RAKIB AHSAN

iamrakib242@gmail.com rab2807.github.io rakib-ahsan-75b172309 rab2807

RESEARCH INTERESTS

Trustworthy AI | Adversarial Prompting | Computer Vision

WORK EXPERIENCE

Lecturer, Department of CSE, United International University

April 2025 – Present

Courses Teaching : Data Structures and Algorithms | Discrete Mathematics | Operating Systems | Operating Systems Lab | Computer Networks Lab

EDUCATION

Bachelor of Science in Computer Science & Engineering

Bangladesh University of Engineering & Technology

February 2020 – March 2025

CGPA : 3.90/4.00 (ranked **26th (top 20%)** among 137 students)

Notable Courses : Computer Security | Machine Learning | Linear Algebra | Operating Systems | Compiler | Computer Networks | Artificial Intelligence | Database

RESEARCH EXPERIENCE

SequentialBreak: Large Language Models Can be Fooled by Embedding Jailbreak Prompts into Sequential Prompt Chains

Trustworthy AI | Large Language Models — 2024 - 2025

- I co-developed a novel blackbox jailbreak attack showing that LLMs tend to overlook harmful context while handling a series of prompts in a single query. While existing jailbreaks often require multiple rounds for success, our attack achieves a high success rate (90% in JBB-Behaviors dataset) against recent LLMs with a single, simple query.
  Our strategy was to embed a harmful prompt within a series of benign ones, causing the model's attention to be diluted across prompts and reducing its focus on the harmful prompt. We showed three different example attack scenarios, where sequential prompt structure can be used to elicit harmful responses. This research was done in collaboration with Penn State University.

Accepted in ACL Student Research Workshop (SRW) 2025 [DOI] [Slides]

A Survey of Attacks and Defenses in LLM Agents and Vision-Language Models

Trustworthy AI | Vision Language Models | Agentic AI — 2025 - Ongoing

- As a follow-up to my previous collaboration, this is an ongoing project on a comprehensive survey examining the landscape of attacks and defenses targeting LLM agents and Vision-Language Models (VLMs). In this study, I aim to explore questions such as: *What additional attack surfaces do multimodal models and AI agents introduce beyond those of unimodal LLMs?* and *To what extent and how can existing attacks and defenses for unimodal LLMs be adapted to multimodal models and agents?*

AttMetNet: Attention-Enhanced Deep Neural Network for Methane Plume Detection in Sentinel-2 Satellite Imagery

Computer Vision | Remote Sensing — 2024 - 2025

- Recognised as *Highly Commended* (Top 10%) in **Global Undergraduate Awards 2025**.
- This is my undergraduate thesis project under Dr. Tanzima Hashem. This is my undergraduate thesis project under Dr. Tanzima Hashem. Methane is the second-largest contributor to greenhouse gas emissions, and its reduction offers a rapid path to limiting global temperature rise. This work focuses on generating segmentation masks of methane plumes from multispectral satellite imagery which is challenging due to limited data and complex plume morphology. We enhanced a U-Net architecture with attention gates and incorporated spectral indices which outperforms the stateof-the-art detection methods in real-world conditions.

Submitted for review in ICLR 2026

RELEVANT COMPETITIONS

IEEE SPS Signal Processing Cup 2025 | [site link]

First Runner-up [arxiv]

- This work focuses on deepfake detection. For binary classification tasks, clearly separating features helps models generalize better to hard-to-classify cases. Our methodology employs MaxViT, CoAtNet, and EVA-02 as backbones, and fine-tuned using supervised contrastive loss. The final prediction is obtained through majority voting on the classification heads of the individual backbones. SupCon loss produces well-separated embedding spaces, while ensembling captures the perspectives from multiple models.

Global Undergraduate Awards | Site Link

Listed as Highly Commended (Top 10%) [link]

- Our undergraduate thesis was recognized as **Highly Commended** in the Computer Science category of the Global Undergraduate Awards (GUA). This distinction is awarded to entries that rank among the **top 10%** of submissions all over the world.

NOTABLE ACADEMIC PROJECTS

Coastline Segmentation with Satellite Imagery | [github] Pytorch SentinelHub OpenCV Earth Engine

- A deep learning semantic segmentation task to differentiate coastline land and water area. 12-channel Sentinel-2 images were pre-processed and fed to different segmentation models for comparison. Unet, Unet with attention and transformer-based segmentation models were becnhmarked.

Deep Neural Network from Scratch | [github] Numpy Seaborn Matplotlib

- A complete neural network framework implemented from scratch using only NumPy without using high-level frameworks like TensorFlow or PyTorch. Implemented forward/backward propagation, activation functions (ReLU, Sigmoid, SoftMax), loss functions (Cross-entropy loss), Optimizers (SGD and Adam), dropout, batch normalization and early stopping.

Penetration Testing with Metasploit | [github] VirtualBox Metasploitable

- Explored multiple Metasploit modules to simulate attacks against vulnerable virtual machines, including FTP on Metasploitable 2, EternalBlue on Windows 7, and an SMB exploit on Windows XP. Based on the findings, a detailed project report was generated.

TCP Adaptive Reno with NS-3 | [github] NS-3

- With NS-3, TCP Adaptive Reno algorithm is implemented that adaptively adjusts its behavior based on network conditions. TCP Adaptive Reno extends TCP Westwood+ with RTT-based congestion estimation and dynamic window sizing. A dumbbell topology was used for comparing TCP algorithms.

TECHNICAL SKILLS

- **Programming Languages:** C/C++, C#, Java, Python, Javascript, x86 Assembly, SQL, Bash
- **Machine Learning** Pytorch, Numpy, Pandas, Gymnasium, Matplotlib, Scikit-learn, Seaborn, OpenCV
- **Tools & Softwares:** Git, Shell, LaTeX, Cisco Packet Tracer, Unity Engine, WireShark, Metasploit
- **Frameworks:** NodeJS, ReactJS, Bootstrap, Tailwind CSS, P5.js

ACADEMIC ACHIEVEMENTS

Dean's List Scholarship

- Recipient of Dean's List Scholarship in every semester for academic excellence.

Merit Scholarship

- Recipient of Merit Scholarship in January 2021, July 2021, January 2022 and July 2023 sessions for achieving top academic result.

LEADERSHIP

BUET CSE Fest 2024 Game Jam

Organizer

- Co-organized the Game Jam event.
- Took an workshop on Game Development using Unity for the same event.