEC), knowledge-graph style threat representations.

Other Semi-supervised learning, pseudo-labeling, knowledge distillation, concept drift and distribution-shift robustness, test-time adaptation, explainable AI.